Activity Report Grenoble - Rhône-Alpes 2018

Edition: 2019-03-11

List of Inria's Research Teams

Project-Team AGORA

ALGorithmes et Optimisation pour Réseaux Autonomes

IN COLLABORATION WITH: Centre of Innovation in Telecommunications and Integration of services

IN PARTNERSHIP WITH:
Institut national des sciences appliquées de Lyon

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Networks and Telecommunications

Table of contents

Project-Team AGORA

Creation of the Team: 2017 January 01, updated into Project-Team: 2018 April 01

Keywords:

Computer Science and Digital Science:
A1.2.1. - Dynamic reconfiguration
A1.2.3. - Routing
A1.2.4. - QoS, performance evaluation
A1.2.5. - Internet of things
A1.2.6. - Sensor networks
A7.1. - Algorithms
A8.2. - Optimization

Other Research Topics and Application Domains:
B3.4.3. - Pollution
B6.2.2. - Radio technology
B6.2.4. - Optic technology
B6.4. - Internet of things
B8.1.2. - Sensor networks for smart buildings
B8.2. - Connected city

1. Team, Visitors, External Collaborators

Research Scientist
Pascale Vicat - Blanc [Inria, Senior Researcher, from Sep 2018, HDR]

Faculty Members
Hervé Rivano [Team leader, INSA Lyon, Professor, HDR]
Walid Bechkit [INSA Lyon, Associate Professor]
Oana Iova [INSA Lyon, Associate Professor]
Razvan Stanica [INSA Lyon, Associate Professor]
Fabrice Valois [INSA Lyon, Professor, HDR]

External Collaborator
Khaled Boussetta [Univ Paris-Nord, HDR]

Technical Staff
Manoel Dahan [Univ de Lyon]
Romain Pujol [INSA Lyon, until Nov 2018]

PhD Students
Yosra Bahri Zguira [University of Sousse, until Dec 2018]
Amjed Belkhiri [INSA Lyon, until Nov 2018]
Abderrahman Ben Khalifa [INSA Lyon]
Ahmed Boubrima [Inria]
Rodrigue Domga Komguem [University of Yaoundé I]
Alexis Duque [Rtone, until Oct 2018]
Mohamed Anis Fekih [Univ de Lyon, from Nov 2018]
Abdoul Aziz Mbacke [Inria, until Oct 2018]
Jad Oueis [INSA Lyon, until Nov 2018]

Romain Pujol [INSA Lyon, from Nov 2018]
Solohaja Rabenjamina [Inria, from Oct 2018]

Post-Doctoral Fellows

Panagiota Katsikouli [Inria]
Junaid Khan [INSA Lyon, until Mar 2018]
Abdoul Aziz Mbacke [INSA Lyon, from Nov 2018]
Elli Zavou [Inria, until Jun 2018]

Visiting Scientist

Abdessamed Sassi [Université Larbi Ben M'hidi – Oum El Bouaghi]

Administrative Assistants

Laetitia Gauthe [Inria, from Dec 2018]
Sophie Karsai [Inria, until Nov 2018]

2. Overall Objectives

2.1. Overall Objectives

The Agora team is positioned in Inria research domain of "Networks, Systems and Services, Distributed Computing" under the theme "Networks and Telecommunications", as a joint team between Inria and INSA Lyon, within the CITI laboratory. The Agora team focus is on the wireless access part of the Internet, where several network architectures and paradigms co-exist: heterogeneous cellular networks, multi-hop wireless networks, long-range low-power connectivity. We work on the deployment of these networking technologies and their combined exploitation, while understanding the spatio-temporal dynamics of users, machines and data.

The deployment of dense networks is challenged by large scale and dense scenarios, with consequences on the optimization of the placement of both the components and functionalities of the network. At the same time, Machine-to-Machine (M2M) communication protocols, designed for running on the Internet of Things (IoT) architectures, need a coherent rethinking to face issues on both saturated cellular networks and fresh deployments of multi-hop wireless networks unable to cover large areas. Exploiting the data carried by the network opens new questions on the network deployment and functioning, by understanding the spatio-temporal dynamics of the users or connected objects.

The main networking fact that motivates the direction of the Agora team is the coming saturation of cellular networks. Even though developed cities can rely on a full coverage of their territory with very high throughput wireless access networks, the expected - and already measured - tremendous growth of mobile data traffic will overwhelm these infrastructures without a dramatic change of communication paradigm.

Beyond cellular networks. The networking functions are today almost only managed through cellular infrastructures. Even current smart-metering network architectures follow a hierarchical organization alike cellular networks. This approach features a number of advantages, including pervasive geographical coverage, seamless connectivity, a good level of security and possibly guaranteed bandwidth and latency. However, this centralized paradigm is over twenty years old now. The access network capacity has already reached its limit, and the explosion of popular, bandwidth-hungry digital services will make the newest technologies, such as LTE, already unable to accommodate the future demand - expected to grow 11-fold in 2018, with respect to 2014. A clear trend is to decentralize the network operation, leveraging network function virtualization so as to make it more pervasive (Small Cells), heterogeneous (HetNets) and self-organizing (SON). Beyond cellular networks, multi-hop wireless mobile networks have been extensively studied in the literature, in particular wireless sensor networks, ad hoc networks, wireless mesh networks and vehicular networks. Such wireless multi-hop solutions met scarce practical success over the last decade, mainly because of the lack of a clear application context and of important use cases. There are however now mature technologies for some specific applications and provide a wealth of connectivity surrounding mobile devices. Combined with the emergence of long range low power technologies dedicated to small traffic IoT applications, one can foresee the emergence of hybrid networks architectures (cellular and multi-hop) that need to be developed and evaluated.

Low cost sensors and density. We also witness the emergence of a new market of sensing devices that is closely related to the industrial effort toward the IoT. Recent breakthrough in micro and nano technologies are indeed enabling dense deployments of low-cost sensing devices that produce reliable enough measurements of physical phenomenon while being energetically autonomous. Density is however challenging network infrastructures deployment and data collection. The deployment of such devices has to be suitable for the application and fitted to the constraints of the environment. Self-organization and self-healing are required for sustainable infrastructure management and operation. Combining all these notions into optimization models is an issue that needs to be addressed to understand and evaluate the relevant networking infrastructures and protocols. On the other hand, density is also an opportunity if one can understand and take advantage of the spatio-temporal characteristics of the data produced and the citizens behavior. Redundancy and correlations are a way to improve on data reliability and network usage.

3. Research Program

3.1. Wireless network deployment

The deployment of networks has fostered a constant research effort for decades, continuously renewed by the evolution of networking technologies. Fundamentally, the deployment problem addresses the trade off between the cost of the network to be minimized or fitted into a budget and the features and services provided by the system, that should reach a target level or be maximized. The variety of cost models and type of features gives rise to a wide scientific field. There are several cost factors of network infrastructure: components (number and capacity), energy, man power (installation and maintenance), etc. The features of the network matter as much as the metric to evaluate them. Coverage and capacity are basic features for wireless networks on which we will focus in the following. One recurrent question is therefore: What is the optimal number and position of network components to deploy so that a given territory is covered and enough networking capacity is provided?

Traditional telecommunication infrastructures were made of dedicated components, each of them providing a given set of functions. However, recently introduced paradigms yield issues on the deployment of network functions. Indeed, the last decade saw a trend towards adding more intelligence within the network. In the case of the access network, the concept of Cloud Radio Access Network (C-RAN) emerged. In the backhaul, the Evolved Packet Core (EPC) network can also benefit from virtualization techniques, as the convergence point for multiple access technologies, as imagined in the case of future 5G networks. The performance limits of a virtualized EPC remain unknown today: Is the delay introduced by this new architecture compatible with the requirements of the mobile applications? How to deploy the different network functions on generic hardware in order to maximize the quality of service?

Network component deployment. In this research direction, we address new issues of the optimal network deployment. In particular, we focus on the deployment of wireless sensor networks for environmental monitoring (e.g. atmospheric pollution). Most of current air quality monitoring systems are using conventional measuring stations, equipped with multiple lab quality sensors. These systems are however massive, inflexible and expensive. An alternative – or complementary – solution is to use low-cost flexible wireless sensor networks. One of the main challenges is to introduce adequate models for the coverage of the phenomenon. Most of the state of the art consider a generic coverage formulation based on detection ranges which are not adapted to environmental sensing. For example, pollution propagation models should take into account the inherently stochastic weather conditions. An issue is to develop adequate formulation and efficient integer linear programming (ILP) models and heuristics able to compute deployments at a relevant scale. In particular, it seems promising to adapt stochastic or robust optimization results of the operational research community in order to deals with uncertainty. Defining the quality of a coverage is also a modeling issue, which depends on the application considered. The detection of anomaly is close to a combinatorial problem. A more difficult objective is to deploy sensors in order to map the phenomenon by interpolation (or other reconstruction mechanisms). This challenge requires interdisciplinary research with fluid mechanics teams who develop numerical models of pollution propagation and practitioners like Air RhoneAlpes.

Regarding the network connectivity, another challenge is to integrate suitable wireless link models accounting for the deployment environment. For example, modeling the integration of sensors in urban areas is challenging due to the presence of neighboring walls and obstacles, as well as moving vehicles and pedestrians that may induce field scattering. Also, the urban constraints and characteristics need to be carefully modeled and considered. Indeed, the urban environment yields constraints or facilities on the deployment of sensor nodes and gateways, such as their embedding within street furniture. Understanding the structure of these spatial constraints is necessary to develop efficient optimization methods able to compute on large scale scenarios.

Network function deployment. In this research direction, we do not address network virtualization per-se, but the algorithmic and architectural challenges that virtualization brings in both radio access and core networks. As a first challenge, we focus on the evaluation of Cloud Radio Access Network solutions. The capacity of a C-RAN architecture and the way this compares to classical RAN is still an open question. The fact that C-RAN enables cooperation between the remote radio heads (RRH) served by the same base-band units (BBU) indicates an improved performance, but at the same time the resulting cells are much larger, which goes against the current trend of increasing capacity through the deployment of small cells. We propose to study the problem both from a user and a network perspective. On the user side, we use standard information theory tools, such as multiple-access channels to model C-RAN scenarios and understand their performance. On the network side, this translates in a resource allocation problem with cooperative base stations. We will extend our previous models for non-cooperative scenarios. Regarding the core network function deployment, we are interested in the specific case of Professional Mobile Radio (PMR) networks. These networks, used for public safety services and in scenarios like post-disaster relief, present the particularity of an EPC formed by a mobile wireless network. Due to its nature, the network can not be pre-planned, and the different EPC functions need to be autonomously deployed on the available network elements. We study the EPC function deployment problem as an optimization problem, constrained by the user capacity requests. User attachment mechanisms will also be proposed, adapted to the network function distribution, the global user demand, and the source/destination of the flows. These challenges are tackled as centralized optimization problems, then extended to the context of real-time decisions. Finally, in order to complete these theoretical work based on ILP models and heuristics, experiments using OpenAir Interface are used to evaluate our proposals.

3.2. Wireless data collection

With an anticipated 11-fold growth between 2014 and 2018, facing the growth of the mobile demand is the foremost challenge for mobile operators. In particular, a 100-fold increase in the number of supported connected devices, mostly newly connected objects with M2M trafic, is expected. A question therefore rises: how to cope with a dense set of M2M low bit rate traffics from energy and computing power constrained devices while classic cellular infrastructure are designed for the sparse high bit rate traffics from powerful devices?

A technological answer to the densification challenge is also embodied by long-range low-power networks such as SigFox, LoRa, NB-IoT, etc. In this context, the idea of offloading cellular traffic to different wireless access technologies is emerging as a very promising solution to relieve the traditional mobile network from its overwhelming load. In fact, offloading is already employed today, and, globally, 45% of total mobile data traffic was offloaded onto the fixed network through Wi-Fi or femtocells in 2013. Device-to-device (D2D) communications in hybrid networks, combining long-range cellular links and short-range technologies, opens even more possibilities. We aim at providing solutions that are missing for efficiently and practically mix multi-hop and cellular networks technologies.

Cellular M2M. Enabling a communication in a cellular network follows two major procedures: a resource allocation demand is first transmitted by the UE which, if successful, is followed by the actual data transmission phase, using dedicated resources allocated by the eNodeB (eNB) to the UE. This procedure was designed specifically for H2H traffic, which is bursty by nature, and it is based on the notions of session and call, activities that keep the user involved for a relatively long time and necessitate the exchange of a series of messages with the network. On the contrary, M2M traffic generates low amounts of data periodically or sporadically. Going through a signaling-heavy random access (RA) procedure to transmit one short message is strongly inefficient for both the M2M devices and the infrastructure.

In the perspective of 5G solutions, we are investigating mechanisms that regulate the M2M traffics in order to obtain good performances while keeping a reasonable quality of service (QoS) for human-to-human (H2H) terminals. The idea of piggybacking the M2M data transmission within one of the RA procedure messages is tempting and it is now considered as the best solution for this type of traffic. This means that the M2M data is transmitted on the shared resources of the RACH, and raises questions regarding the capacity of the RACH, which was not designed for these purposes. In this regard, our analysis of the access capacity of LTE-A RACH procedure has to be adapted to multi-class scenarios, in order to understand the competition between M2M and H2H devices. Modeling based on Markov chains provides trends on system scale performances, while event-based simulations enable the analysis of the distribution of the performances over the different kinds of users. Combining both should give enough insights so as to design relevant regulation techniques and strategies. In particular two open questions that have to be tackled can be stated as: When should access resources be opened to M2M traffics without penalizing H2H performances? Does an eNodeB have a detailed enough knowledge of the system and transmit enough information to UE to regulate the traffics? The objective is to go to the analysis of achievable performances to actual protocols that take into account realistic M2M traffic patterns.

Hybrid networks. The first objective in this research axis is a realistic large-scale performance evaluation of Wi-Fi offloading solutions. While the mechanisms behind Wi-Fi offloading are now clear in the research community, their performance has only been tested in small-scale field tests, covering either small geographical areas (i.e. a few cellular base stations) and/or a small number of specific users (e.g. vehicular users). Instead, we evaluate the offloading performance at a city scale, building on real mobile network traces available in the team. First of all, through our collaboration with Orange Labs, we have access to an accurate characterization of the mobile traffic load at each base station in all major French cities. Second, a data collection application for Android devices has been developed in the team and used by hundreds of users in the Lyon metropolitan area. This application monitors and logs all the Wi-Fi access points in the coverage range of the smartphone, allowing us to build a map of Wi-Fi accessibility in some parts of the city. Combining these two data sources and completing them with simulation studies will allow an accurate evaluation of Wi-Fi offloading solutions over a large area.

On the D2D side, our focus is on the connected objects scenario, where we study the integration of short-range links and long-range technologies such as LTE, SigFox or LoRa. This requires the design of network protocols to discover and group the devices in a certain region. For this, we build on our expertise on clustering sensor and vehicular nodes. The important difference in this case is that the cellular network can assist the clustering formation process. The next step is represented by the selection of the devices that will be using the long-range links on behalf of the entire cluster. With respect to classical cluster head selection problems in ad-hoc networks, our problem distinguishes through device heterogeneity in terms of available communication technologies (not all devices have a long-range connection, or its quality is poor), energy resources (some devices might have energy harvesting capabilities) and expected lifetime. We will evaluate the proposed mechanisms both analytically (clustering problems are generally modeled by dominating set problems in graph theory) and through discrete-event simulation. Prototyping and experimental evaluation in cooperation with our industrial partners is also foreseen in this case.

3.3. Network data exploitation

Mobile devices are continuously interacting with the network infrastructure, and the associated geo-referenced events can be easily logged by the operators, for different purposes, including billing and resource management. This leads to the implicit possibility of monitoring a large percentage of the whole population with minimal cost: no other technology provides today an equivalent coverage. On the networking side, the exploration of data collected within the cellular network can be the enabler of flexible and reconfigurable cellular systems. In order to enable this vision, algorithmic solutions are needed that drive, in concert with the variations in the mobile demand, the establishment, modification, release and relocation of any type of resources in the network. This raises, in turn, the fundamental problem of understanding the mobile demand, and linking it to the resource management processes. More precisely, we contribute to answer questions about the correlation between urban areas and mobile traffic usage, in particular the spatial and temporal causalities in the usage of the mobile network.

In a different type of architecture, the one of wireless sensor networks, the spatio-temporal characteristics of the data that are transported can also be leveraged to improve on the networking performances, e.g. capacity and energy consumption. In several applications (e.g. temperature monitoring, intrusion detection), wireless sensor nodes are prone to transmit redundant or correlated information. This wastes the bandwidth and accelerates the battery depletion. Energy and network capacity savings can be obtained by leveraging spatial and temporal correlation in packet aggregation. Packet transmissions can be reduced with an overhead induced by distributed aggregation algorithms. We aim at designing data aggregation functions that preserve data accuracy and maximize the network lifetime with low assumptions on the network topology and the application.

Mobile data analysis. In this research axis, we delve deeper in the analysis of mobile traffic. In this sense, temporal and spatial usage profiles can be built, by including in our analysis datasets providing service-level usage information. Indeed, previous studies have been generally using call detail records (CDR) or, at best, aggregated packet traffic information. This data is already very useful in many research fields, but fine-grained usage data would allow an even better understanding of the spatiotemporal characteristics of mobile traffic. To achieve this, we exploit datasets made available by Orange Labs, providing information about the network usage for several different mobile services (web, streaming, download, mail, etc.).

To obtain even richer information, we combine this operator-side data with user-side data, collected by a crowdsensing application we developed within the PrivaMov research project. While covering hundreds of thousands of users, operator data only allows to localize the user at the cell level, and only when the user is connected to the network. The crowdsensing application we are using gathers precise GPS user localization data at a high frequency. Combining these two sources of data will allow us to gain insight in possible biases introduced by operator-side data and to infer microscopic properties which, correctly modeled, can be extended to the entire user population, even those for which we do not possess crowdsensed data.

Privacy preservation is an important topic in the field of mobile data analysis. Mobile traffic data anonymization techniques are currently proposed, mainly by adding noise or removing information from the original dataset. While we do not plan to develop anonymization algorithms, we collaborate with teams working on this topic (e.g. Inria Privatics) in order to assess the impact of anonymization techniques on the spatio-temporal properties of mobile traffic data. Through a statistical analysis of both anonymized and non-anonymized data, we hope to better understand the usability of anonymized data for different applications based on the exploration of mobile traffic data.

Data aggregation. Data-aggregation takes benefit from spatial and/or temporal correlation, while preserving the data accuracy. Such correlation comes from the physical phenomenon which is observed. Temporal aggregation is mainly addressed using temporal series (e.g. ARMA) whereas spatial aggregation is now leading by compressive sensing solutions. Our objective is to get rid of the assumption of the knowledge of the network topology properties and the data traffic generated by the application, in particular for dense and massive wireless networks. Note that we focus on data-aggregation with a networking perspective, not with the background of information theory.

The rational design of an aggregation scheme implies understanding data dynamics (statistical characteristics, information representation), algorithmic optimization (aggregator location, minimizing the number of aggregators toward energy efficiency), and network dynamics (routing, medium sharing policies, node activity). We look for designing a complete aggregation chain including both intra-sensor aggregation and inter-sensor aggregation. For this, we characterize the raw data that are collected in order to understand the dynamic behind several key applications. The goal is to provide a taxonomy of the applications according to the data properties in terms of stationarity, dynamic, etc. Then, we aim to design temporal aggregation functions without knowledge of the network topology and without assumptions about the application data. Such functions should be able to self-adapt to the environment evolution. A related issue is the deployment of aggregators into the wireless network to allow spatial aggregation with respect to the energy consumption minimization, capacity saving maximization and distributed algorithm complexity. We therefore look to define dedicated protocols for each aggregation function family.

4. Application Domains

4.1. Smart Cities

One major characteristic of modern societies is that they are prevalently urban. In coherence, the contributions of the Agora team are in particular applied to provide solutions tailored to the emergence of the Internet of Things (IoT) and to Smart Cities applications. A major motivation of the team is the forthcoming explosion of the number of connected devices. In particular, low cost - small data devices are supposed to be densely deployed in our environment, fostering the interest for a convergence of the traditional wireless networking paradigms.

Smart City is a constantly reshaped concept, embracing the future of dense metropolitan areas, with references to efficient and sustainable infrastructure, improving citizens' quality of life and protecting the environment. A consensus on the Smart City philosophy is however that it will be primarily achieved by leveraging a clever integration of Information and Communication Technologies (ICT) in the urban tissue. Indeed, ICTs are enabling an evolution from the current duality between the real world and its digitized counterpart to a continuum in which digital contents and applications are seamlessly interacting with classical infrastructures and services. Smart Cities are often described by the digital services that should be provided which are inherently dependent on dense measurements of the city environment and activities, the collection of these data, their processing into information, and their redistribution. The networking infrastructure plays therefore a critical role in enabling advanced services, in particular the wireless infrastructure supporting density and mobility.

From a wireless networking viewpoint, the digitization of cities can be seen as a paradigm shift extending the Internet of Things (IoT) to a citizen-centric model in order to leverage the massive data collected by pervasive sensors, connected mobiles or fixed devices, and social applications.

5. Highlights of the Year

5.1. Highlights of the Year

- Walid Bechkit holds the PEDR (2017-2021).
- Khaled Boussetta obtained his HDR from the University Paris 13, in December 2018.
- Khaled Boussetta holds the PEDR (2018-2022).
- Khaled Boussetta was promoted MCF *Hors Classe* in September 2018.
- Hervé Rivano holds the PEDR (2017-2021).
- Razvan Stanica holds the PEDR (2016-2020).
- Pascale Vicat Blanc joined Agora as Inria Senior Researcher, in September 2018.

6. New Software and Platforms

6.1. TAPASCologne

Travel and Activity PAtterns Simulation Cologne

KEYWORDS: Mobility - Traces
FUNCTIONAL DESCRIPTION: TAPASCologne is an initiative by the Institute of Transportation Systems at the German Aerospace Center (ITS-DLR), aimed at reproducing, with the highest level of realism possible, car traffic in the greater urban area of the city of Cologne, in Germany.

To that end, different state-of-art data sources and simulation tools are brought together, so to cover all of the specific aspects required for a proper characterization of vehicular traffic:

The street layout of the Cologne urban area is obtained from the OpenStreetMap (OSM) database, The microscopic mobility of vehicles is simulated with the Simulation of Urban Mobility (SUMO) software, The traffic demand information on the macroscopic traffic flows across the Cologne urban area (i.e., the O/D matrix) is derived through the Travel and Activity PAtterns Simulation (TAPAS) methodology, The traffic assignment of the vehicular flows described by the TAPASCologne O/D matrix over the road topology is performed by means of Gawron's dynamic user assignment algorithm.

- Participants: Marco Fiore and Razvan Stanica
- Contact: Marco Fiore
- URL: http://kolntrace.project.citi-lab.fr/#download

6.2. Sense in the City

KEYWORDS: Sensors - Sensors network - Wireless Sensor Networks
FUNCTIONAL DESCRIPTION: Sense in the city is a lightweight experimentation platform for wireless sensor networks in development. The main objective of this platform is to be easily transferable and deployable on the field. It allows a simplified deployment of the code running on the sensors and the collection of logs generated by the instrumentation of the code on a centralized database. In the early stage of the platform, the sensors are powered by small PCs, e.g. Raspberry Pis, but we are investigating the integration of energy harvesting capabilities such as solar panels.

- Participants: Hervé Rivano and Khaled Boussetta
- Contact: Khaled Boussetta

6.3. PrivaMovApp

KEYWORD: Crowd-sensing
FUNCTIONAL DESCRIPTION: Agora is leading the development of an Android application for user data collection purposes. The application is based on the Funf framework, and is currently available on Google Play.

- Participants: Stéphane D'alu, Hervé Rivano, Razvan Stanica and SOLOHAJA RABENJAMINA
- Contact: Razvan Stanica

6.4. WSNet

KEYWORD: Network simulator
FUNCTIONAL DESCRIPTION: WSNet is a modular event-driven simulator targeted to Wireless Sensor Networks. Its main goals are to offer scalabiliy, extensibility and modularity for the integration of new protocols/hardware models and a precise radio medium simulation. We still hope to find the proper resource to make WSNet evolve into a wireless capillary network simulator suitable for conducting simulations at the urban scale.

- Participants: Rodrigue Domga Komguem and Fabrice Valois
- Partner: CEA-LETI
- Contact: Guillaume Chelius
- URL: https://gforge.inria.fr/projects/wsnet-3/

6.5. Platforms

6.5.1. PPAIR Plateforme LoRa - Campus Connecté

The project aims at providing a platform that offers connectivity through a long-range, low-energy network to smart objects. The platform uses LoRa technology, which offers a wide connectivity, covering the entire INSA Lyon campus and providing a data collection service to all campus users. The main purpose of the LoRa plateforme is: *(i)* research (researchers can use it for studying reliability and capacity problems, privacy related challenges, etc.), and *(ii)* teaching (several courses from INSA, especially in the Telecom department can use this platform as a pedagogical tool).
Part of the software is mutualized with the University of Paris 13, where a LoRaWan testbed project is under deployment at the campus of Villetaneuse. This project, is supported by a local BQR and is leaded by Khaled Boussetta. The mutualization of the software tools will allows us to conduct multi sites experiments, at Lyon and at Paris.

6.5.2. UrPolSens Platform

We designed from scratch an energy efficient air pollution sensor network using Atmega micro-controllers and electrochemical air pollution probes. The micro-controller is integrated into a lab-designed printed circuit which includes among others: a high precision ADC, a micro-SD card reader and a radio communication module. The designed nodes measure the nitrogen dioxide (NO2) pollutant in addition to temperature and humidity and transmit data using LoRa to a gateway, which is connected to our servers using a 4G connection. The sensors are also equipped with solar panels in order to extend their lifetime when their batteries are drained. Our platform had been operational in the downtown of the Lyon city with 12 sensor nodes deployed in the Garibaldi street from mid-July to Mid-October 2018.

7. New Results

7.1. Wireless network deployment

Participants : Walid Bechkit, Amjed Belkhiri, Jad Oueis, Hervé Rivano, Razvan Stanica, Fabrice Valois

7.1.1. UAVs positionning

Mobile base stations mounted on unmanned aerial vehicles (UAVs) provide viable wireless coverage solutions in challenging landscapes and conditions, where cellular/WiFi infrastructure is unavailable. Operating multiple such airborne base stations, to ensure reliable user connectivity, demands intelligent control of UAV movements, as poor signal strength and user outage can be catastrophic to mission critical scenarios. In [17], we propose a deep reinforcement learning based solution to tackle the challenges of base stations mobility control. We design an Asynchronous Advantage Actor-Critic (A3C) algorithm that employs a custom reward function, which incorporates SINR and outage events information, and seeks to provide mobile user coverage with the highest possible signal quality. Preliminary results reveal that our solution converges after 4×10 5 steps of training, after which it outperforms a benchmark gradient-based alternative, as we attain 5dB higher median SINR during an entire test mission of 10,000 steps.

7.1.2. Network functions placement

Emerging mobile network architectures (e.g., aerial networks, disaster relief networks) are disrupting the classical careful planning and deployment of mobile networks by requiring specic self-deployment strategies. Such networks, referred to as self-deployable, are formed by interconnected rapidly deployable base stations that have no dedicated backhaul connection towards a traditional core network. Instead, an entity providing essential core network functionalities is co-located with one of the base stations. In [5], we tackle the problem of placing this core network entity within a self-deployable mobile network, i.e., we determine with which of the base stations it must be co-located. We propose a novel centrality metric, the ow centrality, which measures a node capacity of receiving the total amount of ows in the network. We show that in order to maximize the amount of exchanged trac between the base stations and the core network entity, under certain capacity and load distribution constraints, the latter should be co-located with the base station having the maximum ow centrality. We rst compare our proposed metric to other state of the art centralities. Then, we highlight the signicant trac loss occurring when the core network entity is not placed on the node with the maximum ow centrality, which could reach 55

7.1.3. Mobile edge computing orchestration

Orchestrating network and computing resources in Mobile Edge Computing (MEC) is an important item in the networking research agenda. In [12], we propose a novel algorithmic approach to solve the problem of dynamically assigning base stations to MEC facilities, while taking into consideration multiple time-periods, and computing load switching and access latency costs. In particular, leveraging on an existing state of the art on mobile data analytics, we propose a methodology to integrate arbitrary time-period aggregation methods into a network optimization framework. We notably apply simple consecutive time period aggregation and agglomerative hierarchical clustering. Even if the aggregation and optimization methods represent techniques which are different in nature, and whose aim is partially overlapping, we show that they can be integrated in an efficient way. By simulation on real mobile cellular datasets, we show that, thanks to the clustering, we can scale with the number of time-periods considered, that our approach largely outperforms the case without time-period aggregations in terms of MEC access latency, and at which extent the use of clustering and time aggregation affects computing time and solution quality.

7.1.4. On User Mobility in Dynamic Cloud Radio Access Networks

The development of virtualization techniques enables an architectural shift in mobile networks, where resource allocation, or even signal processing, become software functions hosted in a data center. The centralization of computing resources and the dynamic mapping between baseband processing units (BBUs) and remote antennas (RRHs) provide an increased flexibility to mobile operators, with important reductions of operational costs. Most research efforts on Cloud Radio Access Networks (CRAN) consider indeed an operator perspective and network-side performance indicators. The impact of such new paradigms on user experience has been instead overlooked. In [20], we shift the viewpoint, and show that the dynamic assignment of computing resources enabled by CRAN generates a new class of mobile terminal handover that can impair user quality of service. We then propose an algorithm that mitigates the problem, by optimizing the mapping between BBUs and RRHs on a time-varying graph representation of the system. Furthermore, we show that a practical online BBU-RRH mapping algorithm achieves results similar to an oracle-based scheme with perfect knowledge of future traffic demand. We test our algorithms with two large-scale real-world datasets, where the total number of handovers, compared with the current architectures, is reduced by more than 20%. Moreover, if a small tolerance to dropped calls is allowed, 30% less handovers can be obtained.

7.1.5. Wireless sensor network deployment for environmental monitoring

Air pollution has major negative effects on both human health and environment. Thus, air quality monitoring is a main issue in our days. In [9], we focus on the use of mobile WSN to generate high spatio-temporal resolution air quality maps. We address the sensors' online redeployment problem and we propose three redeployment models allowing to assess, with high precision, the air pollution concentrations. Unlike most of existing movement assisted deployment strategies based on network generic characteristics such as coverage and connectivity, our approaches take into account air pollution properties and dispersion models to offer an efficient air quality estimation. First, we introduce our proposition of an optimal integer linear program based on air pollution dispersion characteristics to minimize estimation errors. Then, we propose a local iterative integer linear programming model and a heuristic technique that offer a lower execution time with acceptable estimation quality. We evaluate our models in terms of execution time and estimation quality using a real data set of Lyon City in France. Finally, we compare our models' performances to existing generic redeployment strategies. Results show that our algorithms outperform the existing generic solutions while reducing the maximum estimation error up to 3 times.

7.2. Wireless data collection

Participants : Walid Bechkit, Ahmed Boubrima, Alexis Duque, Abdoul-Aziz Mbacke, Hervé Rivano, Razvan Stanica, Yosra Zguira

7.2.1. RFID paradigm

While RFID technology is gaining increased attention from industrial community deploying different RFID-based applications, it still suffers from reading collisions. As such, many proposals were made by the scientific community to try and alleviate that issue using different techniques either centralized or distributed, mono-channel or multi-channels, TDMA or CSMA. However, the wide range of solutions and their diversity make it hard to have a clear and fair overview of the different works. In [4], we propose a survey of the most relevant and recent known state-of-the- art anti-collision for RFID protocols. It provides a classification and performance evaluation taking into consideration different criteria as well as a guide to choose the best protocol for given applications depending on their constraints or requirements but also in regard to their deployment environments.

7.2.2. Anti-collision and routing protocol for RFID

In the midst of Internet of Things development, a first requirement was tracking and identification of those mentioned "things" which could be done thanks to Radio Frequency Identification. However, since then, the development of RFID allowed a new range of applications among which is remote sensing of environmental values. While RFID can be seen as a more efficient solution than traditional Wireless Sensor Networks, two main issues remain: first reading collisions and second proficient data gathering solution. In [18], we examine the implementation of two applications: for industrial IoT and for smart cities, respectively. Both applications, in regards to their requirements and configuration, challenge the operation of a RFID sensing solution combined with a dynamic wireless data gathering over multi-hops. They require the use of both mobile and fixed readers to cover the extent of deployment area and a quick retrieval of tag information. We propose a distributed cross-layer solution for improving the efficiency of the RFID system in terms of collision and throughput but also its proficiency in terms of tag information routing towards one or multiple sinks. Simulation results show that we can achieve high level of throughput while maintaining a low level of collision and a fairness of reader medium access above 95% in situations where readers can be fix and mobile, while tag information is routed with a data rate of 97% at worst and reliable delays for considered applications.

7.2.3. Routing priority information in RFID

Long being used for identification purposes, a new set of applications is now available thanks to the development of RFID technology. One of which is remote sensing of environmental values using passive RFID tags. This leap forward allowed a more energy efficient and cheaper solution for applications like logistics or urban infrastructure monitoring. Nevertheless, serious issues raised with the use of RFID: (i) reading collisions and (ii) gathering of tag information. Indeed, tags information retrieved by readers have to be transmitted towards a base station through a multi-hop scheme which can interfere with neighboring readers activity. In [19], we propose cross-layer solutions meant for both scheduling of readers' activity to avoid collisions, and a multi-hop routing towards base stations, to gather read tag data. This routing is performed with a data priority aware mechanism allowing end-to-end delay reduction of urgent data packets delivery up to 13% faster compared to standard ones. Using fuzzy logic, we combine several observed metrics to reduce the load of forwarding nodes and improve latency as well as data rate. We validate our proposal running simulations on industrial and urban scenarios.

7.2.4. Data collection in DTN networks

Intelligent Transport Systems (ITS) are an essential part of the global world. They play a substantial role for facing many issues such as traffic jams, high accident rates, unhealthy lifestyles, air pollution, etc. Public bike sharing system is one part of ITS and can be used to collect data from mobiles devices. In this paper, we propose an efficient, *Internet of Bikes*, IoB-DTN routing protocol based on data aggregation which applies the Delay Tolerant Network (DTN) paradigm to Internet of Things (IoT) applications running data collection on urban bike sharing system based sensor network. In [6], we propose and evaluate three variants of IoB-DTN: IoB based on spatial aggregation (IoB-SA), IoB based on temporal aggregation (IoB-TA) and IoB based on spatiotemporal aggregation (IoB-STA). The simulation results show that the three variants offer the best performances regarding several metrics, comparing to IoB-DTN without aggregation and the low-power long-range technology, LoRa type. In an urban application, the choice of the type of which variant of IoB should be used depends on the sensed values.

7.2.5. Data sensing in Internet of Bikes

Following the trend of the Internet of Thing, public transport systems are seen as an efficient bearer of mobile devices to generate and collect data in urban environments. Bicycle sharing system is one part of the city's larger transport system. In [23], we study the *Internet of Bikes* IoB-DTN protocol which applies the Delay Tolerant Network (DTN) paradigm to the Internet of Things (IoT) applications running on urban bike sharing system based sensor network. We evaluate the performances of the protocol with respect to the transmission power. Performances are measured in terms of delivery rate, delivery delay, throughput and energy cost. We also compare the multi-hop IoB-DTN protocol to a low-power wide-area network (LPWAN) technology. LPWAN have been designed to provide cost-effective wide area connectivity for small throughput IoT applications: multiyear lifetime and multi-kilometer range for battery-operated mobile devices. This work aims at providing network designers and managers insights on the most relevant technology for their urban applications that could run on bike sharing systems. To the best of our knowledge, this work is the first to provide a detailed performance comparison between multi-hop and long range DTN-like protocol being applied to mobile network IoT devices running a data collection applications in an urban environment.

7.2.6. Reducing IoT traffic through data aggregation mechanisms

Intelligent Transport Systems (ITS) are an essential part of the global world. They play a substantial role for facing many issues such as traffic jams, high accident rates, unhealthy lifestyles, air pollution, etc. Public bike sharing system is one part of ITS and can be used to collect data from mobiles devices. In this paper, we propose an efficient, " Internet of Bikes " , IoB-DTN routing protocol based on data aggregation which applies the Delay Tolerant Network (DTN) paradigm to Internet of Things (IoT) applications running data collection on urban bike sharing system based sensor network. In [6], we propose and evaluate three variants of IoB-DTN: IoB based on spatial aggregation (IoB-SA), IoB based on temporal aggregation (IoB-TA) and IoB based on spatiotemporal aggregation (IoB-STA). The simulation results show that the three variants offer the best performances regarding several metrics, comparing to IoB-DTN without aggregation and the low-power long-range technology, LoRa type. In an urban application, the choice of the type of which variant of IoB should be used depends on the sensed values.

7.2.7. Environmental modeling

Wireless sensor networks (WSN) are widely used in environmental applications where the aim is to sense a physical parameter such as temperature, humidity, air pollution, etc. Most existing WSN-based environmental monitoring systems use data interpolation based on sensor measurements in order to construct the spatiotemporal field of physical parameters. However, these fields can be also approximated using physical models which simulate the dynamics of physical phenomena. In [11], we focus on the use of wireless sensor networks for the aim of correcting the physical model errors rather than interpolating sensor measurements. We tackle the activity scheduling problem and design an optimization model and a heuristic algorithm in order to select the sensor nodes that should be turned off to extend the lifetime of the network. Our approach is based on data assimilation which allows us to use both measurements and the physical model outputs in the estimation of the spatiotemporal field. We evaluate our approach in the context of air pollution monitoring while using a dataset from the Lyon city, France and considering the characteristics of a monitoring system developed in our lab. We analyze the impact of the nodes' characteristics on the network lifetime and derive guidelines on the optimal scheduling of air pollution sensors.

7.2.8. Multi-robot routing for evolving missions

In [22], we propose Dynamic Multi Robot-Routing (DMRR), as a continuous adaptation of the multi-robot target allocation process (MRTA) to new discovered targets. There are few works addressing dynamic target allocation. Existing methods are lacking the continuous integration of new targets, handling its progressive effects, but also lacking dynamic support (e.g. parallel allocations, participation of new robots). This work proposes a framework for dynamically adapting the existing robot missions to new discovered targets. Missions accumulate targets continuously, so the case of a saturation bound for the mission costs is also considered. Dynamic saturation-based auctioning (DSAT) is proposed for allocating targets, providing lower time complexities (due to parallelism in allocation). Comparison is made with algorithms ranging from greedy to auction-based methods with provable sub-optimality. The algorithms are tested on exhaustive sets of inputs, with random configurations of targets (for DMRR with and without a mission saturation bound). The results for DSAT show that it outperforms state-of-the-art methods, like standard sequential single-item auctioning (SSI) or SSI with regret clearing.

7.2.9. Measuring information using VLC

The use of visible light for bidirectional communication between regular smartphones and the small LEDs integrated in most consumer electronics nowadays raises new challenges. In [13], we enhance the state of the art with an efficient image processing algorithm to accurately detect the LEDs and decode their signal in real time. We propose an efficient decoding algorithm, which can detect the LED position, process and decode the signal on average in 18.4 ms, for each frame, on a Nexus 5 unrooted smartphone. Thus, this implementation is convenient for low latency indoor localization or real-time transmission with a moving receiver. Also, as the ROI detection is the most complex step of the algorithm, scenarios with several transmitters can be envisaged, enabling MIMO-like transmissions. We also present smart mechanisms and protocols to build a robust flash-to-LED communications channel using off- the-shelf smartphones and small LEDs. Our experimental evaluation shows a throughput of 30 bit/s, which is suitable for feedback, wake-up or even some limited communication purposes. We believe that such bidirectional VLC communication system will be a great opportunity for smart and connected consumer electronic products, providing bidirectional smartphone- to-device communication at lower cost.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- We have contracted a first bilateral contract with Total (2018-2019) where we work with the laboratory LQA of Total on the design and the test of autonomous low cost air quality sensors. The Lora-based developed platform is currently deployed et evaluated by LQA.
- We have contracted bilateral cooperation with Rtone, an SME focusing on the connected objects area. This collaboration is associated with the CIFRE PhD grant for Alexis Duque, on the subject of Visible Light Communication.
- We have contracted bilateral cooperation with industrial and academic partners in the context of the PSPC Fed4PMR project (2015-2019). In this context, we are working on the design of new professional mobile radio solutions, compatible with 4G and 5G standards. This collaboration funds the PhD thesis of Jad Oueis, the PhD thesis of Romain Pujol, and a part of the PhD thesis of Abderrahman Ben Khalifa.

8.2. Bilateral Grants with Industry

- Common Laboratory Inria/Nokia Bell Labs - ADR Network Information Theory.
  Agora is part of the ADR Network Information Theory of the common laboratory Inria/Nokia Bell Labs.

- Spie - INSA Lyon IoT Chaire.
  Agora is involved in the SPIE INSA Lyon IoT Chaire, launched in November 2016. The IoT Chaire partially funds the PhD thesis of Abderrahman Ben Khalifa. The PhD thesis work of Alexis Duque and Amjed Belkhiri are also contributing in this structure.
- Volvo - INSA Lyon Chaire.
  Agora is involved in the Volvo Chaire at INSA Lyon, on the area of autonomous electrical distribution vehicle in urban environments. Razvan Stanica is a member in the steering committee of this structure.

9. Partnerships and Cooperations

9.1. Regional Initiatives

- FIL Grant, 2018
  Participants: Hervé Rivano
  The partners of this project, supported by the *Fédération d'Informatique de Lyon*, are: CITI, LIP.
  The goal is to use crow-sensing applications with data collection of Wi-Fi networks which are available in the neighborhood in order *i)* to build a map of the wireless network in terms of performance for the application and *ii)* to optimize the wireless network configuration.
- Labex IMU UrPolSens,10/2015-10/2018
  Participants: Walid Bechkit, Amjed Belkhiri, Ahmed Boubrima, Hervé Rivano
  The partners in this project are Ifsttar, LMFA, EVS, TUBA, and Air Rhone-Alpes, with Inria Agora leading the project.
  UrPolSens deals with the monitoring of air pollution using low-cost sensors interconnected by a wireless networks. Although they are less accurate than the high-end sensors used today, low-cost autonomous air quality sensors allow to achieve a denser spatial granularity and, hopefully, a better monitoring of air pollution. The main objectives of this project are to improve the modeling of air pollution dispersion; propose efficient models to optimize the deployment the sensors while considering the pollution dispersion and the impact of urban environment on communications; deploy a small-scale network for pollution monitoring as a proof of concept; compare the measured and estimated levels of exposure; study the spatial disparities in exposure between urban areas.
- Labex IMU 3M'Air 2018-2021
  Participants: Walid Beckhit, Ahmed Boubrima, Manoel Dahan, Mohamed Anis Fekih, Ichrak Mokhtari, Hervé Rivano.
  The partners in this project are: EVS, LMFA, Métropole de Lyon, Ville de Lyon , Atmo AURA, Météo France, Lyon Météo. Inria Agora is the leader of this project.
  The 3M'Air project explore the potential of participatory sensing to improve local knowledge of air quality and urban heat islands. The main aim of this project is therefore to equip citizens with low-cost mobile sensors and then ensure an efficient real-time data collection and analysis. This allows to obtain a finer spatiotemporal granularity of measurements with lighter installation and operational costs while involving citizens.
- ARC6 Robot fleet mobility under communication constraints, 10/2016-09/2019.
  Participant : Fabrice Valois.
  This work is a joint project with the Inria Chroma research group. Considering a fleet of drones moving in a 3D area, looking for a given target, we focus on how to maintain the wireless connectivity of the network of drones while the drones patrol autonomously. The other partners in this project are University of Grenoble and Viameca.

- Labex IMU Veleval, 10/2017-10/2019
Participant: Hervé Rivano.
The partners in this project are: EVS, LIRIS, LLSETI and CITI, with LAET leading the project.
The goal of this pluridisciplinary project is to study, understand and model the behavior of cyclists in an urban environment with a methodology combining quantitative measurements of mobility traces and image analysis with qualitative informations from reactivation interviews. In particular the input of Agora is to provide crowdsourcing tools for gathering mobility data that are optimized for the practice of urban cycling.

9.2. National Initiatives

9.2.1. ANR

- ANR CANCAN 2019 - 2022 (accepted in 2018, kickoff in February 2019)
Participants: Solohaja Rabenjamina, Razvan Stanica.
The partners in this project are: CEDRIC, Inria, Orange Labs, with Thalès Communications & Security leading the project.
The ANR CANCAN (Content and context based adaptation in mobile networks) targets the following objectives: *i)* collecting novel measurement datasets that describe mobile network data traffic at unprecedented spatial and temporal accuracy levels, and for different mobile services separately. The datasets will be gathered in an operational nationwide network, *ii)* evaluating existing analytics for classification, prediction and anomaly detection within real-world high-detail per-service mobile network data, and tailoring them to the specifications of the management of resources at different network levels, and *iii)* demonstrating the integration of data analytics within next-generation cognitive network architectures in several practical case studies.
- ANR MAESTRO 5G 2019 - 2022 (accepted in 2018, kickoff in February 2019)
Participants: Hervé Rivano, Razvan Stanica.
The partners in this project are: CEDRIC, Inria, L2S, LIA, Nokia Bell Labs, TSP, with Orange Labs leading the project.
The ANR MAESTRO 5G (Management of slices int he radio access of 5G networks) is expected to provide: *i)* a resource allocation framework for slices, integrating heterogeneous QoS requirements and spanning on multiple resources including radio, backhauling/fronthauling and processing resources in the RAN, *ii)* acomplete slice management architecture including provisioning and re-optimization modules and their integration with NFV and SDN strata, *iii)* a business layer for slicing in 5G, *iv)* a demonstrator showing the practical feasibility as well as integration of the major functions and mechanisms proposed by the project, on a 5G Cloud RAN platform. The enhanced platform is expected to support the different 5G services.
- ANR CoWorkWorlds 01/2018 - 12/2020.
Participants: Solohaja Rabenjamina, Razvan Stanica.
The ANR CoWorkWorlds (Sustainability and spatiality in co-workers' mobility practices) project is lead by ENTPE. Its focus is on the study of co-working environments, and more precisely on the mobility behaviour of users of such spaces. Our role in the project is to collect and analyse mobility data from a set of users, using the PrivaMov smartphone application.

9.2.2. DGA

- DGA CLOTHO 10/2016-10/2018.
Participants: Junaid Khan, Romain Pujol, Razvan Stanica, Fabrice Valois
The partners in the DGA CLOTHO project are Traqueur and Sigfox. The objective of the project is to reduce the energy consumption of the device tracking functionality, by taking profit of short-range communications between the tracked objects.

9.2.3. PIA

- PIA ADAGE 07/2016-06/2018.
  Particpants: Elli Zavou, Razvan Stanica
  The partners in the PIA ADAGE project are Orange, LAAS-CNRS and Inria Privatics. The objective of the ADAGE project is to design and evaluate anonymization algorithms for the specific case of mobile traffic data. Our role in the project is focused on evaluating whether the anonymized data is still usable for adaptive networking mechanisms.

9.2.4. Pôle ResCom

- Ongoing participation (since 2006)
  Communication networks, working groups of GDR ASR/RSD, CNRS (http://rescom.inrialpes.fr). Hervé Rivano is member of the scientific committee of ResCom.

9.2.5. EquipEx

- SenseCity
  We have coordinated the participation of several Inria teams to the SenseCity EquipEx. Within the SenseCity project, several small reproduction of 1/3rd scale city surroundings will be built under a climatically controlled environment. Micro and nano sensors will be deployed to experiment on smart cities scenarios, with a particular focus on pollution detection and intelligent transport services. Agora will have the opportunity to tests some of its capillary networking solutions in a very realistic but controlled urban environment. A proof of concept test site has been built in 2015. We have deployed an experiment on low cost sensor network for vehicle detection and one on atmospheric pollution sensor calibration. The operational site is build, the information system is being finalized and the equipment will be inaugurated in April 2018.

9.2.6. Inria Project lab

- CityLab
  Agora is involved in the CityLab Inria Project Lab lead by Valérie Issarny. Within this project, Hervé Rivano co-advises, with Nathalie Mitton (FUN team, Inria Lille-Nord-Europe), the PhD thesis of Abdoul Aziz Mbacke on “Data gathering in sensor and passive RFID with energy harvesting for urban infrastructure monitoring”.

9.3. European Initiatives

9.3.1. Collaborations in European Programs, Except FP7 & H2020

Program: Interreg Med

Project acronym: ESMARTCITY

Project title: Enabling Smarter City in the MED Area through Networking

Duration: 02/2018 - 07/2020

Coordinator: Abruzzo Region, Italy

Other partners: ARIC and RWG (Greece), APEGR (Spain), RAIS (Bosnia and Herzegovina), ENA (Portugal), MCM and PoliMi (Italy), Capergies (France)

Abstract: The project has its primary objective in improving the innovation capacity of MED cities by creating innovation ecosystems, which involve actors of the quadruple helix (Citizens, Businesses Operators, Research, Universities and Public Authorities) , and in applying the Smart City concept, which utilizes digital and energy saving technologies to allow better services for the citizen with less impact on the environment, producing furthermore new employability and living scenarios. To achieve this goal, the project envisages the pilot testing of the Smart City concept to provide specific services to citizens in the field of intelligent urban districts, energy efficiency of buildings and smarter public lighting.

9.4. International Initiatives

9.4.1. Inria International Partners

9.4.1.1. Informal International Partners

- **University of Waterloo, ON, Canada**. Joint publications and visits to/from the group of Prof. Catherine Rosenberg.
- **Nimbus Centre, Cork, Ireland**. Collaboration around LoRa experiments with Dr. Ramona Marfievici.
- **CNR-IEIIT, Turin, Italy**. Joint publications and projects with Dr. Marco Fiore.
- **Trento University, Italy**. Collaboration around routing for IoT networks with the group of Prof. Gian Pietro Picco.
- **University of Edinburgh, UK**. Joint publications and visits to/from the group of Dr. Paul Patras.

9.4.2. Participation in Other International Programs

9.4.2.1. PHC Campus France

- **University of Cluj-Napoca, Romania**. PHC DRONEM (2017-2019) on Monitoring using connected fleet of drones, a collaboration with the group of Prof. Gabriela Czibula.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Abdelmalik Bachir, Professor, Biskra University, Algeria: invited professor at INSA Lyon (July, 2018)
- Josep Paradells Aspas, Professor, Universitat Politecnica de Catalunya, Barcelona, Spain: invited professor at INSA Lyon (October 2018)
- Rui Li, PhD student, University of Edinburgh, Scotland, UK: visiting PhD student (March, 2018)

9.5.1.1. Research Stays Abroad

- Ahmed Boubrima visited the group of Prof. Azzedine Boukerche, University of Ottawa, Canada.
- Mihai Popescu visited the group of Prof. Gabriela Czibula, at University of Cluj-Napoca, Romania (3 periods of 1 month duration: April, June and November 2018).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Khaled Boussetta was co-track chair of Wireless Models and Simulations in the 10th IFIP Wireless Days 2018, April, 2018, Dubai, UAE.

10.1.1.2. Member of the Organizing Committees

- Khaled Boussetta was on the Organizing Committee of 2018 IEEE SOCA/SC2/IOV, 19-22 November, 2018, Paris.

10.1.1.3. Member of the Conference Program Committees

- Walid Bechkit was in the TPC of the followinf conference: IEEE ICC.
- Khaled Boussetta was in the TPC of the following conferences: IEEE CCNC, IEEE GlobeCom, IEEE ICC, Med-Hoc-Net, Wireless Days.

- Oana Iova was in the TPC of the following conferences: IEEE Globecom, ACM AINTEC, AdHoc-Now, EWSN 2018 (track Poster and Demos).
- Razvan Stanica was in the TPC of the following conferences: IEEE ICC, IEEE CCNC, ICIN.
- Fabrice Valois was in the TPC of the following conferences: IEEE ICC, ICT, IEEE Globecom, IEEE VTC Spring, WiSARN.

10.1.1.4. Reviewer

- Oana Iova was a reviewer for the following conferences: IEEE ICC, IEEE WCNC.

10.1.2. Journal

10.1.2.1. Reviewer - Reviewing Activities

- Oana Iova was a reviewer for the following journals: ACM TOSN, Elsevier Ad Hoc Networks, Elsevier Computer Networks, IEEE Transactions on Communications, IEEE Access, Springer Wireless Networks.
- Razvan Stanica was a reviewer for the following journals: IEEE Transactions on Mobile Computing, IEEE Transactions on Intelligent Transportation Systems, IEEE Communications Letters, Pervasive and Mobile Computing, IEEE Access, EURASIP Journal on Wireless Communications and Networking, IET Intelligent Transport Systems.

10.1.3. Invited Talks

- Walid Bechkit was invited speaker at the 1st Winter School on Information Theory and Signal Processing for Internet of Things, Villeurbanne, France.
- Walid Bechkit was invited speaker at the CESER (conseil économique, social et environnemental régional).
- Oana Iova was invited panelist at: 1st Winter School on Information Theory and Signal Processing for Internet of Things, Villeurbanne, France.
- Hervé Rivano was invited speaker at *La Nuit des idées*, Maison de France, Sfax, Tunisia.
- Hervé Rivano was invited speaker at La Myne, Lyon, France.
- Hervé Rivano was invited speaker Conseil de Développement de la Métropole/Amis de l'Université, Lyon, France.
- Hervé Rivano was invited speaker IMERIR, Perpignan, France.
- Hervé Rivano was invited speaker Salon Internet des Objets (SidO), Lyon, France.
- Hervé Rivano was invited speaker réseau ARAMIS, Lyon, France.
- Hervé Rivano was invited speaker at CIVIC Conference, University Constantine III, Algeria.
- Hervé Rivano was invited speaker at TokyoTech, Japan.
- Hervé Rivano was panel organizer on Smart Cities, POP Science Forum, Lyon, France.

10.1.4. Leadership within the Scientific Community

- Walid Bechkit was co-representative of the Labex IMU at the Smart City Expo World Congres (Barcelona).
- Walid Bechkit is a nominated member in the scientific committee of the Fédération d'Informatique de Lyon (FR 2000 CNRS).
- Hervé Rivano is member of the steering committee of the ResCom axis of the RSD CNRS GdR.
- Hervé Rivano is a member of the Scientific Council of TUBA Lyon.
- Hervé Rivano is member ot the steering committee of Ecole Urbaine de Lyon, in charge of Smart and Learning Cities.
- Fabrice Valois is a member of the Scientific Council of the LIMOS-UMR6158 laboratory, Clermont Ferrand.

- Fabrice Valois is member of the Scientific Council of the Labex IMU (Intelligence des Mondes Urbains).

10.1.5. Scientific Expertise

- Khaled Boussetta was reviewer for ANR generic call for projects 2018.
- Hervé Rivano is member of the Metropole de Lyon [R] Challenge, board of scientific experts.
- Hervé Rivano is member of the Scientific Committee of the Digital League Regional Cluster.
- Fabrice Valois was a member in the recruitment committee of a Full Professor in Computer Science at INSA Lyon.
- Fabrice Valois was a member in the recruitment committee of a Full Professor in Computer Science at Université d'Auvergne.

10.1.6. Research Administration

- Walid Bechkit is responsible for seminar organization and scientific animation within the CITI laboratory.
- Khaled Boussetta is member of the steering committee of the MathStic federation at University Paris 13.
- Hervé Rivano is member of the Administration Council of the EquipEx Sense City as representative of Inria.
- Hervé Rivano is president of the CITI laboratory council.
- Razvan Stanica is the CITI laboratory correspondent with the Labex IMU.
- Razvan Stanica is member of the steering committee of the Volvo Chaire at INSA Lyon.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence : Walid Bechkit, IP Networks, 30h, L3, Telecom. Dpt. INSA Lyon.

Licence : Oana Iova, IP Networks, 30h, L3, Telecom. Dpt. INSA Lyon.

Licence : Oana Iova, Medium Access Control, 8h, L3, Telecom. Dpt. INSA Lyon.

Licence : Oana Iova, Network and System Programming, 82h, L3, Telecom. Dpt. INSA Lyon.

Licence : Hervé Rivano, Networking Fundamentals, 2h, L3, Telecom. Dpt. INSA Lyon.

Licence : Hervé Rivano, Algorithms and programming, 165h, L1 - L2, INSA Lyon.

Licence : Hervé Rivano, Sensors data engineering project, 34h, L2, INSA Lyon.

Licence : Razvan Stanica, Network Programming, 90h, L3, Telecom. Dpt. INSA Lyon.

Licence : Fabrice Valois, IP Networks, 42h, L3, Telecom. Dpt. INSA Lyon.

Master : Walid Bechkit, Introduction to wireless sensor networks, 50h, L2, INSA Lyon.

Master : Walid Bechkit, Performance evaluation of telecom networks, 100h, M1, Telecom. Dpt. INSA Lyon.

Master : Walid Bechkit, Cryptography and communication security, 30h, M1, Telecom. Dpt., INSA Lyon.

Master : Walid Bechkit, Wireless networks: architecture and security, 30h, M2, INSA Lyon.

Master : Master : Walid Bechkit, Network Acces Control, 6h, M2, Telecom. Dpt. INSA Lyon.

Master : Oana Iova, Network Routing Protocols, 40h, M1, Telecom. Dpt. INSA Lyon.

Master : Oana Iova, Long Range Networks, 10h , M2, Telecom. Dpt. INSA Lyon

Master : Hervé Rivano, Smart Cities and IoT, 44h, M2, Telecom. Dpt. INSA Lyon.

Master : Razvan Stanica, Mobile Networks, 30h, M1, Telecom. Dpt. INSA Lyon.

Master : Razvan Stanica, Network Science, 10h, M2, Telecom. Dpt. INSA Lyon.

Master : Fabrice Valois, Mobile Networks, 16h, M1, Telecom. Dpt. INSA Lyon.

MOOC : Hervé Rivano, Razvan Stanica, Fabrice Valois, Connectez à Internet vos Objets Intelligents, production started in the context of the ANR Connect-IO project.

Walid Bechkit is in charge of the admission service of the Telecommunication department at INSA Lyon.

Walid Bechkit is an elected member of the Telecommunication department council at INSA Lyon.

Walid Bechkit is the head of the networking teaching team in the Telecommunications department at INSA Lyon, coordinating all the courses in the networking domain.

Khaled Boussetta is the studies director of Apprenticeship Education Program Specialty in Computer Science and Network at Sup'Galilée Engineering School (University Paris 13).

Hervé Rivano is responsible for the coordination of all courses in the Smart Cities and IoT option at the INSA Lyon Telecommunications department.

Herve´ Rivano is referent DSI in the FIMI Dpt., INSA Lyon.

Razvan Stanica is responsible of the research option at the Telecommunications department of INSA Lyon.

Razvan Stanica is vice dean of the Telecommunications department of INSA Lyon, in charge of education related affairs.

10.2.2. Supervision

PhD in progress: Abderrahman Ben Khalifa, Cognitive mechanisms for IoT networks, since 11/2016. Advisors: Hervé Rivano, Razvan Stanica.

PhD in progress: Ahmed Boubrima, Optimal deployment of wireless sensor networks for air pollution monitoring, since 10/2015. Advisors: Walid Bechkit, Hervé Rivano.

PhD in progress : Rodrigue Domga Komguem, Autonomous WSN architectures for road traffic applications, since 11/2012. Advisors: Razvan Stanica, Maurice Tchuente (Univ. Yaoundé, Cameroun), Fabrice Valois.

PhD in progress: Mohamed Anis Fekih, Urban pollution using wireless sensor networks, since 11/2018.Advisors: Walid Bechkit, Hervé Rivano.

PhD in progress: Mihai Popescu, Connectivity constrained mobility in fleets of robots, since 11/2015. Advisors: Olivier Simonin (Inria CHROMA), Anne Spalanzanni (Inria CHROMA), Fabrice Valois.

PhD in progress: Romain Pujol, Data collection in dynamic wireless networks, since 11/2018. Advisors: Razvan Stanica, Fabrice Valois.

PhD in progress: Solohaja Rabenjamnia, Data analysis of cellular trafic, since 11/2018. Advisors: Hervé Rivano, Razvan Stanica.

10.2.3. Juries

- Hervé Rivano was a reviewer in the following PhD defense committee:
  - V. Quintuna Rodriguez, Nouvelle commande réseau IP : Performance des fonctions virtualisées pour une infrastructure programmable, LIP6, Université Pierre et Marie Curie, Paris VI
  - Y. Couble, Optimisation de la gestion des ressources de la voie retour, IRIT, Université de Toulouse
  - L. Molina, Metrics and non-intrusive techniques to characterize Wi-Fi networks, IRISA, Université de Bretagne-Pays de la Loire

  - H. Tall, Load balancing in multichannel data collection wireless sensor networks, LIMOS, Université de Clermont Auvergne
  - H. Pimenta Moraes Junior, A contribution to data sharing in vehicular networks, Heudiasyc, Université Technologique de Compiègne
- Hervé Rivano was a reviewer in the following HDR defense committee:
  - F. Giroire, Nouvelle commande réseau IP : Performance des fonctions virtualisées pour une infrastructure programmable, I3S, Université de Nice Sophia Antipolis
- Fabrice Valois was a reviewer in the following PhD defense committee:
  - E. Morin, Intéroperabilité de protocoles de communication adaptatifs basse consommation pour des réseaux de capteurs, LIG, Université Grenoble Alpes
- Fabrice Valois was a member in the following PhD defense committee:
  - D. Alshamaa, Indoor localization of sensors: application to dependent elderly people, UTT
  - H. Chelle, Contrôle de charge des réseaux IoT : D'une étude théorique à une implantation réelle, IRIT, ENSHEEIT
- Fabrice Valois was a reviewer in the following HDR defense committee:
  - P. Berthou, Vers la dématérialisation des Réseaux Hybrides satellites et terrestre, LAAS, Université Paul Sabatier
- Fabrice Valois was a member in the following HDR defense committee:
  - T. Begin, Contributions to the Performance Modeling of Computer Networks, LIP, Université de Lyon 1
  - K. Boussetta, Dimensionnement, adaptation et placement de fonctionnalités réseaux pour des applications sensibles aux délais, L2TI, Université Paris XIII

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. Duque.*Bidirectional Visible Light Communications for the Internet of Things*, Université de Lyon - INSA Lyon, October 2018, https://hal.inria.fr/tel-01940002

[2] A. A. Mbacké.*Collection and multi-hop forwarding of RFID data for the monitoring of urban infrastructures*, Université de Lille, October 2018, https://hal.inria.fr/tel-01901740

[3] J. Oueis.*Radio Access and Core Functionalities in Self-deployable Mobile Networks*, Université de Lyon - INSA Lyon, November 2018, https://hal.inria.fr/tel-01950198

Articles in International Peer-Reviewed Journal

[4] A. A. Mbacké, N. Mitton, H. Rivano.*A survey of RFID readers anticollision protocols*, in "IEEE Journal of Radio Frequency Identification", March 2018, vol. 2, n^{o} 1, 11 [*DOI :* 10.1109/JRFID.2018.2828094], https://hal.inria.fr/hal-01767311

[5] J. Oueis, V. Conan, D. Lavaux, H. Rivano, R. Stanica, F. Valois.*Core Network Function Placement in Self-Deployable Mobile Networks*, in "Computer Communications", January 2019, vol. 133, p. 12-23 [*DOI :* 10.1016/j.comcom.2018.10.009], https://hal.inria.fr/hal-01928552

[6] Y. ZGUIRA, H. RIVANO, A. MEDDEB.*Internet of Bikes: A DTN Protocol with Data Aggregation for Urban Data Collection*, in "Sensors", August 2018, vol. 19, n° 8, p. 1-39 [*DOI :* 10.3390/s18092819], https://hal.inria.fr/hal-01862155

International Conferences with Proceedings

[7] M. S. ASTEFANOAEI, P. CESARETTI, P. KATSIKOULI, M. GOSWAMI, R. SARKAR.*Multi-resolution sketches and locality sensitive hashing for fast trajectory processing*, in "SIGSPATIAL 2018 - International Conference on Advances in Geographic Information Systems", Seattle, United States, ACM, November 2018, vol. 10 [*DOI :* 10.1145/3274895.3274943], https://hal.inria.fr/hal-01893304

[8] A. BELKHIRI, W. BECHKIT, H. RIVANO.*Virtual Forces based UAV Fleet Mobility Models for Air Pollution Monitoring*, in "LCN 2018 - 43rd Annual IEEE Conference on Local Computer Networks", Chicago, Illinois, United States, IEEE, October 2018, p. 1-4, https://hal.inria.fr/hal-01956717

[9] A. BELKHIRI, W. BECHKIT, H. RIVANO, M. KOUDIL.*Context Aware MWSN Optimal Redeployment Strategies for Air Pollution Timely Monitoring*, in "ICC 2018 - IEEE International Conference on Communications", Kansas, United States, IEEE, May 2018, p. 1-7 [*DOI :* 10.1109/ICC.2018.8422395], https://hal.inria.fr/hal-01780772

[10] A. BOUBRIMA, W. BECHKIT, H. RIVANO, L. SOULHAC.*Leveraging the Potential of WSN for an Efficient Correction of Air Pollution Fine-Grained Simulations*, in "ICCCN 2018 - 27th International Conference on Computer Communications and Networks", Hangzhou, China, IEEE, July 2018, p. 1-9 [*DOI :* 10.1109/ICCCN.2018.8487343], https://hal.inria.fr/hal-01781389

[11] A. BOUBRIMA, A. BOUKERCHE, W. BECHKIT, H. RIVANO.*WSN Scheduling for Energy-Efficient Correction of Environmental Modelling*, in "IEEE MASS 2018 - 15th IEEE International Conference on Mobile Ad-hoc and Sensor Systems", Chengdu, China, October 2018, p. 1-8, https://hal.inria.fr/hal-01865111

[12] A. CESELLI, M. FIORE, A. FURNO, M. PREMOLI, S. SECCI, R. STANICA.*Prescriptive Analytics for MEC Orchestration*, in "IFIP Networking 2018", Zurich, Switzerland, May 2018, p. 1-9, https://hal.sorbonne-universite.fr/hal-01740816

[13] A. DUQUE, R. STANICA, H. RIVANO, A. DESPORTES.*Decoding Methods in LED-to-Smartphone Bidirectional Communication for the IoT*, in "Global LiFi Congress - 1st edition", Paris, France, February 2018, p. 1-6 [*DOI :* 10.23919/GLC.2018.8319118], https://hal.inria.fr/hal-01683629

[14] A. DUQUE, R. STANICA, H. RIVANO, C. GOURSAUD, A. DESPORTES.*Poster: Insights into RGB-LED to Smartphone Communication*, in "EWSN 2018 - International Conference on Embedded Wireless Systems and Networks", Madrid, Spain, ACM, February 2018, p. 173-174, https://hal.inria.fr/hal-01683605

[15] P. KATSIKOULI, M. S. ASTEFANOAEI, R. SARKAR.*Distributed Mining of Popular Paths in Road Networks*, in "DCOSS 2018 - International Conference on Distributed Computing in Sensor Systems", New York City, United States, IEEE, June 2018, p. 1-8, https://hal.inria.fr/hal-01768273

[16] J. A. KHAN, C. WESTPHAL, Y. GHAMRI-DOUDANE.*A Popularity-aware Centrality Metric for Content Placement in Information Centric Networks*, in "ICNC 2018 - International Conference on Computing, Networking and Communication", Maui, Hawaii, United States, March 2018, p. 1-7, https://hal.inria.fr/hal-01620062

[17] R. LI, C. ZHANG, P. PATRAS, R. STANICA, F. VALOIS.*Learning Driven Mobility Control of Airborne Base Stations in Emergency Networks*, in "WAIN 2018 - Workshop on Artificial Intelligence in Networks", Toulouse, France, December 2018, https://hal.inria.fr/hal-01927252

[18] A. A. MBACKÉ, N. MITTON, H. RIVANO.*DACAR: Distributed & Adaptable Crosslayer Anticollision and Routing protocol for RFID*, in "AdHoc-Now 2018 - 17th International Conference on Ad Hoc Networks and Wireless", St Malo, France, Springer, September 2018, p. 226-238 [*DOI :* 10.1007/978-3-030-00247-3_21], https://hal.inria.fr/hal-01819767

[19] A. A. MBACKÉ, N. MITTON, H. RIVANO.*Using Fuzzy Logic for data priority aware collection in RFID sensing wireless networks*, in "PIMRC 2018 - IEEE International Symposium on Personal, Indoor and Mobile Radio Communications", Bologne, Italy, September 2018, p. 1-5, https://hal.inria.fr/hal-01830721

[20] D. NABOULSI, A. MERMOURI, R. STANICA, H. RIVANO, M. FIORE.*On User Mobility in Dynamic Cloud Radio Access Networks*, in "INFOCOM 2018 – 37th Annual IEEE International Conference on Computer Communications", Honolulu, United States, IEEE, April 2018, p. 1-9, https://hal.inria.fr/hal-01767560

[21] J. OUEIS, R. STANICA, F. VALOIS.*Virtualized Local Core Network Functions Placement in Mobile Networks*, in "WCNC 2019 - IEEE Wireless Communications and Networking Conference", Marrakech, Morocco, IEEE, April 2019, p. 1-6, https://hal.inria.fr/hal-01988294

[22] M.-I. POPESCU, O. SIMONIN, A. SPALANZANI, F. VALOIS.*DMRR: Dynamic Multi-Robot Routing for Evolving Missions*, in "DMAP 2018 - 6th Workshop on Distributed and Multi-Agent Planning", Delft, Netherlands, June 2018, https://hal.inria.fr/hal-01942845

[23] Y. ZGUIRA, H. RIVANO, A. MEDDEB.*A Comparative Evaluation of the Performance of the multi-hop IoB-DTN routing protocol*, in "AdHoc Now 2018 - 17th International Conference on Ad Hoc Networks and Wireless", Saint Malo, France, September 2018, p. 1-12, https://hal.inria.fr/hal-01887573

[24] Y. ZGUIRA, H. RIVANO, A. MEDDEB.*For An Efficient Internet of Bikes : A DTN Routing Protocol Based On Data Aggregation Approach*, in "PEWASUN 2018 - 15th ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks", Montreal, Canada, October 2018 [*DOI :* 10.1145/3243046.3243048], https://hal.inria.fr/hal-01887567

[25] Y. ZGUIRA, H. RIVANO, A. MEDDEB.*IoB-DTN: a lightweight DTN protocol for mobile IoT Applications to smart bike sharing systems*, in "WD 2018 - Wirelesse Days", Dubai, United Arab Emirates, IEEE, April 2018, p. 131-136 [*DOI :* 10.1109/WD.2018.8361708], https://hal.archives-ouvertes.fr/hal-01772664

National Conferences with Proceeding

[26] A. A. MBACKÉ, N. MITTON, H. RIVANO.*Une vision plus durable et fiable de la ville intelligente grâce à la RFID*, in "CORES 2018 - Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication", Roscoff, France, May 2018, p. 1-4, https://hal.archives-ouvertes.fr/hal-01784572

[27] D. NABOULSI, A. MERMOURI, R. STANICA, H. RIVANO, M. FIORE.*Adaptation des Réseaux d'Accès Radio Virtualisés à la Mobilité des Utilisateurs*, in "ALGOTEL 2018 - 20èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications", Roscoff, France, May 2018, p. 1-4, https://hal.inria.fr/hal-01781620

Conferences without Proceedings

[28] C. ASCHAN-LEYGONIE, W. BECHKIT, A. BOUBRIMA, C. CHAPPAZ, I. CLOSTRE, M. DAHAN, A. FREI, C. HARPET, L. HERRMANN, B. LOEILLET, H. RIVANO, L. SOULHAC.*Réseaux de capteurs sans fil pour la caractérisation fine de la pollution de l'air*, in "CIVIC 2018 - Colloque International sur les Villes Intelligentes à Constantine", Constantine, Algeria, November 2018, https://hal.archives-ouvertes.fr/hal-01938583

[29] P. KATSIKOULI, A. CARNEIRO VIANA, M. FIORE, A. TARABLE.*L' étude de la fréquence d' échantillonnage des mouvements des humains*, in "CORES 2018 - Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication", Roscoff, France, May 2018, p. 1-4, https://hal.archives-ouvertes.fr/hal-01784475

Research Reports

[30] Y. ZGUIRA, H. RIVANO.*Performance evaluation of "Internet-of-Bikes" IoB-DTN routing protocol and IoB-Long range*, Institut national des sciences appliquées de Lyon, May 2018, p. 1-9, https://hal.inria.fr/hal-01803214

Project-Team AIRSEA

mathematics and computing applied to oceanic and atmospheric flows

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

IN PARTNERSHIP WITH:
Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Earth, Environmental and Energy Sciences

Table of contents

Project-Team AIRSEA

Creation of the Team: 2015 January 01, updated into Project-Team: 2016 April 01

Keywords:

Computer Science and Digital Science:

A3.1.8. - Big data (production, storage, transfer)
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling
A6.1.4. - Multiscale modeling
A6.1.5. - Multiphysics modeling
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A6.2.7. - High performance computing
A6.3.1. - Inverse problems
A6.3.2. - Data assimilation
A6.3.4. - Model reduction

Other Research Topics and Application Domains:

B3.2. - Climate and meteorology
B3.3.2. - Water: sea & ocean, lake & river
B3.3.4. - Atmosphere
B3.4.1. - Natural risks
B4.3.2. - Hydro-energy
B4.3.3. - Wind energy
B9.11.1. - Environmental risks

1. Team, Visitors, External Collaborators

Research Scientists

Laurent Debreu [Team leader, Inria, Researcher, HDR]
Christopher Eldred [Inria, Starting Research position]
Eugene Kazantsev [Inria, Researcher]
Florian Lemarié [Inria, Researcher]
Gurvan Madec [CNRS, Senior Researcher, from Dec 2018, HDR]
Farshid Nazari [Inria, Starting Research position]
Arthur Vidard [Inria, Researcher, HDR]
Olivier Zahm [Inria, Researcher, from Feb 2018]

Faculty Members

Elise Arnaud [Univ Grenoble Alpes, Associate Professor]
Eric Blayo-Nogret [Univ Grenoble Alpes, Professor, HDR]
Christine Kazantsev [Univ Grenoble Alpes, Associate Professor]
François-Xavier Le Dimet [Univ Grenoble Alpes, Emeritus, HDR]
Maëlle Nodet [Univ Grenoble Alpes, Associate Professor, HDR]
Clémentine Prieur [Univ Grenoble Alpes, Professor, HDR]

Technical Staff
Maurice Bremond [Inria]
Roland Patoum [Inria, until May 2018]
Laurence Viry [Univ Grenoble Alpes]

PhD Students
Charles Pelletier [Inria, until Feb 2018]
Mohamed Reda El Amri [IFPEN]
Maria Belen Heredia Guzman [IRSTEA]
Adrien Hirvoas [IFPEN, from Apr 2018]
Philomene Le Gall [Univ Grenoble Alpes, from Oct 2018]
Emilie Duval [Inria, from Oct 2018]
Long Long Li [China Scholarship Council]
Arthur Macherey [Ecole centrale de Nantes, from Apr 2018]
Sophie Thery [Univ Grenoble Alpes]
Victor Trappler [Univ Grenoble Alpes]

Post-Doctoral Fellows
Matthieu Brachet [Inria, from Nov 2018]
Charles Pelletier [Inria, from Apr 2018 until Jun 2018]

Visiting Scientist
Nicholas Kevlahan [McMaster Univ, Canada, from Sep 2018]

Administrative Assistant
Annie Simon [Inria]

2. Overall Objectives

2.1. Overall Objectives

The general scope of the AIRSEA project-team is to develop *mathematical and computational methods for the modeling of oceanic and atmospheric flows*. The mathematical tools used involve both *deterministic and statistical approaches*. The main research topics cover a) modeling and coupling b) model reduction for sensitivity analysis, coupling and multiscale optimizations c) sensitivity analysis, parameter estimation and risk assessment d) algorithms for high performance computing. The range of application is from climate modeling to the prediction of extreme events.

3. Research Program

3.1. Introduction

Recent events have raised questions regarding the social and economic implications of anthropic alterations of the Earth system, i.e. climate change and the associated risks of increasing extreme events. Ocean and atmosphere, coupled with other components (continent and ice) are the building blocks of the Earth system. A better understanding of the ocean atmosphere system is a key ingredient for improving prediction of such events. Numerical models are essential tools to understand processes, and simulate and forecast events at various space and time scales. Geophysical flows generally have a number of characteristics that make it difficult to model them. This justifies the development of specifically adapted mathematical methods:

- Geophysical flows are strongly non-linear. Therefore, they exhibit interactions between different scales, and unresolved small scales (smaller than mesh size) of the flows have to be **parameterized** in the equations.

- Geophysical fluids are non closed systems. They are open-ended in their scope for including and dynamically coupling different physical processes (e.g., atmosphere, ocean, continental water, etc). **Coupling** algorithms are thus of primary importance to account for potentially significant feedback.
- Numerical models contain parameters which cannot be estimated accurately either because they are difficult to measure or because they represent some poorly known subgrid phenomena. There is thus a need for **dealing with uncertainties**. This is further complicated by the turbulent nature of geophysical fluids.
- The computational cost of geophysical flow simulations is huge, thus requiring the use of **reduced models, multiscale methods** and the design of algorithms ready for **high performance computing** platforms.

Our scientific objectives are divided into four major points. The first objective focuses on developing advanced mathematical methods for both the ocean and atmosphere, and the coupling of these two components. The second objective is to investigate the derivation and use of model reduction to face problems associated with the numerical cost of our applications. The third objective is directed toward the management of uncertainty in numerical simulations. The last objective deals with efficient numerical algorithms for new computing platforms. As mentioned above, the targeted applications cover oceanic and atmospheric modeling and related extreme events using a hierarchy of models of increasing complexity.

3.2. Modeling for oceanic and atmospheric flows

Current numerical oceanic and atmospheric models suffer from a number of well-identified problems. These problems are mainly related to lack of horizontal and vertical resolution, thus requiring the parameterization of unresolved (subgrid scale) processes and control of discretization errors in order to fulfill criteria related to the particular underlying physics of rotating and strongly stratified flows. Oceanic and atmospheric coupled models are increasingly used in a wide range of applications from global to regional scales. Assessment of the reliability of those coupled models is an emerging topic as the spread among the solutions of existing models (e.g., for climate change predictions) has not been reduced with the new generation models when compared to the older ones.

Advanced methods for modeling 3D rotating and stratified flows The continuous increase of computational power and the resulting finer grid resolutions have triggered a recent regain of interest in numerical methods and their relation to physical processes. Going beyond present knowledge requires a better understanding of numerical dispersion/dissipation ranges and their connection to model fine scales. Removing the leading order truncation error of numerical schemes is thus an active topic of research and each mathematical tool has to adapt to the characteristics of three dimensional stratified and rotating flows. Studying the link between discretization errors and subgrid scale parameterizations is also arguably one of the main challenges.

Complexity of the geometry, boundary layers, strong stratification and lack of resolution are the main sources of discretization errors in the numerical simulation of geophysical flows. This emphasizes the importance of the definition of the computational grids (and coordinate systems) both in horizontal and vertical directions, and the necessity of truly multi resolution approaches. At the same time, the role of the small scale dynamics on large scale circulation has to be taken into account. Such parameterizations may be of deterministic as well as stochastic nature and both approaches are taken by the AIRSEA team. The design of numerical schemes consistent with the parameterizations is also arguably one of the main challenges for the coming years. This work is complementary and linked to that on parameters estimation described in 3.4.

Ocean Atmosphere interactions and formulation of coupled models State-of-the-art climate models (CMs) are complex systems under continuous development. A fundamental aspect of climate modeling is the representation of air-sea interactions. This covers a large range of issues: parameterizations of atmospheric and oceanic boundary layers, estimation of air-sea fluxes, time-space numerical schemes, non conforming grids, coupling algorithms ...Many developments related to these different aspects were performed over the last 10-15 years, but were in general conducted independently of each other.

The aim of our work is to revisit and enrich several aspects of the representation of air-sea interactions in CMs, paying special attention to their overall consistency with appropriate mathematical tools. We intend to work consistently on the physics and numerics. Using the theoretical framework of global-in-time Schwarz methods, our aim is to analyze the mathematical formulation of the parameterizations in a coupling perspective. From this study, we expect improved predictability in coupled models (this aspect will be studied using techniques described in 3.4). Complementary work on space-time nonconformities and acceleration of convergence of Schwarz-like iterative methods (see 6.1.1) are also conducted.

3.3. Model reduction / multiscale algorithms

The high computational cost of the applications is a common and major concern to have in mind when deriving new methodological approaches. This cost increases dramatically with the use of sensitivity analysis or parameter estimation methods, and more generally with methods that require a potentially large number of model integrations.

A dimension reduction, using either stochastic or deterministic methods, is a way to reduce significantly the number of degrees of freedom, and therefore the calculation time, of a numerical model.

Model reduction Reduction methods can be deterministic (proper orthogonal decomposition, other reduced bases) or stochastic (polynomial chaos, Gaussian processes, kriging), and both fields of research are very active. Choosing one method over another strongly depends on the targeted application, which can be as varied as real-time computation, sensitivity analysis (see e.g., section 6.3.1) or optimisation for parameter estimation (see below).

Our goals are multiple, but they share a common need for certified error bounds on the output. Our team has a 4-year history of working on certified reduction methods and has a unique positioning at the interface between deterministic and stochastic approaches. Thus, it seems interesting to conduct a thorough comparison of the two alternatives in the context of sensitivity analysis. Efforts will also be directed toward the development of efficient greedy algorithms for the reduction, and the derivation of goal-oriented sharp error bounds for non linear models and/or non linear outputs of interest. This will be complementary to our work on the deterministic reduction of parametrized viscous Burgers and Shallow Water equations where the objective is to obtain sharp error bounds to provide confidence intervals for the estimation of sensitivity indices.

Reduced models for coupling applications Global and regional high-resolution oceanic models are either coupled to an atmospheric model or forced at the air-sea interface by fluxes computed empirically preventing proper physical feedback between the two media. Thanks to high-resolution observational studies, the existence of air-sea interactions at oceanic mesoscales (i.e., at $\mathcal{O}(1km)$ scales) have been unambiguously shown. Those interactions can be represented in coupled models only if the oceanic and atmospheric models are run on the same high-resolution computational grid, and are absent in a forced mode. Fully coupled models at high-resolution are seldom used because of their prohibitive computational cost. The derivation of a reduced model as an alternative between a forced mode and the use of a full atmospheric model is an open problem.

Multiphysics coupling often requires iterative methods to obtain a mathematically correct numerical solution. To mitigate the cost of the iterations, we will investigate the possibility of using reduced-order models for the iterative process. We will consider different ways of deriving a reduced model: coarsening of the resolution, degradation of the physics and/or numerical schemes, or simplification of the governing equations. At a mathematical level, we will strive to study the well-posedness and the convergence properties when reduced models are used. Indeed, running an atmospheric model at the same resolution as the ocean model is generally too expensive to be manageable, even for moderate resolution applications. To account for important fine-scale interactions in the computation of the air-sea boundary condition, the objective is to derive a simplified boundary layer model that is able to represent important 3D turbulent features in the marine atmospheric boundary layer.

Reduced models for multiscale optimization The field of multigrid methods for optimisation has known a tremendous development over the past few decades. However, it has not been applied to oceanic and atmospheric problems apart from some crude (non-converging) approximations or applications to simplified and low dimensional models. This is mainly due to the high complexity of such models and to the difficulty in handling several grids at the same time. Moreover, due to complex boundaries and physical phenomena, the grid interactions and transfer operators are not trivial to define.

Multigrid solvers (or multigrid preconditioners) are efficient methods for the solution of variational data assimilation problems. We would like to take advantage of these methods to tackle the optimization problem in high dimensional space. High dimensional control space is obtained when dealing with parameter fields estimation, or with control of the full 4D (space time) trajectory. It is important since it enables us to take into account model errors. In that case, multigrid methods can be used to solve the large scales of the problem at a lower cost, this being potentially coupled with a scale decomposition of the variables themselves.

3.4. Dealing with uncertainties

There are many sources of uncertainties in numerical models. They are due to imperfect external forcing, poorly known parameters, missing physics and discretization errors. Studying these uncertainties and their impact on the simulations is a challenge, mostly because of the high dimensionality and non-linear nature of the systems. To deal with these uncertainties we work on three axes of research, which are linked: sensitivity analysis, parameter estimation and risk assessment. They are based on either stochastic or deterministic methods.

Sensitivity analysis Sensitivity analysis (SA), which links uncertainty in the model inputs to uncertainty in the model outputs, is a powerful tool for model design and validation. First, it can be a pre-stage for parameter estimation (see 3.4), allowing for the selection of the more significant parameters. Second, SA permits understanding and quantifying (possibly non-linear) interactions induced by the different processes defining e.g., realistic ocean atmosphere models. Finally SA allows for validation of models, checking that the estimated sensitivities are consistent with what is expected by the theory. On ocean, atmosphere and coupled systems, only first order deterministic SA are performed, neglecting the initialization process (data assimilation). AIRSEA members and collaborators proposed to use second order information to provide consistent sensitivity measures, but so far it has only been applied to simple academic systems. Metamodels are now commonly used, due to the cost induced by each evaluation of complex numerical models: mostly Gaussian processes, whose probabilistic framework allows for the development of specific adaptive designs, and polynomial chaos not only in the context of intrusive Galerkin approaches but also in a black-box approach. Until recently, global SA was based primarily on a set of engineering practices. New mathematical and methodological developments have led to the numerical computation of Sobol' indices, with confidence intervals assessing for both metamodel and estimation errors. Approaches have also been extended to the case of dependent entries, functional inputs and/or output and stochastic numerical codes. Other types of indices and generalizations of Sobol' indices have also been introduced.

Concerning the stochastic approach to SA we plan to work with parameters that show spatio-temporal dependencies and to continue toward more realistic applications where the input space is of huge dimension with highly correlated components. Sensitivity analysis for dependent inputs also introduces new challenges. In our applicative context, it would seem prudent to carefully learn the spatio-temporal dependences before running a global SA. In the deterministic framework we focus on second order approaches where the sought sensitivities are related to the optimality system rather than to the model; i.e., we consider the whole forecasting system (model plus initialization through data assimilation).
All these methods allow for computing sensitivities and more importantly a posteriori error statistics.

Parameter estimation Advanced parameter estimation methods are barely used in ocean, atmosphere and coupled systems, mostly due to a difficulty of deriving adequate response functions, a lack of knowledge of these methods in the ocean-atmosphere community, and also to the huge associated computing costs. In the presence of strong uncertainties on the model but also on parameter values, simulation and inference are closely associated. Filtering for data assimilation and Approximate Bayesian Computation (ABC) are two examples of such association.

Stochastic approach can be compared with the deterministic approach, which allows to determine the sensitivity of the flow to parameters and optimize their values relying on data assimilation. This approach is already shown to be capable of selecting a reduced space of the most influent parameters in the local parameter space and to adapt their values in view of correcting errors committed by the numerical approximation. This approach assumes the use of automatic differentiation of the source code with respect to the model parameters, and optimization of the obtained raw code.

AIRSEA assembles all the required expertise to tackle these difficulties. As mentioned previously, the choice of parameterization schemes and their tuning has a significant impact on the result of model simulations. Our research will focus on parameter estimation for parameterized Partial Differential Equations (PDEs) and also for parameterized Stochastic Differential Equations (SDEs). Deterministic approaches are based on optimal control methods and are local in the parameter space (i.e., the result depends on the starting point of the estimation) but thanks to adjoint methods they can cope with a large number of unknowns that can also vary in space and time. Multiscale optimization techniques as described in 6.2 will be one of the tools used. This in turn can be used either to propose a better (and smaller) parameter set or as a criterion for discriminating parameterization schemes. Statistical methods are global in the parameter state but may suffer from the curse of dimensionality. However, the notion of parameter can also be extended to functional parameters. We may consider as parameter a functional entity such as a boundary condition on time, or a probability density function in a stationary regime. For these purposes, non-parametric estimation will also be considered as an alternative.

Risk assessment Risk assessment in the multivariate setting suffers from a lack of consensus on the choice of indicators. Moreover, once the indicators are designed, it still remains to develop estimation procedures, efficient even for high risk levels. Recent developments for the assessment of financial risk have to be considered with caution as methods may differ pertaining to general financial decisions or environmental risk assessment. Modeling and quantifying uncertainties related to extreme events is of central interest in environmental sciences. In relation to our scientific targets, risk assessment is very important in several areas: hydrological extreme events, cyclone intensity, storm surges...Environmental risks most of the time involve several aspects which are often correlated. Moreover, even in the ideal case where the focus is on a single risk source, we have to face the temporal and spatial nature of environmental extreme events. The study of extremes within a spatio-temporal framework remains an emerging field where the development of adapted statistical methods could lead to major progress in terms of geophysical understanding and risk assessment thus coupling data and model information for risk assessment.

Based on the above considerations we aim to answer the following scientific questions: how to measure risk in a multivariate/spatial framework? How to estimate risk in a non stationary context? How to reduce dimension (see 3.3) for a better estimation of spatial risk?

Extreme events are rare, which means there is little data available to make inferences of risk measures. Risk assessment based on observation therefore relies on multivariate extreme value theory. Interacting particle systems for the analysis of rare events is commonly used in the community of computer experiments. An open question is the pertinence of such tools for the evaluation of environmental risk.
Most numerical models are unable to accurately reproduce extreme events. There is therefore a real need to develop efficient assimilation methods for the coupling of numerical models and extreme data.

3.5. High performance computing

Methods for sensitivity analysis, parameter estimation and risk assessment are extremely costly due to the necessary number of model evaluations. This number of simulations require considerable computational resources, depends on the complexity of the application, the number of input variables and desired quality of approximations. To this aim, the AIRSEA team is an intensive user of HPC computing platforms, particularly grid computing platforms. The associated grid deployment has to take into account the scheduling of a huge number of computational requests and the links with data-management between these requests, all of these as automatically as possible. In addition, there is an increasing need to propose efficient numerical algorithms specifically designed for new (or future) computing architectures and this is part of our scientific objectives.

According to the computational cost of our applications, the evolution of high performance computing platforms has to be taken into account for several reasons. While our applications are able to exploit space parallelism to its full extent (oceanic and atmospheric models are traditionally based on a spatial domain decomposition method), the spatial discretization step size limits the efficiency of traditional parallel methods. Thus the inherent parallelism is modest, particularly for the case of relative coarse resolution but with very long integration time (e.g., climate modeling). Paths toward new programming paradigms are thus needed. As a step in that direction, we plan to focus our research on parallel in time methods.

New numerical algorithms for high performance computing Parallel in time methods can be classified into three main groups. In the first group, we find methods using parallelism across the method, such as parallel integrators for ordinary differential equations. The second group considers parallelism across the problem. Falling into this category are methods such as waveform relaxation where the space-time system is decomposed into a set of subsystems which can then be solved independently using some form of relaxation techniques or multigrid reduction in time. The third group of methods focuses on parallelism across the steps. One of the best known algorithms in this family is parareal. Other methods combining the strengths of those listed above (e.g., PFASST) are currently under investigation in the community.

Parallel in time methods are iterative methods that may require a large number of iteration before convergence. Our first focus will be on the convergence analysis of parallel in time (Parareal / Schwarz) methods for the equation systems of oceanic and atmospheric models. Our second objective will be on the construction of fast (approximate) integrators for these systems. This part is naturally linked to the model reduction methods of section (6.2.1). Fast approximate integrators are required both in the Schwarz algorithm (where a first guess of the boundary conditions is required) and in the Parareal algorithm (where the fast integrator is used to connect the different time windows). Our main application of these methods will be on climate (i.e., very long time) simulations. Our second application of parallel in time methods will be in the context of optimization methods. In fact, one of the major drawbacks of the optimal control techniques used in 3.4 is a lack of intrinsic parallelism in comparison with ensemble methods. Here, parallel in time methods also offer ways to better efficiency. The mathematical key point is centered on how to efficiently couple two iterative methods (i.e., parallel in time and optimization methods).

4. Application Domains

4.1. The Ocean-Atmosphere System

The evolution of natural systems, in the short, mid, or long term, has extremely important consequences for both the global Earth system and humanity. Forecasting this evolution is thus a major challenge from the scientific, economic, and human viewpoints.

Humanity has to face the problem of **global warming**, brought on by the emission of greenhouse gases from human activities. This warming will probably cause huge changes at global and regional scales, in terms of climate, vegetation and biodiversity, with major consequences for local populations. Research has therefore been conducted over the past 15 to 20 years in an effort to model the Earth's climate and forecast its evolution in the 21st century in response to anthropic action.

With regard to short-term forecasts, the best and oldest example is of course **weather forecasting**. Meteorological services have been providing daily short-term forecasts for several decades which are of crucial importance for numerous human activities.

Numerous other problems can also be mentioned, like **seasonal weather forecasting** (to enable powerful phenomena like an El Niño event or a drought period to be anticipated a few months in advance), **operational oceanography** (short-term forecasts of the evolution of the ocean system to provide services for the fishing industry, ship routing, defense, or the fight against marine pollution) or the prediction of **floods**.

As mentioned previously, mathematical and numerical tools are omnipresent and play a fundamental role in these areas of research. In this context, the vocation of AIRSEA is not to carry out numerical prediction, but to address mathematical issues raised by the development of prediction systems for these application fields, in close collaboration with geophysicists.

5. New Software and Platforms

5.1. AGRIF

Adaptive Grid Refinement In Fortran
KEYWORD: Mesh refinement
SCIENTIFIC DESCRIPTION: AGRIF is a Fortran 90 package for the integration of full adaptive mesh refinement (AMR) features within a multidimensional finite difference model written in Fortran. Its main objective is to simplify the integration of AMR potentialities within an existing model with minimal changes. Capabilities of this package include the management of an arbitrary number of grids, horizontal and/or vertical refinements, dynamic regridding, parallelization of the grids interactions on distributed memory computers. AGRIF requires the model to be discretized on a structured grid, like it is typically done in ocean or atmosphere modelling.
NEWS OF THE YEAR: In 2017, the multiresolution capabilities of the AGRIF software have been extended to be able to treat a much larger number of grids. In particular, the load balancing algorithms have been greatly improved.

- Participants: Roland Patoum and Laurent Debreu
- Contact: Laurent Debreu
- Publications: Numerical and experimental approach for a better physical description of submesoscale processes : A north-western Mediterranean Sea case - AGRIF: Adaptive Grid Refinement in Fortran
- URL: http://www-ljk.imag.fr/MOISE/AGRIF

5.2. BALAISE

Bilbliothèque d'Assimilation Lagrangienne Adaptée aux Images Séquencées en Environnement
KEYWORDS: Multi-scale analysis - Data assimilation - Optimal control
FUNCTIONAL DESCRIPTION: BALAISE (Bilbliothèque d'Assimilation Lagrangienne Adaptée aux Images Séquencées en Environnement) is a test bed for image data assimilation. It includes a shallow water model, a multi-scale decomposition library and an assimilation suite.

- Contact: Patrick Vidard

5.3. NEMOVAR

Variational data assimilation for NEMO
KEYWORDS: Oceanography - Data assimilation - Adjoint method - Optimal control
FUNCTIONAL DESCRIPTION: NEMOVAR is a state-of-the-art multi-incremental variational data assimilation system with both 3D and 4D var capabilities, and which is designed to work with NEMO on the native ORCA grids. The background error covariance matrix is modelled using balance operators for the multivariate component and a diffusion operator for the univariate component. It can also be formulated as a linear combination of covariance models to take into account multiple correlation length scales associated with ocean variability on different scales. NEMOVAR has recently been enhanced with the addition of ensemble data assimilation and multi-grid assimilation capabilities. It is used operationnaly in both ECMWF and the Met Office (UK)

- Partners: CERFACS - ECMWF - Met Office
- Contact: Patrick Vidard

5.4. Sensitivity

FUNCTIONAL DESCRIPTION: This package is useful for conducting sensitivity analysis of complex computer codes.

- Contact: Laurent Gilquin
- URL: https://cran.r-project.org/web/packages/sensitivity/index.html

6. New Results

6.1. Modeling for Oceanic and Atmospheric flows

6.1.1. Coupling Methods for Oceanic and Atmospheric Models

Participants: Eric Blayo, Florian Lemarié, Sophie Thery.

Coupling methods routinely used in regional and global climate models do not provide the exact solution to the ocean-atmosphere problem, but an approximation of one [72]. For the last few years we have been actively working on the analysis of ocean-atmosphere coupling both in terms of its continuous and numerical formulation.Our activities over the last few years can be divided into four general topics

1. *Stability and consistency analysis of existing coupling methods*: in [72] we showed that the usual methods used in the context of ocean-atmosphere coupling are prone to splitting errors because they correspond to only one iteration of an iterative process without reaching convergence. Moreover, those methods have an additional condition for the coupling to be stable even if unconditionally stable time stepping algorithms are used. This last remark was further studied recently in [47] and it turned out to be a major source of instability in atmosphere-snow coupling.
2. *Study of physics-dynamics coupling*: during the PhD-thesis of Charles Pelletier (funded by Inria and defended on Feb. 15, 2018, [2]) the scope was on including the formulation of physical parameterizations in the theoretical analysis of the coupling, in particular the parameterization schemes to compute air-sea fluxes [79]. To do so, a metamodel representative of the behavior of the full parameterization but with a continuous form easier to manipulate has been derived thanks to a sensitivity analysis. This metamodel is more adequate to conduct the mathematical analysis of the coupling while being physically satisfactory [80]. In parallel we have contributed to a general review gathering the main international specialists on the topic [64]. More recently we have started to work specifically on the discretization methods for the parameterization of planetary boundary layers in climate models [27] which takes the form of an nonstationary nonlinear parabolic equation. The objective is to derive a discretization for which we could prove nonlinear stability criteria and show robustness to large variations in parabolic Courant number while being consistent with our knowledge of the underlying physical principles (e.g. the Monin-Obukhov theory in the surface layer).
3. *Design of a coupled single column model*: in order to focus on specific problems of ocean-atmosphere coupling, a work on simplified equation sets has been started. The aim is to implement a one-dimensional (in the vertical direction) coupled model with physical parameterizations representative of those used in realistic models. Thanks to this simplified coupled model the objective is to develop a benchmark suite for coupled models evaluation. Last year the single column oceanic and atmospheric components have been developed and coupled during the PhD-thesis of Rémi Pellerej (defended on Mar. 26, 2018) and in the framework of the SIMBAD project [17]. A publication describing this model is currently in preparation for the Geoscientific Model Development journal.

4. *Analysis of air-sea-wave interactions in realistic high-resolution realistic simulations*: part of our activity has been in collaboration with atmosphericists and physical oceanographers to study the impact on some modeling assumptions (e.g. [73]) in large-scale realistic ocean-atmosphere coupled simulations [14]. Moreover, within the ALBATROS project, we have contributed to the development of a 2-way coupling between an ocean global circulation model (NEMO) with a surface wave model (WW3). Such coupling is not straightforward to implement since it requires modifications of the governing equations, boundary conditions and subgrid scale closures in the oceanic model. A paper is currently under review in Geoscientific Model Development journal on that topic.
5. *Efficient coupling methods*: we have been developing coupling approaches for several years, based on so-called Schwarz algorithms. In particular, we addressed the development of efficient interface conditions for multi-physics problems representative of air-sea coupling [28] (paper in preparation). This work is done in the framework of S. Théry PhD (started in fall 2017).

These topics are addressed through strong collaborations between the applied mathematicians and the climate community (Meteo-France, Ifremer, LMD, and LOCEAN). Indeed, Our work on ocean-atmosphere coupling has steadily matured over the last few years and has reached a point where it triggered interest from the climate community. Through the funding of the COCOA ANR project (started in January 2017, PI: E. Blayo), Airsea team members play a major role in the structuration of a multi-disciplinary scientific community working on ocean-atmosphere coupling spanning a broad range from mathematical theory to practical implementations in climate models. An expected outcome of this project should be the design of a benchmark suite of idealized coupled test cases representative of known issues in coupled models. Such idealized test cases should motivate further collaborations at an international level.

6.1.2. Numerical Schemes for Ocean Modelling

Participants: Eric Blayo, Matthieu Brachet, Laurent Debreu, Emilie Duval, Nicholas Kevlahan, Florian Lemarié, Christopher Eldred, Farshid Nazari.

The increase of model resolution naturally leads to the representation of a wider energy spectrum. As a result, in recent years, the understanding of oceanic submesoscale dynamics has significantly improved. However, dissipation in submesoscale models remains dominated by numerical constraints rather than physical ones. Effective resolution is limited by the numerical dissipation range, which is a function of the model numerical filters (assuming that dispersive numerical modes are efficiently removed). A review paper on coastal ocean models has been written with German colleagues and has been published in Ocean Modelling ([11]).
F. Lemarié and L. Debreu (with H. Burchard, K. Klingbeil and J. Sainte-Marie) have organized the international COMMODORE workshop on numerical methods for oceanic models (Paris, Sept. 17-19, 2018). https://commodore2018.sciencesconf.org/, see [12] for a summary of the scientific discussions

With the increase of resolution, the hydrostatic assumption becomes less valid and the AIRSEA group also works on the development of non-hydrostatic ocean models. The treatment of non-hydrostatic incompressible flows leads to a 3D elliptic system for pressure that can be ill conditioned in particular with non geopotential vertical coordinates. That is why we flavor the use of the non-hydrostatic compressible equations that removes the need for a 3D resolution at the price of reincluding acoustic waves [24].
In addition, Emilie Duval started her PhD in September 2018 on the coupling between the hydrostatic incompressible and non-hydrostatic compressible equations.
The team is involved in the HEAT (Highly Efficient ATmospheric Modelling) ANR project. This project aims at developing a new atmospheric dynamical core (DYNAMICO) discretized on an icosahedral grid. This project is in collaboration with Ecole Polytechnique, Meteo-France, LMD, LSCE and CERFACS. This year we worked on dispersion analysis of compatible Galerkin schemes for shallow water model both in 1D ([5]) and 2D ([39]). In addition, we worked on the discrete formulation of the thermal rotating shallow water equations. This formulation, based on quasi-Hamiltonian discretizations methods, allows for the first time total mass, buoyancy and energy conservation to machine precision ([4]).

6.1.3. Data assimilation for coupled models

In the context of operational meteorology and oceanography, forecast skills heavily rely on proper combination of model prediction and available observations via data assimilation techniques. Historically, numerical weather prediction is made separately for the ocean and the atmosphere in an uncoupled way. However, in recent years, fully coupled ocean-atmosphere models are increasingly used in operational centers to improve the reliability of seasonal forecasts and tropical cyclones predictions. For coupled problems, the use of separated data assimilation schemes in each medium is not satisfactory since the result of such assimilation process is generally inconsistent across the interface, thus leading to unacceptable artefacts. Hence, there is a strong need for adapting existing data assimilation techniques to the coupled framework. As part of our ERACLIM2 contribution, R. Pellerej started a PhD on that topic late 2014 and defended it early 2018 [1]. Three general data assimilation algorithms, based on variational data assimilation techniques, have been developed and applied to a single column coupled model. The dynamical equations of the considered problem are coupled using an iterative Schwarz domain decomposition method. The aim is to properly take into account the coupling in the assimilation process in order to obtain a coupled solution close to the observations while satisfying the physical conditions across the air-sea interface. Results shows significant improvement compared to the usual approach on this simple system. The aforementioned system has been coded within the OOPS framework (Object Oriented Prediction System) in order to ease the transfer to more complex/realistic models.

Finally, CASIS, a new collaborative project with Mercator Océan has started late 2017 in order to extend developments to iterative Kalman smoother data assimilation scheme, in the framework of a coupled ocean-atmospheric boundary layer context.

6.1.4. Optimal control of grids and schemes for ocean model.

Participants: Laurent Debreu, Eugene Kazantsev.

In [33], variational data assimilation technique is applied to a simple bidimensional wave equation that simulates propagation of internal gravity waves in the ocean in order to control grids and numerical schemes. Grid steps of the vertical grid, Brunt-Vaisala frequency and approximation of the horizontal derivative were used as control parameters either separately or in the joint control. Obtained results show that optimized parameters may partially compensate errors committed by numerical scheme due to insufficient grid resolution.

Optimal vertical grid steps and coefficients in horizontal derivative approximation found in the variational control procedure allow us to get the model solution that is rather close to the solution of the reference model. The error in the wave velocity on the coarse grid is mostly compensated in experiments with joint control of parameters while the error in the wave amplitude occurs to be more difficult to correct.

However, optimal grid steps and discretization schemes may be in a disagreement with requirements of other model physics and additional analysis of obtained optimized parameters from the point of view of they agreement with the model is necessary.

6.1.5. Nonhydrostatic Modeling

Participants: Eric Blayo, Laurent Debreu, Emilie Duval.

In the context of the French initiative CROCO (Coastal and Regional Ocean COmmunity model, https://www.croco-ocean.org) for the development of a new oceanic modeling system, Emilie Duval started a PhD (Oct. 2018) focused on the design of methods to couple local nonhydrostatic models to larger scale hydrostatic ones. Such a coupling is quite delicate from a mathematical point of view, due to the different nature of hydrostatic and nonhydrostatic equations (where the vertical velocity is either a diagnostic or a prognostic variable).

6.2. Model reduction / multiscale algorithms

6.2.1. Model Order Reduction

Participants: Mohamed Reda El Amri, Youssef Marzouk, Maëlle Nodet, Clémentine Prieur, Alessio Spantini, Olivier Zahm.

Another point developed in the team for sensitivity analysis is model reduction. To be more precise regarding model reduction, the aim is to reduce the number of unknown variables (to be computed by the model), using a well chosen basis. Instead of discretizing the model over a huge grid (with millions of points), the state vector of the model is projected on the subspace spanned by this basis (of a far lesser dimension). The choice of the basis is of course crucial and implies the success or failure of the reduced model. Various model reduction methods offer various choices of basis functions. A well-known method is called “proper orthogonal decomposition" or “principal component analysis". More recent and sophisticated methods also exist and may be studied, depending on the needs raised by the theoretical study. Model reduction is a natural way to overcome difficulties due to huge computational times due to discretizations on fine grids. In [68], the authors present a reduced basis offline/online procedure for viscous Burgers initial boundary value problem, enabling efficient approximate computation of the solutions of this equation for parametrized viscosity and initial and boundary value data. This procedure comes with a fast-evaluated rigorous error bound certifying the approximation procedure. The numerical experiments in the paper show significant computational savings, as well as efficiency of the error bound.

When a metamodel is used (for example reduced basis metamodel, but also kriging, regression, ...) for estimating sensitivity indices by Monte Carlo type estimation, a twofold error appears: a sampling error and a metamodel error. Deriving confidence intervals taking into account these two sources of uncertainties is of great interest. We obtained results particularly well fitted for reduced basis metamodels [69]. In [66], the authors provide asymptotic confidence intervals in the double limit where the sample size goes to infinity and the metamodel converges to the true model. These results were also adapted to problems related to more general models such as Shallow-Water equations, in the context of the control of an open channel [70].

When considering parameter-dependent PDE, it happens that the quantity of interest is not the PDE’s solution but a linear functional of it. In [67], we have proposed a probabilistic error bound for the reduced output of interest (goal-oriented error bound). By probabilistic we mean that this bound may be violated with small probability. The bound is efficiently and explicitly computable, and we show on different examples that this error bound is sharper than existing ones.

A collaboration has been started with Christophe Prieur (Gipsa-Lab) on the very challenging issue of sensitivity of a controlled system to its control parameters [70]. In [71], we propose a generalization of the probabilistic goal-oriented error estimation in [67] to parameter-dependent nonlinear problems. One aims at applying such results in the previous context of sensitivity of a controlled system.

More recently, in the context of the Inria associate team UNQUESTIONABLE, we have extended the focus of the axis on model order reduction. Our objectives are to understand the kinds of low-dimensional structure that may be present in important geophysical models; and to exploit this low-dimensional structure in order to extend Bayesian approaches to high-dimensional inverse problems, such as those encountered in geophysical applications. Our recent and future efforts are/will be concerned with parameter space dimension reduction techniques, low- rank structures in geophysical models and transport maps tools for probability measure approximation. At the moment, scientific progress has been achieved in different directions, as detailed below: A first paper [45] has been submitted on gradient- based dimension reduction of vector-valued functions. Multivariate functions encountered in high-dimensional uncertainty quantification problems often vary most strongly along a few dominant directions in the input parameter space. In this work, we propose a gradient-based method for detecting these directions and using them to construct ridge approximations of such functions, in the case where the functions are vector-valued. The methodology consists of minimizing an upper bound on the approximation error, obtained by subspace Poincaré inequalities. We have provided a thorough mathematical analysis in the case where the parameter space is equipped with a Gaussian probability measure. A second work [46] has been submitted, which proposes a dimension reduction technique for Bayesian inverse problems with nonlinear forward operators, non-Gaussian priors, and non-Gaussian observation noise. In this work, the likelihood function is approximated by a ridge function, i.e., a map which depends non-trivially only on a few linear combinations of the parameters. The ridge approximation is built by minimizing an upper bound on the Kullback-Leibler divergence between the posterior distribution and its approximation. This bound, obtained via logarithmic Sobolev inequalities, allows one to certify the error of the posterior approximation. A sample-based approximation of the upper bound is also proposed. In the framework of the PhD thesis of Reda El Amri, a work on data-driven stochastic inversion via functional quantization was submitted. In this paper [36], a new methodology is proposed for solving stochastic inversion problems through computer experiments, the stochasticity being driven by functional random variables. Main tools are a new greedy algorithm for functional quantization, and the adaptation of Stepwise Uncertainty Reduction techniques.

6.3. Dealing with uncertainties

6.3.1. Sensitivity Analysis

Participants: Elise Arnaud, Eric Blayo, Laurent Gilquin, Maria Belén Heredia, François-Xavier Le Dimet, Clémentine Prieur, Laurence Viry.

6.3.1.1. Scientific context

Forecasting geophysical systems require complex models, which sometimes need to be coupled, and which make use of data assimilation. The objective of this project is, for a given output of such a system, to identify the most influential parameters, and to evaluate the effect of uncertainty in input parameters on model output. Existing stochastic tools are not well suited for high dimension problems (in particular time-dependent problems), while deterministic tools are fully applicable but only provide limited information. So the challenge is to gather expertise on one hand on numerical approximation and control of Partial Differential Equations, and on the other hand on stochastic methods for sensitivity analysis, in order to develop and design innovative stochastic solutions to study high dimension models and to propose new hybrid approaches combining the stochastic and deterministic methods.

6.3.2. Extensions of the replication method for the estimation of Sobol' indices

Participants: Elise Arnaud, Eric Blayo, Laurent Gilquin, Alexandre Janon, Clémentine Prieur.

Sensitivity analysis studies how the uncertainty on an output of a mathematical model can be attributed to sources of uncertainty among the inputs. Global sensitivity analysis of complex and expensive mathematical models is a common practice to identify influent inputs and detect the potential interactions between them. Among the large number of available approaches, the variance-based method introduced by Sobol' allows to calculate sensitivity indices called Sobol' indices. Each index gives an estimation of the influence of an individual input or a group of inputs. These indices give an estimation of how the output uncertainty can be apportioned to the uncertainty in the inputs. One can distinguish first-order indices that estimate the main effect from each input or group of inputs from higher-order indices that estimate the corresponding order of interactions between inputs. This estimation procedure requires a significant number of model runs, number that has a polynomial growth rate with respect to the input space dimension. This cost can be prohibitive for time consuming models and only a few number of runs is not enough to retrieve accurate informations about the model inputs.

The use of replicated designs to estimate first-order Sobol' indices has the major advantage of reducing drastically the estimation cost as the number of runs n becomes independent of the input space dimension. The generalization to closed second-order Sobol' indices relies on the replication of randomized orthogonal arrays. However, if the input space is not properly explored, that is if n is too small, the Sobol' indices estimates may not be accurate enough. Gaining in efficiency and assessing the estimate precision still remains an issue, all the more important when one is dealing with limited computational budget.

We designed an approach to render the replication method iterative, enabling the required number of evaluations to be controlled. With this approach, more accurate Sobol' estimates are obtained while recycling previous sets of model evaluations. Its main characteristic is to rely on iterative construction of stratified designs, latin hypercubes and orthogonal arrays [61]

In [7] a new strategy to estimate the full set of first-order and second-order Sobol' indices with only two replicated designs based on orthogonal arrays of strength two. Such a procedure increases the precision of the estimation for a given computation budget. A bootstrap procedure for producing confidence intervals, that are compared to asymptotic ones in the case of first-order indices, is also proposed.

The replicated designs strategy for global sensitivity analysis was also implemented in the applied framework of marine biogeochemical modeling, making use of distributed computing environments [43].

6.3.3. Sensitivity analysis with dependent inputs

An important challenge for stochastic sensitivity analysis is to develop methodologies which work for dependent inputs. For the moment, there does not exist conclusive results in that direction. Our aim is to define an analogue of Hoeffding decomposition [65] in the case where input parameters are correlated. Clémentine Prieur supervised Gaëlle Chastaing's PhD thesis on the topic (defended in September 2013) [53]. We obtained first results [54], deriving a general functional ANOVA for dependent inputs, allowing defining new variance based sensitivity indices for correlated inputs. We then adapted various algorithms for the estimation of these new indices. These algorithms make the assumption that among the potential interactions, only few are significant. Two papers have been recently accepted [52], [55]. We also considered the estimation of groups Sobol' indices, with a procedure based on replicated designs [63]. These indices provide information at the level of groups, and not at a finer level, but their interpretation is still rigorous.

Céline Helbert and Clémentine Prieur supervised the PhD thesis of Simon Nanty (funded by CEA Cadarache, and defended in October, 2015). The subject of the thesis is the analysis of uncertainties for numerical codes with temporal and spatio-temporal input variables, with application to safety and impact calculation studies. This study implied functional dependent inputs. A first step was the modeling of these inputs [75]. The whole methodology proposed during the PhD is presented in [76].

More recently, the Shapley value, from econometrics, was proposed as an alternative to quantify the importance of random input variables to a function. Owen [77] derived Shapley value importance for independent inputs and showed that it is bracketed between two different Sobol' indices. Song et al. [82] recently advocated the use of Shapley value for the case of dependent inputs. In a very recent work [78], in collaboration with Art Owen (Standford's University), we show that Shapley value removes the conceptual problems of functional ANOVA for dependent inputs. We do this with some simple examples where Shapley value leads to intuitively reasonable nearly closed form values. We also investigated further the properties of Shapley effects in [41].

6.3.4. Global sensitivity analysis for parametrized stochastic differential equations

Participant: Clémentine Prieur.

Many models are stochastic in nature, and some of them may be driven by parametrized stochastic differential equations. It is important for applications to propose a strategy to perform global sensitivity analysis for such models, in presence of uncertainties on the parameters. In collaboration with Pierre Etoré (DATA department in Grenoble), Clémentine Prieur proposed an approach based on Feynman-Kac formulas [40].

6.3.5. Parameter control in presence of uncertainties: robust estimation of bottom friction

Participants: Victor Trappler, Elise Arnaud, Laurent Debreu, Arthur Vidard.

Many physical phenomena are modelled numerically in order to better understand and/or to predict their behaviour. However, some complex and small scale phenomena can not be fully represented in the models. The introduction of ad-hoc correcting terms, can represent these unresolved processes, but they need to be properly estimated.

A good example of this type of problem is the estimation of bottom friction parameters of the ocean floor. This is important because it affects the general circulation. This is particularly the case in coastal areas, especially for its influence on wave breaking. Because of its strong spatial disparity, it is impossible to estimate the bottom friction by direct observation, so it requires to do so indirectly by observing its effects on surface movement. This task is further complicated by the presence of uncertainty in certain other characteristics linking the bottom and the surface (eg boundary conditions). The techniques currently used to adjust these settings are very basic and do not take into account these uncertainties, thereby increasing the error in this estimate.

Classical methods of parameter estimation usually imply the minimisation of an objective function, that measures the error between some observations and the results obtained by a numerical model. In the presence of uncertainties, the minimisation is not straightforward, as the output of the model depends on those uncontrolled inputs and on the control parameter as well. That is why we will aim at minimising the objective function, to get an estimation of the control parameter that is robust to the uncertainties.

The definition of robustness differs depending of the context in which it is used. In this work, two different notions of robustness are considered: robustness by minimising the mean and variance, and robustness based on the distribution of the minimisers of the function. This information on the location of the minimisers is not a novel idea, as it had been applied as a criterion in sequential Bayesian optimisation. However, the constraint of optimality is here relaxed to define a new estimate. To evaluate this estimation, a toy model of a coastal area has been implemented. The control parameter is the bottom friction, upon which classical methods of estimation are applied in a simulation-estimation experiment. The model is then modified to include uncertainties on the boundary conditions in order to apply robust control methods.

6.3.6. Development of a data assimilation method for the calibration and continuous update of wind turbines digital twins

Participants: Adrien Hirvoas, Elise Arnaud, Clémentine Prieur, Arthur Vidard.

In the context of the energy transition, wind power generation is developing rapidly in France and worldwide. Research and innovation on wind resource characterisation, turbin control, coupled mechanical modelling of wind systems or technological development of offshore wind turbines floaters are current research topics.

In particular, the monitoring and the maintenance of wind turbine is becoming a major issue. Current solutions do not take full advantage of the large amount of data provided by sensors placed on modern wind turbines in production. These data could be advantageously used in order to refine the predictions of production, the life of the structure, the control strategies and the planning of maintenance. In this context, it is interesting to optimally combine production data and numerical models in order to obtain highly reliable models of wind turbines. This process is of interest to many industrial and academic groups and is known in many fields of the industry, including the wind industry, as "digital twin”.

The objective of Adrien Hirvoas’s PhD work is to develop of data assimilation methodology to build the "digital twin" of an onshore wind turbine. Based on measurements, the data assimilation should allow to reduce the uncertainties of the physical parameters of the numerical model developed during the design phase to obtain a highly reliable model. Various ensemble data assimilation approches are currently under consideration to address the problem.

This work in done in collaboration with IFPEN.

6.3.7. Non-Parametric Estimation for Kinetic Diffusions

Participants: Clémentine Prieur, Jose Raphael Leon Ramos.

This research is the subject of a collaboration with Chile and Uruguay. More precisely, we started working with Venezuela. Due to the crisis in Venezuela, our main collaborator on that topic moved to Uruguay.

We are focusing our attention on models derived from the linear Fokker-Planck equation. From a probabilistic viewpoint, these models have received particular attention in recent years, since they are a basic example for hypercoercivity. In fact, even though completely degenerated, these models are hypoelliptic and still verify some properties of coercivity, in a broad sense of the word. Such models often appear in the fields of mechanics, finance and even biology. For such models we believe it appropriate to build statistical non-parametric estimation tools. Initial results have been obtained for the estimation of invariant density, in conditions guaranteeing its existence and unicity [48] and when only partial observational data are available. A paper on the non parametric estimation of the drift has been accepted recently [49] (see Samson et al., 2012, for results for parametric models). As far as the estimation of the diffusion term is concerned, a paper has been accepted [49], in collaboration with J.R. Leon (Montevideo, Uruguay) and P. Cattiaux (Toulouse). Recursive estimators have been also proposed by the same authors in [50], also recently accepted. In a recent collaboration with Adeline Samson from the statistics department in the Lab, we considered adaptive estimation, that is we proposed a data-driven procedure for the choice of the bandwidth parameters.

In [51], we focused on damping Hamiltonian systems under the so-called fluctuation-dissipation condition. Idea in that paper were re-used with applications to neuroscience in [74].

Note that Professor Jose R. Leon (Caracas, Venezuela, Montevideo, Uruguay) was funded by an international Inria Chair, allowing to collaborate further on parameter estimation.

We recently proposed a paper on the use of the Euler scheme for inference purposes, considering reflected diffusions.This paper could be extended to the hypoelliptic framework.

We also have a collaboration with Karine Bertin (Valparaiso, Chile), Nicolas Klutchnikoff (Université Rennes) and Jose R. León (Montevideo, Uruguay) funded by a MATHAMSUD project (2016-2017) and by the LIA/CNRS (2018). We are interested in new adaptive estimators for invariant densities on bounded domains [32], and would like to extend that results to hypo-elliptic diffusions.

6.3.8. Multivariate Risk Indicators

Participants: Clémentine Prieur, Patricia Tencaliec.

Studying risks in a spatio-temporal context is a very broad field of research and one that lies at the heart of current concerns at a number of levels (hydrological risk, nuclear risk, financial risk etc.). Stochastic tools for risk analysis must be able to provide a means of determining both the intensity and probability of occurrence of damaging events such as e.g. extreme floods, earthquakes or avalanches. It is important to be able to develop effective methodologies to prevent natural hazards, including e.g. the construction of barrages.

Different risk measures have been proposed in the one-dimensional framework . The most classical ones are the return level (equivalent to the Value at Risk in finance), or the mean excess function (equivalent to the Conditional Tail Expectation CTE). However, most of the time there are multiple risk factors, whose dependence structure has to be taken into account when designing suitable risk estimators. Relatively recent regulation (such as Basel II for banks or Solvency II for insurance) has been a strong driver for the development of realistic spatio-temporal dependence models, as well as for the development of multivariate risk measurements that effectively account for these dependencies.

We refer to [56] for a review of recent extensions of the notion of return level to the multivariate framework. In the context of environmental risk, [81] proposed a generalization of the concept of return period in dimension greater than or equal to two. Michele et al. proposed in a recent study [57] to take into account the duration and not only the intensity of an event for designing what they call the dynamic return period. However, few studies address the issues of statistical inference in the multivariate context. In [58], [60], we proposed non parametric estimators of a multivariate extension of the CTE. As might be expected, the properties of these estimators deteriorate when considering extreme risk levels. In collaboration with Elena Di Bernardino (CNAM, Paris), Clémentine Prieur is working on the extrapolation of the above results to extreme risk levels [35]. This paper has now been accepted for publication.

Elena Di Bernardino, Véronique Maume-Deschamps (Univ. Lyon 1) and Clémentine Prieur also derived an estimator for bivariate tail [59]. The study of tail behavior is of great importance to assess risk.

With Anne-Catherine Favre (LTHE, Grenoble), Clémentine Prieur supervised the PhD thesis of Patricia Tencaliec. We are working on risk assessment, concerning flood data for the Durance drainage basin (France). The PhD thesis started in October 2013 and was defended in February 2017. A first paper on data reconstruction has been accepted [83]. It was a necessary step as the initial series contained many missing data. A second paper is in revision, considering the modeling of precipitation amount with semi-parametric models, modeling both the bulk of the distribution and the tails, but avoiding the arbitrary choice of a threshold. We work in collaboration with Philippe Naveau (LSCE, Paris).

6.4. Assimilation of Images

Participants: Elise Arnaud, François-Xavier Le Dimet, Maëlle Nodet, Arthur Vidard, Long Li.

6.4.1. Direct assimilation of image sequences

At the present time the observation of Earth from space is done by more than thirty satellites. These platforms provide two kinds of observational information:

- Eulerian information as radiance measurements: the radiative properties of the earth and its fluid envelops. These data can be plugged into numerical models by solving some inverse problems.
- Lagrangian information: the movement of fronts and vortices give information on the dynamics of the fluid. Presently this information is scarcely used in meteorology by following small cumulus clouds and using them as Lagrangian tracers, but the selection of these clouds must be done by hand and the altitude of the selected clouds must be known. This is done by using the temperature of the top of the cloud.

Our current developments are targeted at the use of « Level Sets » methods to describe the evolution of the images. The advantage of this approach is that it permits, thanks to the level sets function, to consider the images as a state variable of the problem. We have derived an Optimality System including the level sets of the images. This approach is being applied to the tracking of oceanic oil spills in the framework of a Long Li's Phd in co-supervision with

A collaborative project started with C. Lauvernet (IRSTEA) in order to make use of our image assimilation strategies on the control of pesticide transfer.

6.4.2. Optimal transport for image assimilation

We investigate the use of optimal transport based distances for data assimilation, and in particular for assimilating dense data such as images. The PhD thesis of N. Feyeux studied the impact of using the Wasserstein distance in place of the classical Euclidean distance (pixel to pixel comparison). In a simplified one dimensional framework, we showed that the Wasserstein distance is indeed promising. Data assimilation experiments with the Shallow Water model have been performed and confirm the interest of the Wasserstein distance. Results have been presented at conferences and seminars and a paper has been published at NPG [6].

6.5. Land Use and Transport Models Calibration

Participants: Thomas Capelle, Laurent Gilquin, Clémentine Prieur, Arthur Vidard, Peter Sturm, Elise Arnaud.

Given the complexity of modern urban areas, designing sustainable policies calls for more than sheer expert knowledge. This is especially true of transport or land use policies, because of the strong interplay between the land use and the transportation systems. Land use and transport integrated (LUTI) modelling offers invaluable analysis tools for planners working on transportation and urban projects. Yet, very few local authorities in charge of planning make use of these strategic models. The explanation lies first in the difficulty to calibrate these models, second in the lack of confidence in their results, which itself stems from the absence of any well-defined validation procedure. Our expertise in such matters will probably be valuable for improving the reliability of these models. To that purpose we participated to the building up of the ANR project CITiES led by the STEEP EPI. This project started early 2013 and two PhD about sensitivity analysis and calibration were launched late 2013. Laurent Gilquin defended his PhD in October 2016 [62] and Thomas Capelle defended his in April 2017 and published his latest results in [3].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

A 2-year contract with Mercator-Ocean on the thematic "The AGRIF software in the NEMO European ocean model": see 5.1

Contract with IFPEN (Institut Français du pétrole et des énergies nouvelles), for the supervision of a PhD (Adrien Hirvoas). Research subject: Development of a data assimilation method for the calibration and continuous update of wind turbines digital twins

The Chair OQUAIDO – for "Optimisation et QUAntification d'Incertitudes pour les Données Onéreuses" in French – is the chair in applied mathematics held at Mines Saint-Étienne (France). It aims at gathering academical and technological partners to work on problems involving costly-to-evaluate numerical simulators for uncertainty quantification, optimization and inverse problems. This Chair, created in January 2016, is the continuation of the projects DICE and ReDICE which respectively covered the periods 2006-2009 and 2011-2015. Reda El Amri's PhD thesis is funded by OQUAIDO.

A 3-year contract (from June 2016 to June 2019) named ALBATROSS with Mercator-Ocean on the topic « Interaction océan, vagues, atmosphère à haute résolution » (PI: F. Lemarié).

8. Partnerships and Cooperations

8.1. Regional Initiatives

C. Prieur is co-leader of work-package 3 of the cross-disciplinary-project Trajectories from Idex Grenoble.

8.2. National Initiatives

8.2.1. ANR

A 4-year contract : ANR COCOA (COmprehensive Coupling approach for the Ocean and the Atmosphere). PI: E. Blayo. (Jan. 2017 - Dec. 2020). Other partners: Laboratoire des Sciences du Climat et de l'Environnement (UMR8212, Gif-sur-Yvette), Laboratoire de Météorologie Dynamique (UMR8539, Paris), Laboratoire d'Océanographie Physique et Spatiale (UMR6523, Brest), Centre National de Recherche Météorologique (UMR3589, Toulouse), Cerfacs (Toulouse). This project aims at revisiting the overall representation of air-sea interactions in coupled ocean-atmosphere models, and particularly in climate models, by coherently considering physical, mathematical, numerical and algorithmic aspects.

A 4-year contract : ANR HEAT (Highly Efficient ATmospheric modelling) http://www.agence-nationale-recherche.fr/?Project=ANR-14-CE23-0010.

A 4-year contract : ANR ADOM (Asynchronous Domain decomposition methods)

A 5-year contract with the French Navy (SHOM) on the improvment of the CROCO ocean model http://www.croco-ocean.org.

C. Prieur and E. Arnaud are involved as experts in project High-Tune http://www.agence-nationale-recherche.fr/Projet-ANR-16-CE01-0010 funded by ANR.

8.2.2. Other Initiatives

A. Vidard leads a group of projects gathering multiple partners in France and UK on the topic "Variational Data Assimilation for the NEMO/OPA9 Ocean Model", see 5.3.

C. Prieur chaired GdR MASCOT NUM 2010-2017, in which are also involved M. Nodet, E. Blayo, C. Helbert, E. Arnaud, L. Viry, S. Nanty, L. Gilquin. She is still strong involved in thie group (co-chair) http://www.gdr-mascotnum.fr/doku.php.

LEFE/GMMC CASIS, Coupled Assimilation Strategies for the Initialisation of an ocean- atmospheric boundary layer System, A. Vidard en collaboration avec Mercator océan

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

H2020 project IMMERSE (Improving Models for Marine EnviRonment SErvices) is funded from 2018-12-01 to 2022-11-30 (Inria contact: Florian Lemarié, coordinator: J. Le Sommer, CNRS). The overarching goal of the project is to ensure that the Copernicus Marine Environment Monitoring Service (CMEMS) will have continuing access to world-class marine modelling tools for its next generation systems while leveraging advances in space and information technologies, therefore allowing it to address the ever-increasing and evolving demands for marine monitoring and prediction in the 2020s and beyond.

See also https://cordis.europa.eu/project/rcn/218810/factsheet/fr

8.3.2. Collaborations with Major European Organizations

Partner: European Center for Medium Range Weather Forecast. Reading (UK)

World leading Numerical Weather Center, that include an ocean analysis section in order to provide ocean initial condition for the coupled ocean atmosphere forecast. They play a significant role in the NEMOVAR project in which we are also partner.

Partner: Met Office (U.K) National British Numerical Weather and Oceanographic service. Exceter (UK).

We do have a strong collaboration with their ocean initialization team through both our NEMO, NEMO-ASSIM and NEMOVAR activities. They also are our partner in the NEMOVAR consortium.

8.4. International Initiatives

8.4.1. Inria Associate Teams Not Involved in an Inria International Labs

8.4.1.1. UNQUESTIONABLE

Title: UNcertainty QUantification is ESenTIal for OceaNic & Atmospheric flows proBLEms.

International Partner:

Massachusetts Institute of Technology (United States) - Aerospace Computational Design Laboratory - Youssef Marzouk

Start year: 2018

See also: https://team.inria.fr/unquestionable/

The ability to understand and predict the behavior of geophysical flows is of greatest importance, due to its strong societal impact. Numerical models are essential to describe the evolution of the system (ocean + atmosphere), and involve a large number of parameters, whose knowledge is sometimes really poor. The reliability of the numerical predictions thus requires a step of parameter identification. The Inria-AIRSEA team has a strong expertise in variational approaches for inverse problems. An alternative is the use of particle filters, whose main advantage is their ability to tackle non-gaussian frameworks. However, particle filters suffer from the curse of dimensionality. The main objective of the collaboration we propose between the Inria-AIRSEA team and the MIT UQ group is the understanding of potential low-dimensional structure underlying geophysical applications, then the exploitation of such structures to extend particle filter to high-dimensional applications.

F. Lemarié is involved in the Inria associate team NEMOLOCO with Santiago University (Chile)

8.4.2. Inria International Partners

8.4.2.1. Informal International Partners

C. Prieur collaborates with Jose R. Leon (Universdad de la república de Uruguay, Montevideo).

C. Prieur collaborates with K. Bertin (CIMFAV, Valparaíso).

F. Lemarié and L. Debreu collaborate with Hans Burchard and Knut Klingbeil from the Leibniz-Institut für Ostseeforschung in Warnemünde (Germany).

8.5. International Research Visitors

8.5.1. Visits of International Scientists

Tiangang Cui, associate professor at Monash University (Melbourne, Australia), has visited the AIRSEA team during two weeks in December 2018. The purpose of this visit is to continue the collaboration with O. Zahm on the question of the dimension reduction for Bayesian inverse problems. Tiangang Cui also gave a presentation in the "Bayes in Grenoble" seminar at Inria-Montbonnot (3rd Dec. 2018).

Nicholas Kevlahan, from McMaster University (Canada) is a visiting scientist of the AIRSEA team for 10 months starting September 2018.

8.5.2. Visits to International Teams

Olivier Zahm was invited by Prof. Fabio Nobile to spend one week at EPFL to discuss the possible future collaboration

Clémentine Prieur visited Durham (US) in the framework of the SAMSI program on Quasi-Monte Carlo and High-Dimensional Sampling Methods for Applied Mathematics (QMC). She was invited for a tutorial at the Opening Workshop : August 28 – September 1, 2018. I take part to several working groups of the program.

Clémentine Prieur visited the Isaac Newton Institute for Mathematical Sciences in Cambridge in June 2018. She was invited in the framework of a semester on Uncertainty quantification for complex systems : theory and methodologies.

F.-X. Le Dimet visited the Florida State University during 2 weeks in May 2018. He made one presentation

F.-X. Le Dimet visited the University of Wisconsin during one week in June 2018. He made 2 presentations.

F.-X. Le Dimet visited the Harbin Institute of Technology during two weeks. He gave a 16 hours cours on the Data Assimilation.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

E. Blayo, E. Cosme and A. Vidard organized a one-week school “Introduction to data assimilation” for doctoral students (Jan. 8-12, 2018).

9.1.1.2. Member of the Organizing Committees

E. Blayo and A. Vidard were members of the organizing committee of the 7th National Conference on Data Assimilation (CNA 2018), held in Rennes Sept. 26-28, 2018.

E. Blayo organized with Pr Hansong Tang (City University of New York) a mini-symposium on domain decomposition and model coupling issues for oceanic and atmospheric flows, within the 25th International Domain Decomposition Conference, DD XXV (St. John’s, Canada, July 23-27, 2018).

L. Debreu was the co-organizer of the DRAKKAR workshop on global ocean simulations based on the NEMO platform. http://pp.ige-grenoble.fr/pageperso/barnierb/WEBDRAKKAR2018/

F. Lemarié and L. Debreu (with H. Burchard, K. Klingbeil and J. Sainte-Marie) have organized the international COMMODORE workshop on numerical methods for oceanic models (Paris, Sept. 17-19, 2018). https://commodore2018.sciencesconf.org/, see [12] for a summary of the scientific discussions.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

Clémentine Prieur was a member of the confrence program committee of the annual conference of the GdR MASCOT NUM 2018, Nantes (France), march 21-23, 2018.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

F. Lemarié is associate editor of the Journal of Advances in Modeling Earth Systems (JAMES)

C. Prieur is associate editor of the Annales Mathématiques Blaise Pascal, as far as of the journal Computational & Applied Mathematics, which was conceived as the main scientific publication of SBMAC (Brazilian Society of Computational and Applied Mathematics).

9.1.3.2. Reviewer - Reviewing Activities

F. Lemarié received a certificate for outstanding contribution in reviewing in recognition of the contributions made to the Ocean Modelling journal

F. Lemarié: reviewer for Journal of Computational and Applied Mathematics

E. Blayo: reviewer for Journal of Scientific Computing, Communications in Computational Physics

O. Zahm reviewed papers for: ESAIM: Mathematical Modelling and Numerical Analysis (M2AN), Journal Of Computational Physics, Journal on Uncertainty Quantification (JUQ), International Journal for Numerical Methods in Engineering, Aerospace Science and Technology, Computational and Applied Mathematics, Engineering Computations.

9.1.4. Invited Talks

E. Arnaud, M. Nodet. (Se) tromper avec les chiffres. Colloque "Sciences et esprit critique, interroger les certitudes", Maison pour la sciences, Académie de Grenoble, 8 nov. 2018, Grenoble

E. Arnaud. Estimation de paramètre sous incertitude. atelier incertitude du projet Idex trajectories, 17 mai 2018, Grenoble

F. Lemarié:

- Comod Workshop on "Coastal Ocean Modelling" in Hamburg (Germany) [16]
- Colloque de Bilan et de Prospective du programme LEFE in Clermond-Ferrand (France) [17]

C. Prieur:

- L'école Mathématiques pour l'énergie Nucléaire organied in partnership with GDR MaNu, July, 2-6, 2018.
- 13th International Conference in Monte Carlo & Quasi-Monte Carlo in Scientific Computing, July 1-6, 2018, Rennes, France.
- International workshop on Design of Experiments : New Challenges, April 30-May 4th, 2018, Luminy, France.
- Journée du groupe SIGMA of the SMAI, November, 30 2018, Jussieu (Paris).
- SIAM UQ 2018, 16-19 avril 2018, Garden Grove, Californie, US, invitation to participate in the "Advances in Global Sensitivity Analysis" session.

- Séminaire de Statistique d'Avignon, December, 10, 2018.
- Talk at the Issac Newton Institute in Cambridge (UK, June 2018).

9.1.5. Leadership within the Scientific Community

E. Blayo was the chair of the CNRS-INSU research program LEFE-MANU on mathematical and numerical methods for ocean and atmosphere http://www.insu.cnrs.fr/lefe until March 2018.

L. Debreu is the chair of the CNRS-INSU research program LEFE-MANU on mathematical and numerical methods for ocean and atmosphere http://www.insu.cnrs.fr/lefe since April 2018.

L. Debreu is the coordinator of the national group COMODO (Numerical Models in Oceanography).

L. Debreu is a member of the steering committee of the CROCO ocean model https://www.croco-ocean.org

C. Prieur chairs GdR MASCOT NUM. http://www.gdr-mascotnum.fr/doku.php.

9.1.6. Scientific Expertise

E. Arnaud, scientific evaluation for the AAP ANR 2018

E. Arnaud, C. Prieur, scientific expertise for the ANR project HIGH-TUNE

E. Arnaud, evaluation for the AAP Idex formation UGA 2018

F. Lemarié is a member of the CROCO (https://www.croco-ocean.org/) scientific committee in charge of the « numerical methods » topic.

F. Lemarié is a member of the NEMO (https://www.nemo-ocean.eu/) Developers Committee as external expert.

9.1.7. Research Administration

E. Blayo is a deputy director of the Jean Kuntzmann Lab.

E. Arnaud is in charge of the MAD (Modèles et algorithmes déterministes) department of Laboratoire Jean Kuntzmann

L. Debreu is a member of the scientific evaluation committee of the French Research Institute for Development (IRD).

L. Debreu is the chair of the French LEFE/MANU program on Applied mathematics and numerical methods for the ocean and the atmosphere. http://www.insu.cnrs.fr/lefe/presentation-manu

C. Prieur is a member of the Scientific Council of the Mathematical Society of France (SMF).

C. Prieur is a member of the Research Council of UGA.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

License: E. Arnaud, Mathematics for engineer, 50h, L1, Univ. Grenoble Alpes, France.

License: E. Arnaud, statisctics for biologists, 40h, L2, Univ. Grenoble Alpes, France.

Licence: M Nodet, Mathematiques pour l'ingenieur, 50h, L1, Univ. Grenoble Alpes

Licence: E. Blayo, Analyse approfondie, 80h, L1, Univ. Grenoble Alpes.

Licence : C.Kazantsev, Mathématiques outils pour les sciences et l'ingénierie, 76h, L1, Univ. Grenoble Alpes, France

Licence : C.Kazantsev, Mathématiques pour les sciences de l'ingénieur, 60h, L2, Univ. Grenoble Alpes, France

Master: E. Arnaud, Advising students on apprenticeship, 28h, M2, Univ. Grenoble Alpes, France.

Master: E. Arnaud, Inverse problem and data assimilation, 28h, M2, Univ. Grenoble Alpes, France.

Master: C. Prieur and O. Zahm, Model Exploration for Approximation of Complex, High-Dimensional Problems, 18h, M2, Univ. Grenoble Alpes-ENSIMAG, France.

Master: L. Debreu, Numerical methods for ocean models, 14h, M2, ENSTA Bretagne engineer schoold, Brest.

Master: E. Blayo, PDEs and numerical methods, 43h, M1, Univ. Grenoble Alpes and Ensimag engineer school.

Master: M Nodet Equations aux derivees partielles travaux pratiques, 18h, M1, Univ. Grenoble Alpes

Master: M Nodet Methodes inverses, 18h, M2, Univ. Grenoble, Alpes

Master: M Nodet taught was responsible for the 1st year of the applied mathematics master of Univ. Grenoble, Alpes

Doctorat: E. Blayo and A. Vidard, Introduction to data assimilation, 20h, Univ. Grenoble Alpes

Doctorat: M. Nodet, Data assimilation, 3h, Univ. Grenoble, Alpes

Doctorat: L. Debreu co-organized a one week doctoral training session on numerical modelling of atmospheric and oceanic flows (with F. Hourdin (LMD, Paris), G. Roullet (UBO, Brest) and T. Dubos (Ecole Polytechnique, Paris)).)

Doctorat: F.-X. Le Dimet, Data Assimilation, 16h, Harbin Institute of Technology.

E-learning : M.Nodet, Videos level L1, youtube channel https://www.youtube.com/channel/UCfYhfa8eKzdiM_MYAOoeMiQ

E-learning : E. Arnaud, Mathematics for engineer, L1, Pedagogical resources on http://math.u-ga.fr and videos for http://tinyurl.com/youtube-mat207

E-learning :E. Arnaud, Inverse problem and data assimilation, L2, Pedagogical resources on http://math.u-ga.fr.

9.2.2. Supervision

Intern: Natalie Noun, Characterization of coupling errors in ocean-atmosphere coupled models, M2R, mathématiques appliquées, Université Lyon 1, 6 months, A. Vidard and F. Lemarié.

PhD in progress: Victor Trappler, Parameter control in presence of uncertainties, October 2017, E. Arnaud, L. Debreu and A. Vidard.

PhD in progress: Adrien Hirvoas, Development of a data assimilation method for the calibration and continuous update of wind turbines digital twins, May 2018, E. Arnaud, C. Prieur, F. Caleyron

PhD in progress : Sophie Théry, Numerical study of coupling algorithms and boundary layer parameterizations in climate models. October 2017, E. Blayo and F. Lemarié.

PhD in progress : Emilie Duval, Coupling hydrostatic and nonhydrostatic ocean circulation models. October 2018, L. Debreu and E. Blayo.

PhD in progress: Long Li, Assimilation d'image pour le suivi de polluants, September 2017, A. Vidard, J.-W. Ma (Harbin University, China).

PhD in progress: Reda El Amri, Analyse d'incertitudes et de robustesse pour les modèles à entrées et sorties fonctionnelles, April 2016, Clémentine Prieur, Céline Helbert (Centrale Lyon), funded by IFPEN, in the OQUAIDO chair program.

PhD in progress: Maria Belén Heredia, A generic Bayesian approach for the calibration of advanced snow avalanche models with application to real-time risk assessment conditional to snow conditions, October 2017, Nicolas Eckert (IRSTEA Grenoble), Clémentine Prieur, funded by the OSUG@2020 Labex.

PhD in progress: Philomène Le Gall, Nonparametric non stationarity tests for extremes, October 2018, Clémentine Prieur, Anne-Catherine Favre (IGE, Grenoble), Philippe Naveau (LSCE, Paris), funded by the cross-disciplinary-project Trajectories from Idex Grenoble.

PhD in progress: Arthur Macherey, Uncertainty quantification methods for models described by stochastic differential equations or partial differential equations with a probabilistic interpretation, April 2018, Clémentine Prieur, Anthony Nouy (Ecole Centrale de Nantes), Marie Billaud Friess (Ecole Centrale de Nantes), funded by Inria and Ecole Centrale de Nantes.

PhD : Rémi Pellerej, Étude et développement d'algorithmes d'assimilation de données variationnelle adaptés aux modèles couplés océan-atmosphère, Université Grenoble-Alpes, Mars 2018, A. Vidard, F. Lemarié.

PhD : Charles Pelletier, Etude mathématique du problème de couplage océan-atmosphère incluant les échelles turbulentes, Université Grenoble Alpes, February 15, 2018, E. Blayo and F. Lemarié

9.2.3. Juries

E. Blayo:

- February 27, 2018: HDR thesis of Igor Gejadze, Univ. Grenoble Alpes (president)
- March 26, 2018: PhD thesis of Rémi Pellerej, Univ. Grenoble Alpes (president)
- June 28, 2018: PhD thesis of Pedro Colombo, Univ. Grenoble Alpes (president)
- July 3, 2018: PhD thesis of Matthieu Brachet, Université de Lorraine (referee)
- September 21, 2018: PhD thesis of Joseph Bellier, Univ. Grenoble Alpes (president)
- September 25, 2018: PhD thesis of Alexandre Vieira, Univ. Grenoble Alpes (president)

L. Debreu:

- November 12, 2018: PhD thesis of Thibaud Vandamme, Univ. of Toulouse (referee).

E. Arnaud, in charge of ATER recruitment in computer sciences, University Grenoble Alpes

E. Arnaud, member of a recrutment comitee for a "Maitre de conférences", University Grenoble Alpes

M. Nodet was a reviewer for the PhD of R. Ventura

M. Nodet was an expert at the TIPE jury (oral examination for the admission to Grandes Ecoles)

9.3. Popularization

9.3.1. Internal or external Inria responsibilities

E.Kazantsev is a memeber of the Local Commission for Permanent Formation of Inria Grenoble - Rhône-Alpes.

9.3.2. Articles and contents

Blog : E. Blayo, Les big data peuvent-ils faire la pluie et le beau temps ?, October 25, 2018, http://binaire.blog.lemonde.fr/2018/10/25/les-big-data-peuvent-ils-faire-la-pluie-et-le-beau-temps/Binaire and https://theconversation.com/quand-les-big-data-font-la-pluie-et-le-beau-temps-105608The conversation

Interviews : E. Blayo, Carnets de campagne - France Inter, September 27, 2018

E. Arnaud, "Sans le numérique, pas de prévision météo !" Dossier C'est quoi le numérique. "Le dauphiné des enfants", numéro 14, novembre-décembre 2018, http://www.ledauphinedesenfants.ledauphine.com/numero-archive/dde2018_14#

In collaboration with teachers, M. Nodet wrote an article about how to set up a "glaciers" for secondary school teachers in Reperes IREM journal [31]

9.3.3. Education

E. Arnaud, animation of a "Laboratoire des mathématiques" in Lycée Pablo Neruda (Saint Martin d'Hères), on the subject of math and critical thinking.

C.Kazantsev is a member of an IREM group for creation of scientific activities for professional development of secondary schools teachers.

C.Kazantsev is the creator and animator of the mathematical club for motivated pupils aged from 8 up to 18 years. 10 meetings for about 20 pupils of 2 hours each have been organized on sunday morning during the year.

C.Kazantsev participated in the meeting "Espace Mathématique Francophone" EMF2018, Gennevilliers, October, 22-26, 2018. Presentation of animations and participation in the working group "Étude des processus de vulgarisation".

C.Kazantsev participated in the elaboration and planning of the permanent exposition of the "Salle de thé" group at the meeting IHP2020, 30 october, Institut Henri Poincaré, Paris.

9.3.4. Interventions

- National events:

  E. Arnaud, M. Nodet. (Se) tromper avec les chiffres. Colloque "Sciences et esprit critique, interroger les certitudes", Maison pour la sciences, Académie de Grenoble, 8 nov. 2018, Grenoble - intervention in a training days for teachers from Education Nationale

  M. Nodet facilitated a yearly math club with three secondary schools around Grenoble, aiming pupils to research open projects ("Math en Jeans" initiative)

  M. Nodet facilitated a workshop about "Math modelling, example with oceanography" for high school and secondary school teachers

  M. Nodet welcomed schoolchildren for short-time internship and offered a small workshop about oceanography

  E. Blayo gave several outreach talks, in particular for middle school and high school students, and for more general audiences.

  Ch. Kazantsev and E. Blayo are strongly involved in the creation of "La Grange des maths", a science popularization center that will be located in Varces (south of Grenoble), which will offer a huge variety of mathematical hands-on exhibits. See http://www.la-grange-des-maths.fr/

  Ch. Kazantsev participated in the "Fête de la Science", Parvis de Sciences, October, 13.

  Ch. Kazantsev participated in the "Fille et Maths" day, Decembre, 5, with the presentation of the "La Grange des Maths" center and its activities.

  Ch. Kazantsev presented the exposition "La recherche se prend au jeu" at the "Culture and mathematical games Salon" by invitation of the Henri Poincaré Institute, Place St Sulpice, Paris, 24-27 May.

- Public exhibitions

  C.Kazantsev participated in the "Oriel des Maths" and in the "Forum des associations" with the presentation of the "La Grange des Maths" center and its activities. Varces, March, 11.

  C.Kazantsev participated at the "Raout de Domène" with the presentation of mathematical animations. Domene, September, 3.

  C.Kazantsev participated at the "Maths en ville" with the presentation of the "La Grange des Maths" center and its activities. Grenoble, November, 28.

- In educational institutions

  L. Debreu gave an introductary lecture on the finite element method at the Ecole des Pupilles de l'AIR (superior mathematics)

  C.Kazantsev presented mathematical animations to pupils of the Poussous scholl in Varces for about 80 children during 3 hours. Varces, March, 15.

9.3.5. *Creation of media or tools for science outreach*

C.Kazantsev participated in the edition of the Teachers notebooks which explain and advise how to use the "La Grange Suitcases" (sets of mathematical games, problems and animations) destined for primary and secondary schools teachers as well as for the general public.

C.Kazantsev participated in the creation of mathematical activities that can be autonomously used by schoolchildren of primary and secondary schools and by the general public.

E. Arnaud, in charge of the UGA Idex project math@uga : implementation of a collaborative moodle platform http://math.u-ga.fr to share pedagogical resources within teachers and towards students.

E. Arnaud, participation to UGA Idex projects Caseine and data@ugat

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] R. PELLEREJ.*Study and development of some variational data assimilation methods suitable for ocean-atmophere coupled models*, Université Grenoble Alpes, March 2018, https://tel.archives-ouvertes.fr/tel-01806281

[2] C. PELLETIER.*Mathematical study of the air-sea coupling problem including turbulent scale effects*, Université Grenoble Alpes, February 2018, https://tel.archives-ouvertes.fr/tel-01717274

Articles in International Peer-Reviewed Journal

[3] T. CAPELLE, P. STURM, A. VIDARD, B. MORTON.*Calibration of the Tranus Land Use Module: Optimisation-Based Algorithms, their Validation, and Parameter Selection by Statistical Model Selection*, in "Computers, Environment and Urban Systems", 2018 [*DOI :* 10.1016/J.COMPENVURBSYS.2017.04.009], https://hal.inria.fr/hal-01519654

[4] C. ELDRED, T. DUBOS, E. KRITSIKIS.*A Quasi-Hamiltonian Discretization of the Thermal Shallow Water Equations*, in "Journal of Computational Physics", October 2018, p. 1-53 [*DOI :* 10.1016/J.JCP.2018.10.038], https://hal.inria.fr/hal-01847698

[5] C. ELDRED, D. LE ROUX.*Dispersion analysis of compatible Galerkin schemes for the 1D shallow water model*, in "Journal of Computational Physics", October 2018, vol. 371, p. 779-800 [*DOI :* 10.1016/J.JCP.2018.06.007], https://hal.archives-ouvertes.fr/hal-01669048

[6] N. FEYEUX, A. VIDARD, M. NODET.*Optimal transport for variational data assimilation*, in "Nonlinear Processes in Geophysics", January 2018, vol. 25, n^{o} 1, p. 55-66 [*DOI :* 10.5194/NPG-25-55-2018], https://hal.archives-ouvertes.fr/hal-01342193

[7] L. GILQUIN, E. ARNAUD, C. PRIEUR, A. JANON.*Making best use of permutations to compute sensitivity indices with replicated orthogonal arrays*, in "Reliability Engineering and System Safety", October 2018, p. 1-12 [*DOI :* 10.1016/J.RESS.2018.09.010], https://hal.inria.fr/hal-01558915

[8] M. GROSS, H. WAN, P. J. RASCH, P. M. CALDWELL, D. L. WILLIAMSON, D. KLOCKE, C. JABLONOWSKI, D. R. THATCHER, N. WOOD, M. CULLEN, B. BEARE, M. WILLETT, F. LEMARIÉ, E. BLAYO, S. MALARDEL, P. TERMONIA, A. GASSMANN, P. H. LAURITZEN, H. JOHANSEN, C. M. ZARZYCKI, K. SAKAGUCHI, R. LEUNG.*Recent progress and review of Physics Dynamics Coupling in geophysical models*, in "Monthly Weather Review", August 2018, https://arxiv.org/abs/1605.06480 [*DOI :* 10.1175/MWR-D-17-0345.1], https://hal.inria.fr/hal-01323768

[9] A. JANON, M. NODET, C. PRIEUR, C. PRIEUR.*Goal-oriented error estimation for parameter-dependent nonlinear problems*, in "ESAIM: Mathematical Modelling and Numerical Analysis", July 2018, vol. 52, n^{o} 2, p. 705-728 [*DOI :* 10.1051/M2AN/2018003], https://hal.archives-ouvertes.fr/hal-01290887

[10] L. A. JIMÉNEZ RUGAMA, L. GILQUIN.*Reliable error estimation for Sobol' indices*, in "Statistics and Computing", July 2018, vol. 28, n^{o} 4, p. 725–738 [*DOI :* 10.1007/S11222-017-9759-1], https://hal.inria.fr/hal-01358067

[11] K. KLINGBEIL, F. LEMARIÉ, L. DEBREU, H. BURCHARD.*The numerics of hydrostatic structured-grid coastal ocean models: state of the art and future perspectives*, in "Ocean Modelling", May 2018, vol. 125, p. 80-105 [*DOI :* 10.1016/J.OCEMOD.2018.01.007], https://hal.inria.fr/hal-01443357

[12] F. LEMARIÉ, H. BURCHARD, L. DEBREU, K. KLINGBEIL, J. SAINTE-MARIE.*Advancing dynamical cores of oceanic models across all scales*, in "Bulletin of the American Meteorological Society", November 2018 [*DOI :* 10.1175/BAMS-D-18-0303.1], https://hal.inria.fr/hal-01939057

[13] L. LI, A. VIDARD, F.-X. LE DIMET, J. MA.*Topological data assimilation using Wasserstein distance*, in "Inverse Problems", January 2019, vol. 35, n^{o} 1, 015006 [*DOI :* 10.1088/1361-6420/AAE993], https://hal.inria.fr/hal-01960206

[14] V. OERDER, F. COLAS, V. ECHEVIN, S. MASSON, F. LEMARIÉ.*Impacts of the Mesoscale Ocean-Atmosphere Coupling on the Peru-Chile Ocean Dynamics: The Current-Induced Wind Stress Modulation*, in "Journal of Geophysical Research. Oceans", February 2018, vol. 123, n^{o} 2, p. 812-833 [*DOI :* 10.1002/2017JC013294], https://hal.inria.fr/hal-01661645

[15] P. TENCALIEC, A.-C. FAVRE, P. NAVEAU, C. PRIEUR.*Flexible semiparametric Generalized Pareto modeling of the entire range of rainfall amount*, in "Environmetrics", 2018, p. 1-22, https://hal.inria.fr/hal-01709061

Invited Conferences

[16] F. AUCLAIR, R. BENSHILA, L. DEBREU, N. DUCOUSSO, F. DUMAS, P. MARCHESIELLO, F. LEMARIÉ.*Some Recent Developments around the CROCO Initiative for Complex Regional to Coastal Modeling*, in "Comod Workshop on Coastal Ocean Modelling", Hambourg, Germany, February 2018, https://hal.inria.fr/hal-01947670

[17] F. LEMARIÉ, G. SAMSON, J.-L. REDELSPERGER, G. MADEC, H. GIORDANI, R. BOURDALLE-BADIE, Y. DRILLET.*PPR SIMBAD: en quête d'une nouvelle méthodologie de représentation des échanges air-mer dans les modèles opérationnels globaux d'océan à haute-résolution*, in "Colloque de Bilan et de Prospective du programme LEFE", Clermond-Ferrand, France, March 2018, https://hal.inria.fr/hal-01947683

[18] O. ZAHM, P. CONSTANTINE, C. PRIEUR, Y. MARZOUK.*Certified dimension reduction of the input parameter space of vector-valued functions*, in "INI Workshop UNQW03", Cambridge, United Kingdom, March 2018, https://hal.inria.fr/hal-01955776

[19] O. ZAHM, P. CONSTANTINE, C. PRIEUR, Y. MARZOUK.*Certified dimension reduction of the input parameter space of vector-valued functions*, in "FrontUQ 18 - Frontiers of Uncertainty Quantification", Pavie, Italy, September 2018, https://hal.inria.fr/hal-01955806

[20] O. ZAHM, Y. MARZOUK, C. PRIEUR, P. CONSTANTINE.*Certified dimension reduction of the input parameter space of Bayesian inverse problems*, in "IMS Vilnius - 12th International Vilnius Conference on Probability Theory and Mathematical Statistics", Vilnius, Lithuania, July 2018, https://hal.inria.fr/hal-01955800

[21] O. ZAHM.*Certified dimension reduction of the input parameter space of multivariate functions*, in "Journées EDP Auvergne-Rhône-Alpes", Grenoble, France, November 2018, https://hal.inria.fr/hal-01955812

[22] O. ZAHM.*Detecting and exploiting the low-effective dimension of multivariate problems using gradient information*, in "Séminaire MATHICSE, EPFL", Lausanne, Switzerland, November 2018, https://hal.inria.fr/hal-01955818

[23] O. ZAHM.*Dimension reduction of the input parameter space of vector-valued functions*, in "SIAM-UQ 2018 - SIAM Conference on Uncertainty Quantification", Los Angeles, United States, April 2018, https://hal.inria.fr/hal-01955795

Conferences without Proceedings

[24] F. AUCLAIR, L. DEBREU.*A non-hydrostatic non Boussinesq algorithm for free surface ocean modelling*, in "COMMODORE: Community for the numerical modeling of the global, regional and coastal ocean", Paris, France, September 2018, https://hal.inria.fr/hal-01961579

[25] E. BLAYO, F. LEMARIÉ, C. PELLETIER, S. THÉRY.*Toward improved ocean-atmosphere coupling algorithms*, in "25th international conference on Domain Decomposition Methods", St. John's, Canada, July 2018, https://hal.inria.fr/hal-01951472

[26] E. BLAYO, A. ROUSSEAU.*Coupling hydrostatic and nonhydrostatic Navier-Stokes flows using a Schwarz algorithm*, in "25th international conference on Domain Decomposition Methods", St. John's, Canada, July 2018, https://hal.inria.fr/hal-01951485

[27] F. LEMARIÉ.*On the discretization of vertical diffusion in the turbulent surface and planetary boundary layers*, in "3rd workshop on Physics Dynamics Coupling (PDC18)", Reading, United Kingdom, July 2018, https://hal.inria.fr/hal-01947691

[28] S. THÉRY, E. BLAYO, F. LEMARIÉ.*Algorithmes de Schwarz et conditions absorbantes pour le couplage océan-atmosphère*, in "Congrès National d'Analyse Numérique", Cap d'Agde, France, May 2018, https://hal.inria.fr/hal-01947885

[29] O. ZAHM.*Dimension reduction of the input parameter space of vector-valued functions*, in "MoRePaS 2018 - Model Reduction of Parametrized Systems IV", Nantes, France, April 2018, https://hal.inria.fr/hal-01955788

Scientific Popularization

[30] E. BLAYO.*Les big data peuvent-ils faire la pluie et le beau temps ?*, Le Monde, October 2018, https://hal.inria.fr/hal-01951505

[31] S. DEWYSPELAERE, M. NODET, J. CHARTON, P. GARAT, F. LETUE, C. PE`S, V. WALES.*Exemple d'EPI au collège : l'évolution des glaciers*, in "Repères IREM", July 2018, n^{o} 112, https://hal.inria.fr/hal-01619788

Other Publications

[32] K. BERTIN, N. KLUTCHNIKOFF, J. LEÓN, C. PRIEUR.*Adaptive density estimation on bounded domains under mixing conditions*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01934913

[33] L. DEBREU, E. KAZANTSEV.*Optimal control of grids and schemes for the inertial gravity waves equation*, January 2019, working paper or preprint, https://hal.inria.fr/hal-01968678

[34] J. DEMANGE, L. DEBREU, P. MARCHESIELLO, F. LEMARIÉ, E. BLAYO, C. ELDRED.*Stability analysis of split-explicit free surface ocean models: implication of the depth-independent barotropic mode approximation*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01947706

[35] E. DI BERNARDINO, C. PRIEUR.*Estimation of the Multivariate Conditional-Tail-Expectation for extreme risk levels: illustrations on environmental data-sets*, March 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01524536

[36] M. R. EL AMRI, C. HELBERT, O. LEPREUX, M. MUNOZ ZUNIGA, C. PRIEUR, D. SINOQUET.*Data-driven stochastic inversion under functional uncertainties*, February 2018, working paper or preprint, https://hal.inria.fr/hal-01704189

[37] C. ELDRED, W. BAUER.*Variational and Hamiltonian Formulations of Geophysical Fluids using Split Exterior Calculus*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01895935

[38] C. ELDRED, F. GAY-BALMAZ.*Single and Double Generator Bracket Formulations of Geophysical Fluids with Irreversible Processes*, November 2018, https://arxiv.org/abs/1811.11609 - working paper or preprint, https://hal.inria.fr/hal-01937941

[39] C. ELDRED, D. LE ROUX.*Dispersion analysis of compatible Galerkin schemes on quadrilaterals for shallow water models*, November 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01916382

[40] P. ETORÉ, C. PRIEUR, D. K. PHAM, L. LI.*Global sensitivity analysis for models described by stochastic differential equations*, November 2018, https://arxiv.org/abs/1811.08101 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01926919

[41] B. IOOSS, C. PRIEUR.*Shapley effects for sensitivity analysis with correlated inputs: comparisons with Sobol' indices, numerical estimation and applications*, May 2018, https://arxiv.org/abs/1707.01334 - working paper or preprint, https://hal.inria.fr/hal-01556303

[42] R. LAM, O. ZAHM, Y. MARZOUK, K. WILLCOX.*Multifidelity Dimension Reduction via Active Subspaces*, September 2018, https://arxiv.org/abs/1809.05567 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01875946

[43] C. PRIEUR, L. VIRY, E. BLAYO, J.-M. BRANKART.*A global sensitivity analysis approach for marine biogeochemical modeling*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01952797

[44] K. SMETANA, O. ZAHM, A. T. PATERA.*Randomized residual-based error estimators for parametrized equations*, July 2018, https://arxiv.org/abs/1807.10489 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01851462

[45] O. ZAHM, P. CONSTANTINE, C. PRIEUR, Y. MARZOUK.*Gradient-based dimension reduction of multivariate vector-valued functions*, December 2018, https://arxiv.org/abs/1801.07922 - working paper or preprint, https://hal.inria.fr/hal-01701425

[46] O. ZAHM, T. CUI, K. LAW, A. SPANTINI, Y. MARZOUK.*Certified dimension reduction in nonlinear Bayesian inverse problems*, July 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01834039

References in notes

[47] A. BELJAARS, E. DUTRA, G. BALSAMO, F. LEMARIÉ.*On the numerical stability of surface-atmosphere coupling in weather and climate models*, in "Geoscientific Model Development Discussions", 2017, vol. 10, n^{o} 2, p. 977-989 [*DOI :* 10.5194/GMD-10-977-2017], https://hal.inria.fr/hal-01406623

[48] P. CATTIAUX, J. R. LEON, C. PRIEUR.*Estimation for Stochastic Damping Hamiltonian Systems under Partial Observation. I. Invariant density*, in "Stochastic Processes and their Applications", March 2014, vol. 124, n^{o} 3, p. 1236-1260 [*DOI :* 10.1016/J.SPA.2013.10.008], https://hal.archives-ouvertes.fr/hal-00739136

[49] P. CATTIAUX, J. R. LEON, C. PRIEUR.*Estimation for Stochastic Damping Hamiltonian Systems under Partial Observation. II Drift term*, in "ALEA (Latin American Journal of Probability and Statistics)", 2014, vol. 11, n^{o} 1, p. 359-384, https://hal.archives-ouvertes.fr/hal-00877054

[50] P. CATTIAUX, J. R. LEON, C. PRIEUR.*Recursive Estimation for Stochastic Damping Hamiltonian Systems*, in "Journal of Nonparametric Statistics", 2015, vol. 27, n^{o} 3, p. 401-424, https://hal.archives-ouvertes.fr/hal-01071252

[51] P. CATTIAUX, J. R. LEÓN, A. PINEDA CENTENO, C. PRIEUR.*An overlook on statistical inference issues for stochastic damping Hamiltonian systems under the fluctuation-dissipation condition*, in "Statistics", 2017, vol. 51, n^{o} 1, p. 11-29 [*DOI :* 10.1080/02331888.2016.1259807], https://hal.archives-ouvertes.fr/hal-01405427

[52] M. CHAMPION, G. CHASTAING, S. GADAT, C. PRIEUR.*L2 Boosting on generalized Hoeffding decomposition for dependent variables. Application to Sensitivity Analysis*, 2013, 48 pages, 7 Figures

[53] G. CHASTAING.*Generalized Sobol sensitivity indices for dependent variables*, Université de Grenoble, September 2013, https://tel.archives-ouvertes.fr/tel-00930229

[54] G. CHASTAING, F. GAMBOA, C. PRIEUR.*Generalized Hoeffding-Sobol Decomposition for Dependent Variables - Application to Sensitivity Analysis*, in "Electronic Journal of Statistics", December 2012, vol. 6, p. 2420-2448 [*DOI :* 10.1214/12-EJS749], http://hal.archives-ouvertes.fr/hal-00649404

[55] G. CHASTAING, C. PRIEUR, F. GAMBOA.*Generalized Sobol sensitivity indices for dependent variables: numerical methods*, March 2013, http://hal.inria.fr/hal-00801628

[56] A. COUSIN, E. DI BERNARDINO.*On multivariate extensions of Value-at-Risk*, in "J. Multivariate Anal.", 2013, vol. 119, p. 32–46, http://dx.doi.org/10.1016/j.jmva.2013.03.016

[57] C. DE MICHELE, G. SALVADORI, R. VEZZOLI, S. PECORA.*Multivariate assessment of droughts: Frequency analysis and dynamic return period*, in "Water Resources Research", 2013, vol. 49, n^o 10, p. 6985–6994

[58] E. DI BERNARDINO, T. LALOË, V. MAUME-DESCHAMPS, C. PRIEUR.*Plug-in estimation of level sets in a non-compact setting with applications in multivariate risk theory*, in "ESAIM: Probability and Statistics", February 2013, vol. 17, p. 236-256 [*DOI :* 10.1051/PS/2011161], https://hal.archives-ouvertes.fr/hal-00580624

[59] E. DI BERNARDINO, V. MAUME-DESCHAMPS, C. PRIEUR.*Estimating Bivariate Tail: a copula based approach*, in "Journal of Multivariate Analysis", August 2013, vol. 119, p. 81-100 [*DOI :* 10.1016/J.JMVA.2013.03.020], https://hal.archives-ouvertes.fr/hal-00475386

[60] E. DI BERNARDINO, C. PRIEUR.*Estimation of Multivariate Conditional Tail Expectation using Kendall's Process*, in "Journal of Nonparametric Statistics", March 2014, vol. 26, n^o 2, p. 241-267 [*DOI :* 10.1080/10485252.2014.889137], https://hal.archives-ouvertes.fr/hal-00740340

[61] L. GILQUIN, E. ARNAUD, C. PRIEUR, H. MONOD.*Recursive estimation procedure of Sobol' indices based on replicated designs*, January 2016, working paper or preprint, https://hal.inria.fr/hal-01291769

[62] L. GILQUIN.*Monte Carlo and quasi-Monte Carlo sampling methods for the estimation of Sobol' indices. Application to a LUTI model*, Université Grenoble Alpes, October 2016, https://hal.inria.fr/tel-01403914

[63] L. GILQUIN, C. PRIEUR, E. ARNAUD.*Replication procedure for grouped Sobol' indices estimation in dependent uncertainty spaces*, in "Information and Inference", August 2015, vol. 4, n^o 4, p. 354-379 [*DOI :* 10.1093/IMAIAI/IAV010], https://hal.inria.fr/hal-01045034

[64] M. GROSS, H. WAN, P. J. RASCH, P. M. CALDWELL, D. L. WILLIAMSON, D. KLOCKE, C. JABLONOWSKI, D. R. THATCHER, N. WOOD, M. CULLEN, B. BEARE, M. WILLETT, F. LEMARIÉ, E. BLAYO, S. MALARDEL, P. TERMONIA, A. GASSMANN, P. H. LAURITZEN, H. JOHANSEN, C. M. ZARZYCKI, K. SAKAGUCHI, R. LEUNG.*Recent progress and review of Physics Dynamics Coupling in geophysical models*, May 2016, working paper or preprint, https://hal.inria.fr/hal-01323768

[65] W. HOEFFDING.*A class of statistics with asymptotically normal distribution*, in "Ann. Math. Statistics", 1948, vol. 19, p. 293–325

[66] A. JANON, T. KLEIN, A. LAGNOUX-RENAUDIE, M. NODET, C. PRIEUR.*Asymptotic normality and efficiency of two Sobol index estimators*, in "ESAIM: Probability and Statistics", October 2014, vol. 18, p. 342-364 [*DOI :* 10.1051/PS/2013040], https://hal.inria.fr/hal-00665048

[67] A. JANON, M. NODET, C. PRIEUR.*Goal-oriented error estimation for reduced basis method, with application to certified sensitivity analysis*, http://hal.archives-ouvertes.fr/hal-00721616

[68] A. JANON, M. NODET, C. PRIEUR.*Certified reduced-basis solutions of viscous Burgers equation parametrized by initial and boundary values*, in "ESAIM: Mathematical Modelling and Numerical Analysis", March 2013, vol. 47, n^o 2, p. 317-348 [*DOI :* 10.1051/M2AN/2012029], http://hal.inria.fr/inria-00524727

[69] A. JANON, M. NODET, C. PRIEUR.*Uncertainties assessment in global sensitivity indices estimation from metamodels*, in "International Journal for Uncertainty Quantification", 2014, vol. 4, n^{o} 1, p. 21-36 [*DOI :* 10.1615/INT.J.UNCERTAINTYQUANTIFICATION.2012004291], https://hal.inria.fr/inria-00567977

[70] A. JANON, M. NODET, C. PRIEUR, C. PRIEUR.*Global sensitivity analysis for the boundary control of an open channel*, in "Mathematics of Control, Signals, and Systems", March 2016, vol. 28, n^{o} 1, p. 6:1-27 [*DOI :* 10.1007/S00498-015-0151-4], https://hal.archives-ouvertes.fr/hal-01065886

[71] A. JANON, M. NODET, C. PRIEUR, C. PRIEUR.*Goal-oriented error estimation for fast approximations of nonlinear problems*, GIPSA-lab, 2016, Rapport interne de GIPSA-lab, https://hal.archives-ouvertes.fr/hal-01290887

[72] F. LEMARIÉ, E. BLAYO, L. DEBREU.*Analysis of ocean-atmosphere coupling algorithms : consistency and stability*, in "Procedia Computer Science", 2015, vol. 51, p. 2066–2075 [*DOI :* 10.1016/J.PROCS.2015.05.473], https://hal.inria.fr/hal-01174132

[73] F. LEMARIÉ.*Numerical modification of atmospheric models to include the feedback of oceanic currents on air-sea fluxes in ocean-atmosphere coupled models*, Inria Grenoble - Rhône-Alpes ; Laboratoire Jean Kuntzmann ; Universite de Grenoble I - Joseph Fourier ; Inria, August 2015, n^{o} RT-0464, https://hal.inria.fr/hal-01184711

[74] J. R. LEON, A. SAMSON.*Hypoelliptic stochastic FitzHugh-Nagumo neuronal model: mixing, up-crossing and estimation of the spike rate*, in "Annals of Applied Probability", 2017, https://hal.archives-ouvertes.fr/hal-01492590

[75] S. NANTY, C. HELBERT, A. MARREL, N. PÉROT, C. PRIEUR.*Uncertainy quantification for functional dependent random variables*, 2014, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01075840

[76] S. NANTY, C. HELBERT, A. MARREL, N. PÉROT, C. PRIEUR.*Sampling, metamodelling and sensitivity analysis of numerical simulators with functional stochastic inputs*, in "SIAM/ASA Journal on Uncertainty Quantification", May 2016, vol. 4, n^{o} 1, p. 636-659 [*DOI :* 10.1137/15M1033319], https://hal.archives-ouvertes.fr/hal-01187162

[77] A. B. OWEN.*Sobol' indices and Shapley value*, in "Journal on Uncertainty Quantification", 2014, vol. 2, p. 245–251

[78] A. B. OWEN, C. PRIEUR.*On Shapley value for measuring importance of dependent inputs*, in "SIAM/ASA Journal on Uncertainty Quantification", September 2017, vol. 51, n^{o} 1, p. 986–1002 [*DOI :* 10.1137/16M1097717], https://hal.archives-ouvertes.fr/hal-01379188

[79] C. PELLETIER, F. LEMARIÉ, E. BLAYO.*A theoretical study of a simplified air-sea coupling problem including turbulent parameterizations*, in "COUPLED PROBLEMS 2017 - VII International Conference on Computational Methods for Coupled Problems in Science and Engineering", Rhodes, Greece, M. PAPADRAKAKIS, E. OÑATE, B. SCHREFLER (editors), International Center for Numerical Methods in Engineering (CIMNE) , June 2017, p. 38-49, https://hal.archives-ouvertes.fr/hal-01659443

[80] C. PELLETIER, F. LEMARIÉ, E. BLAYO.*Sensitivity analysis and metamodels for the bulk parameterization of turbulent air-sea fluxes*, in "Quarterly Journal of the Royal Meteorological Society", December 2017 [*DOI :* 10.1002/QJ.3233], https://hal.inria.fr/hal-01663668

[81] G. SALVADORI, C. DE MICHELE, F. DURANTE.*On the return period and design in a multivariate framework*, in "Hydrology and Earth System Sciences", 2011, vol. 15, n^{o} 11, p. 3293–3305

[82] E. SONG, B. L. NELSON, J. STAUM.*Shapley Effects for Global Sensitivity Analysis: Theory and Computation*, Northwestern University, 2015

[83] P. TENCALIEC, A.-C. FAVRE, C. PRIEUR, T. MATHEVET.*Reconstruction of missing daily streamflow data using dynamic regression models*, in "Water Resources Research", December 2015, vol. 51, n^{o} 12, p. 9447–9463 [*DOI :* 10.1002/2015WR017399], https://hal.inria.fr/hal-01245238

Project-Team ARIC

Arithmetic and Computing

IN COLLABORATION WITH: Laboratoire de l'Informatique du Parallélisme (LIP)

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Lyon

Université Claude Bernard (Lyon 1)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Algorithmics, Computer Algebra and Cryptology

Table of contents

Project-Team ARIC

Creation of the Team: 2012 January 01, updated into Project-Team: 2013 January 01

Keywords:

Computer Science and Digital Science:

A1.1. - Architectures
A2.4. - Formal method for verification, reliability, certification
A4. - Security and privacy
A7. - Theory of computation
A8. - Mathematics of computing

Other Research Topics and Application Domains:

B9.5. - Sciences
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

Bruno Salvy [Team leader, Inria, Senior Researcher]
Nicolas Brisebarre [CNRS, Researcher, HDR]
Claude-Pierre Jeannerod [Inria, Researcher]
Vincent Lefèvre [Inria, Researcher]
Benoît Libert [CNRS, Senior Researcher, HDR]
Jean-Michel Muller [CNRS, Senior Researcher, HDR]
Alain Passelègue [Inria, Researcher, from Oct 2018]
Nathalie Revol [Inria, Researcher]
Gilles Villard [CNRS, Senior Researcher, HDR]

Faculty Members

Guillaume Hanrot [École Normale Supérieure de Lyon, Professor, HDR]
Fabien Laguillaumie [Univ. Claude Bernard, Professor, HDR]
Nicolas Louvet [Univ. Claude Bernard, Associate Professor]
Damien Stehlé [École Normale Supérieure de Lyon, Professor, HDR]

Technical Staff

Laurent Grémy [École Normale Supérieure de Lyon]
Joris Picot [École Normale Supérieure de Lyon]

PhD Students

Florent Bréhard [École Normale Supérieure de Lyon]
Adel Hamdi [Orange Labs]
Fabrice Mouhartem [École Normale Supérieure de Lyon]
Huyen Nguyen [École Normale Supérieure de Lyon, intern, then PhD student from Sep 2018]
Alice Pellet–Mary [École Normale Supérieure de Lyon]
Chen Qian [École Normale Supérieure de Rennes]
Miruna Rosca [Bitdefender Romania, École Normale Supérieure de Lyon]
Radu Titiu [Bitdefender Romania, École Normale Supérieure de Lyon]
Ida Tucker [École Normale Supérieure de Lyon]
Weiqiang Wen [École Normale Supérieure de Lyon, until Aug 2018]

Post-Doctoral Fellows

Chitchanok Chuengsatiansup [Inria]
Junqing Gong [École Normale Supérieure de Lyon]
Alonso Gonzalez [École Normale Supérieure de Lyon]
Gottfried Herold [École Normale Supérieure de Lyon]
Elena Kirshanova [École Normale Supérieure de Lyon]
Changmin Lee [Univ. de Lyon, from Oct 2018]
Anastasiia Volkova Lozanova [Inria]
Alexandre Wallet [École Normale Supérieure de Lyon]

Visiting Scientists

Shi Bai [École Normale Supérieure de Lyon, May–July 2018]
Lloyd Nicholas Trefethen [École Normale Supérieure de Lyon, until Jul 2018]
Warwick Tucker [Univ. de Lyon, until Jul 2018]

Administrative Assistants

Nelly Amsellem [École Normale Supérieure de Lyon, from Feb 2018]
Kadiatou Bangoura [École Normale Supérieure de Lyon]
Chiraz Benamor [École Normale Supérieure de Lyon]
Evelyne Blesle [Inria, until Jan 2018]

2. Overall Objectives

2.1. Overall Objectives

The overall objective of AriC (Arithmetic and Computing) is, through computer arithmetic and computational mathematics, to improve computing at large.

A major challenge in modeling and scientific computing is the simultaneous mastery of hardware capabilities, software design, and mathematical algorithms for the efficiency of the computation. Further, performance relates as much to efficiency as to reliability, requiring progress on automatic proofs, certificates and code generation. In this context, computer arithmetic and mathematical algorithms are the keystones of AriC. Our approach conciliates fundamental studies, practical performance and qualitative aspects, with a shared strategy going from high-level problem specifications and normalization actions, to computer arithmetic and the lowest-level details of implementations.
We focus on the following lines of action:

- Design and integration of new methods and tools for mathematical program specification, certification, security, and guarantees on numerical results. Some main ingredients here are: the interleaving of formal proofs, computer arithmetic and computer algebra; error analysis and computation of certified error bounds; the study of the relationship between performance and numerical quality; and on the cryptology aspects, focus on the practicality of existing protocols and design of more powerful lattice-based primitives.
- Generalization of a hybrid symbolic-numeric trend, and interplay between arithmetics for both improving and controlling numerical approaches (symbolic $\rightarrow$ numeric), and accelerating exact solutions (symbolic $\longleftarrow$ numeric). This trend, especially in the symbolic computation community, has acquired a strategic role for the future of scientific computing. The integration in AriC of computer arithmetic, reliable computing, and algebraic computing is expected to lead to a deeper understanding of the problem and novel solutions.
- Mathematical and algorithmic foundations of computing. We address algorithmic complexity and fundamental aspects of approximation, polynomial and matrix algebra, and lattice-based cryptology. Practical questions concern the design of high performance and reliable computing kernels, thanks to optimized computer arithmetic operators and an improved adequacy between arithmetic bricks and higher level ones.

According to the application domains that we target and our main fields of expertise, these lines of actions are declined in three themes with specific objectives. These themes also correspond to complementary angles for addressing the general computing challenge stated at the beginning of this introduction:

- **Efficient approximation methods** (§3.1). Here lies the question of interleaving formal proofs, computer arithmetic and computer algebra, for significantly extending the range of functions whose reliable evaluation can be optimized.
- **Lattices: algorithms and cryptology** (§3.2). Long term goals are to go beyond the current design paradigm in basis reduction, and to demonstrate the superiority of lattice-based cryptography over contemporary public-key cryptographic approaches.
- **Algebraic computing and high performance kernels** (§3.3). The problem is to keep the algorithm and software designs in line with the scales of computational capabilities and application needs, by simultaneously working on the structural and the computer arithmetic levels.

3. Research Program

3.1. Efficient approximation methods

3.1.1. Computer algebra generation of certified approximations

We plan to focus on the generation of certified and efficient approximations for solutions of linear differential equations. These functions cover many classical mathematical functions and many more can be built by combining them. One classical target area is the numerical evaluation of elementary or special functions. This is currently performed by code specifically handcrafted for each function. The computation of approximations and the error analysis are major steps of this process that we want to automate, in order to reduce the probability of errors, to allow one to implement "rare functions", to quickly adapt a function library to a new context: new processor, new requirements – either in terms of speed or accuracy.

In order to significantly extend the current range of functions under consideration, several methods originating from approximation theory have to be considered (divergent asymptotic expansions; Chebyshev or generalized Fourier expansions; Padé approximants; fixed point iterations for integral operators). We have done preliminary work on some of them. Our plan is to revisit them all from the points of view of effectivity, computational complexity (exploiting linear differential equations to obtain efficient algorithms), as well as in their ability to produce provable error bounds. This work is to constitute a major progress towards the automatic generation of code for moderate or arbitrary precision evaluation with good efficiency. Other useful, if not critical, applications are certified quadrature, the determination of certified trajectories of spatial objects and many more important questions in optimal control theory.

3.1.2. Digital Signal Processing

As computer arithmeticians, a wide and important target for us is the design of efficient and certified linear filters in digital signal processing (DSP). Actually, following the advent of MATLAB as the major tool for filter design, the DSP experts now systematically delegate to MATLAB all the part of the design related to numerical issues. And yet, various key MATLAB routines are neither optimized, nor certified. Therefore, there is a lot of room for enhancing numerous DSP numerical implementations and there exist several promising approaches to do so.

The main challenge that we want to address over the next period is the development and the implementation of optimal methods for rounding the coefficients involved in the design of the filter. If done in a naive way, this rounding may lead to a significant loss of performance. We will study in particular FIR and IIR filters.

3.1.3. Table Maker's Dilemma (TMD)

Implementing "ultimately accurate" functions (i.e., rounded to nearest) requires either the knowledge of hardest-to-round cases, or an as tight as possible lower bound on the distance between the image of a floating-point number by the function and the middle of two consecutive floating-point numbers. Obtaining such results is a challenge. Several computer manufacturers have contacted us to obtain new cases. One of our current solutions for obtaining hardest-to-round cases is based on Lefèvre's algorithm. We aim at rewriting the current implementations of this algorithm, and giving formal proofs of their correction.

We plan to use uniform polynomial approximation and diophantine techniques in order to tackle the case of the IEEE quad precision, and continue to use analytic number theory techniques (exponential sums estimates) for counting the hardest-to-round cases.

3.2. Lattices: algorithms and cryptology

Lattice-based cryptography (LBC) is an utterly promising, attractive (and competitive) research ground in cryptography, thanks to a combination of unmatched properties:

- **Improved performance.** LBC primitives have low asymptotic costs, but remain cumbersome in practice (e.g., for parameters achieving security against computations of up to 2100 bit operations). To address this limitation, a whole branch of LBC has evolved where security relies on the restriction of lattice problems to a family of more structured lattices called *ideal lattices*. Primitives based on such lattices can have quasi-optimal costs (i.e., quasi-constant amortized complexities), outperforming all contemporary primitives. This asymptotic performance sometimes translates into practice, as exemplified by NTRUEncrypt.
- **Improved security.** First, lattice problems seem to remain hard even for quantum computers. Moreover, the security of most of LBC holds under the assumption that standard lattice problems are hard in the worst case. Oppositely, contemporary cryptography assumes that specific problems are hard with high probability, for some precise input distributions. Many of these problems were artificially introduced for serving as a security foundation of new primitives.
- **Improved flexibility.** The master primitives (encryption, signature) can all be realized based on worst-case (ideal) lattice assumptions. More evolved primitives such as ID-based encryption (where the public key of a recipient can be publicly derived from its identity) and group signatures, that were the playing-ground of pairing-based cryptography (a subfield of elliptic curve cryptography), can also be realized in the LBC framework, although less efficiently and with restricted security properties. More intriguingly, lattices have enabled long-wished-for primitives. The most notable example is homomorphic encryption, enabling computations on encrypted data. It is the appropriate tool to securely outsource computations, and will help overcome the privacy concerns that are slowing down the rise of the cloud.

We work on three directions, detailed now.

3.2.1. Lattice algorithms

All known lattice reduction algorithms follow the same design principle: perform a sequence of small elementary steps transforming a current basis of the input lattice, where these steps are driven by the Gram-Schmidt orthogonalisation of the current basis.

In the short term, we will fully exploit this paradigm, and hopefully lower the cost of reduction algorithms with respect to the lattice dimension. We aim at asymptotically fast algorithms with complexity bounds closer to those of basic and normal form problems (matrix multiplication, Hermite normal form). In the same vein, we plan to investigate the parallelism potential of these algorithms.

Our long term goal is to go beyond the current design paradigm, to reach better trade-offs between run-time and shortness of the output bases. To reach this objective, we first plan to strengthen our understanding of the interplay between lattice reduction and numerical linear algebra (how far can we push the idea of working on approximations of a basis?), to assess the necessity of using the Gram-Schmidt orthogonalisation (e.g., to obtain a weakening of LLL-reduction that would work up to some stage, and save computations), and to determine whether working on generating sets can lead to more efficient algorithms than manipulating bases. We will also study algorithms for finding shortest non-zero vectors in lattices, and in particular look for quantum accelerations.

We will implement and distribute all algorithmic improvements, e.g., within the fplll library. We are interested in high performance lattice reduction computations (see application domains below), in particular in connection with/continuation of the HPAC ANR project (algebraic computing and high performance consortium).

3.2.2. Lattice-based cryptography

Our long term goal is to demonstrate the superiority of lattice-based cryptography over contemporary public-key cryptographic approaches. For this, we will 1- Strengthen its security foundations, 2- Drastically improve the performance of its primitives, and 3- Show that lattices allow to devise advanced and elaborate primitives.

The practical security foundations will be strengthened by the improved understanding of the limits of lattice reduction algorithms (see above). On the theoretical side, we plan to attack two major open problems: Are ideal lattices (lattices corresponding to ideals in rings of integers of number fields) computationally as hard to handle as arbitrary lattices? What is the quantum hardness of lattice problems?

Lattice-based primitives involve two types of operations: sampling from discrete Gaussian distributions (with lattice supports), and arithmetic in polynomial rings such as $(\mathbb{Z}/q\mathbb{Z})[x]/(x^n+1)$ with n a power of 2. When such polynomials are used (which is the case in all primitives that have the potential to be practical), then the underlying algorithmic problem that is assumed hard involves ideal lattices. This is why it is crucial to precisely understand the hardness of lattice problems for this family. We will work on improving both types of operations, both in software and in hardware, concentrating on values of q and n providing security. As these problems are very arithmetic in nature, this will naturally be a source of collaboration with the other themes of the AriC team.

Our main objective in terms of cryptographic functionality will be to determine the extent to which lattices can help securing cloud services. For example, is there a way for users to delegate computations on their outsourced dataset while minimizing what the server eventually learns about their data? Can servers compute on encrypted data in an efficiently verifiable manner? Can users retrieve their files and query remote databases anonymously provided they hold appropriate credentials? Lattice-based cryptography is the only approach so far that has allowed to make progress into those directions. We will investigate the practicality of the current constructions, the extension of their properties, and the design of more powerful primitives, such as functional encryption (allowing the recipient to learn only a function of the plaintext message). To achieve these goals, we will in particular focus on cryptographic multilinear maps.

This research axis of AriC is gaining strength thanks to the recruitment of Benoit Libert. We will be particularly interested in the practical and operational impacts, and for this reason we envision a collaboration with an industrial partner.

3.2.3. Application domains

- Diophantine equations. Lattice reduction algorithms can be used to solve diophantine equations, and in particular to find simultaneous rational approximations to real numbers. We plan to investigate the interplay between this algorithmic task, the task of finding integer relations between real numbers, and lattice reduction. A related question is to devise LLL-reduction algorithms that exploit specific shapes of input bases.
- Communications. We will continue our collaboration with Cong Ling (Imperial College) on the use of lattices in communications. We plan to work on the wiretap channel over a fading channel (modeling cell phone communications in a fast moving environment). The current approaches rely

on ideal lattices, and we hope to be able to find new approaches thanks to our expertise on them due to their use in lattice-based cryptography. We will also tackle the problem of sampling vectors from Gaussian distributions with lattice support, for a very small standard deviation parameter. This would significantly improve current schemes for communication schemes based on lattices, as well as several cryptographic primitives.

- Cryptanalysis of variants of RSA. Lattices have been used extensively to break variants of the RSA encryption scheme, via Coppersmith's method to find small roots of polynomials. We plan to work with Nadia Heninger (U. of Pennsylvania) on improving these attacks, to make them more practical. This is an excellent test case for testing the practicality of LLL-type algorithm. Nadia Heninger has a strong experience in large scale cryptanalysis based on Coppersmith's method (http://smartfacts.cr.yp.to/)

3.3. Algebraic computing and high performance kernels

The main theme here is the study of fundamental operations ("kernels") on a hierarchy of symbolic or numeric data types spanning integers, floating-point numbers, polynomials, power series, as well as matrices of all these. Fundamental operations include basic arithmetic (e.g., how to multiply or how to invert) common to all such data, as well as more specific ones (change of representation/conversions, GCDs, determinants, etc.). For such operations, which are ubiquitous and at the very core of computing (be it numerical, symbolic, or hybrid numeric-symbolic), our goal is to ensure both high performance and reliability.

3.3.1. Algorithms

On the symbolic side, we will focus on the design and complexity analysis of algorithms for matrices over various domains (fields, polynomials, integers) and possibly with specific properties (structure). So far, our algorithmic improvements for polynomial matrices and structured matrices have been obtained in a rather independent way. Both types are well known to have much in common, but this is sometimes not reflected by the complexities obtained, especially for applications in cryptology and coding theory. Our goal in this area is thus to explore these connections further, to provide a more unified treatment, and eventually bridge these complexity gaps, A first step towards this goal will be the design of enhanced algorithms for various generalizations of Hermite-Padé approximation; in the context of list decoding, this should in particular make it possible to match or even improve over the structured-matrix approach, which is so far the fastest known.

On the other hand we will focus on the design of algorithms for certified computing. We will study the use of various representations, such as mid-rad for classical interval arithmetic, or affine arithmetic. We will explore the impact of precision tuning in intermediate computations, possibly dynamically, on the accuracy of the results (e.g. for iterative refinement and Newton iterations). We will continue to revisit and improve the classical error bounds of numerical linear algebra in the light of the subtleties of IEEE floating-point arithmetic.

Our goals in linear algebra and lattice basis reduction that have been detailed above in Section 3.2 will be achieved in the light of a hybrid symbolic-numeric approach.

3.3.2. Computer arithmetic

We aim at providing tight error bounds for basic "buiding blocks" of numerical computing. Examples are complex arithmetic (in the continuity of what we have already done), Fourier transforms.

We will also work on the interplay between floating-point and integer arithmetics. Currently, small numerical kernels like an exponential or a 2×2 determinant are typically written using exclusively one of these two kinds of arithmetic. However, modern processors now have hardware support for both floating-point and integer arithmetics, often with vector (SIMD) extensions, and an important question is how to make the best use of all such capabilities to optimize for both accuracy and efficiency.

A third direction will be to work on algorithms for performing correctly-rounded arithmetic operations in medium precision as efficiently and reliably as possible. Indeed, many numerical problems require higher precision than the conventional floating-point (single, double) formats. One solution is to use multiple precision libraries, such as GNU MPFR, which allow the manipulation of very high precision numbers, but their generality (they are able to handle numbers with millions of digits) is a quite heavy alternative when high performance is needed. Our objective here is thus to design a multiple precision arithmetic library that would allow to tackle problems where a precision of a few hundred bits is sufficient, but which have strong performance requirements. Applications include the process of long-term iteration of chaotic dynamical systems ranging from the classical Henon map to calculations of planetary orbits. The designed algorithms will be formally proved.

Finally, our work on the IEEE 1788 standard leads naturally to the development of associated reference libraries for interval arithmetic. A first direction will be to implement IEEE 1788 interval arithmetic within MPFI, our library for interval arithmetic using the arbitrary precision floating-point arithmetic provided by MPFR: indeed, MPFI has been originally developed with definitions and handling of exceptions which are not compliant with IEEE 1788. Another one will be to provide efficient support for multiple-precision intervals, in mid-rad representation and by developing MPFR-based code-generation tools aimed at handling families of functions.

3.3.3. High-performance algorithms and software

The algorithmic developments for medium precision floating-point arithmetic discussed above will lead to high performance implementations on GPUs. As a follow-up of the HPAC project (which ended in December 2015) we shall pursue the design and implementation of high performance linear algebra primitives and algorithms.

4. Application Domains

4.1. Floating-point and Validated Numerics

Our expertise on validated numerics is useful to analyze and improve, and guarantee the quality of numerical results in a wide range of applications including:

- scientific simulation;
- global optimization;
- control theory.

Much of our work, in particular the development of correctly rounded elementary functions, is critical to the

- reproducibility of floating-point computations.

4.2. Cryptography, Cryptology, Communication Theory

Lattice reduction algorithms have direct applications in

- public-key cryptography;
- diophantine equations;
- communications theory.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

Damien Stehlé was nominated IUF junior member.

5.1.2. Book

Publication of the second edition of the "Handbook of Floating-Point Arithmetic" [43].
BEST PAPERS AWARDS :
[42]
G. VILLARD.*On Computing the Resultant of Generic Bivariate Polynomials*, in "ISSAC 2018, 43rd International Symposium on Symbolic and Algebraic Computation, New York, USA, July 16-19, 2018", New York, United States, July 2018, https://hal.archives-ouvertes.fr/hal-01921369

6. New Software and Platforms

6.1. FPLLL

KEYWORDS: Euclidean Lattices - Computer algebra system (CAS) - Cryptography
SCIENTIFIC DESCRIPTION: The fplll library is used or has been adapted to be integrated within several mathematical computation systems such as Magma, Sage, and PariGP. It is also used for cryptanalytic purposes, to test the resistance of cryptographic primitives.
FUNCTIONAL DESCRIPTION: fplll contains implementations of several lattice algorithms. The implementation relies on floating-point orthogonalization, and LLL is central to the code, hence the name.

It includes implementations of floating-point LLL reduction algorithms, offering different speed/guarantees ratios. It contains a 'wrapper' choosing the estimated best sequence of variants in order to provide a guaranteed output as fast as possible. In the case of the wrapper, the succession of variants is oblivious to the user.

It includes an implementation of the BKZ reduction algorithm, including the BKZ-2.0 improvements (extreme enumeration pruning, pre-processing of blocks, early termination). Additionally, Slide reduction and self dual BKZ are supported.

It also includes a floating-point implementation of the Kannan-Fincke-Pohst algorithm that finds a shortest non-zero lattice vector. For the same task, the GaussSieve algorithm is also available in fplll. Finally, it contains a variant of the enumeration algorithm that computes a lattice vector closest to a given vector belonging to the real span of the lattice.

- Author: Damien Stehlé
- Contact: Damien Stehlé
- URL: https://github.com/fplll/fplll

6.2. Gfun

generating functions package
KEYWORD: Symbolic computation
FUNCTIONAL DESCRIPTION: Gfun is a Maple package for the manipulation of linear recurrence or differential equations. It provides tools for guessing a sequence or a series from its first terms, for manipulating rigorously solutions of linear differential or recurrence equations, using the equation as a data-structure.

- Contact: Bruno Salvy
- URL: http://perso.ens-lyon.fr/bruno.salvy/software/the-gfun-package/

6.3. GNU-MPFR

KEYWORDS: Multiple-Precision - Floating-point - Correct Rounding

FUNCTIONAL DESCRIPTION: GNU MPFR is an efficient arbitrary-precision floating-point library with well-defined semantics (copying the good ideas from the IEEE 754 standard), in particular correct rounding in 5 rounding modes. It provides about 80 mathematical functions, in addition to utility functions (assignments, conversions...). Special data (Not a Number, infinities, signed zeros) are handled like in the IEEE 754 standard. GNU MPFR is based on the mpn and mpz layers of the GMP library.

- Participants: Guillaume Hanrot, Paul Zimmermann, Philippe Théveny and Vincent Lefèvre
- Contact: Vincent Lefèvre
- Publications: Correctly Rounded Arbitrary-Precision Floating-Point Summation - Optimized Binary64 and Binary128 Arithmetic with GNU MPFR - Évaluation rapide de fonctions hypergéométriques - Arbitrary Precision Error Analysis for computing $\zeta(s)$ with the Cohen-Olivier algorithm: Complete description of the real case and preliminary report on the general case - MPFR: A Multiple-Precision Binary Floating-Point Library with Correct Rounding. - The Generic Multiple-Precision Floating-Point Addition With Exact Rounding (as in the MPFR Library)
- URL: https://www.mpfr.org/

6.4. Sipe

KEYWORDS: Floating-point - Correct Rounding
FUNCTIONAL DESCRIPTION: Sipe is a mini-library in the form of a C header file, to perform radix-2 floating-point computations in very low precisions with correct rounding, either to nearest or toward zero. The goal of such a tool is to do proofs of algorithms/properties or computations of tight error bounds in these precisions by exhaustive tests, in order to try to generalize them to higher precisions. The currently supported operations are addition, subtraction, multiplication (possibly with the error term), fused multiply-add/subtract (FMA/FMS), and miscellaneous comparisons and conversions. Sipe provides two implementations of these operations, with the same API and the same behavior: one based on integer arithmetic, and a new one based on floating-point arithmetic.

- Participant: Vincent Lefèvre
- Contact: Vincent Lefèvre
- Publications: SIPE: Small Integer Plus Exponent - Sipe: a Mini-Library for Very Low Precision Computations with Correct Rounding
- URL: https://www.vinc17.net/research/sipe/

6.5. LinBox

KEYWORD: Exact linear algebra
FUNCTIONAL DESCRIPTION: LinBox is an open-source C++ template library for exact, high-performance linear algebra computations. It is considered as the reference library for numerous computations (such as linear system solving, rank, characteristic polynomial, Smith normal forms,...) over finite fields and integers with dense, sparse, and structured matrices.

- Participants: Clément Pernet and Thierry Gautier
- Contact: Clément Pernet
- URL: http://linalg.org/

6.6. HPLLL

KEYWORDS: Euclidean Lattices - Computer algebra system (CAS)
FUNCTIONAL DESCRIPTION: Software library for linear algebra and Euclidean lattice problems

- Contact: Gilles Villard
- URL: http://perso.ens-lyon.fr/gilles.villard/hplll/

7. New Results

7.1. Efficient approximation methods

7.1.1. A High Throughput Polynomial and Rational Function Approximations Evaluator

In [21] we present an automatic method for the evaluation of functions via polynomial or rational approximations and its hardware implementation, on FPGAs. These approximations are evaluated using Ercegovac's iterative E-method adapted for FPGA implementation. The polynomial and rational function coefficients are optimized such that they satisfy the constraints of the E-method. We present several examples of practical interest; in each case a resource-efficient approximation is proposed and comparisons are made with alternative approaches.

7.1.2. Continued fractions in power series fields

In [4], we explicitly describe a noteworthy transcendental continued fraction in the field of power series over $\mathbb{Q}$, having irrationality measure equal to 3. This continued fraction is a generating function of a particular sequence in the set $\{1, 2\}$. The origin of this sequence, whose study was initiated in a recent paper, is to be found in another continued fraction, in the field of power series over $\mathbb{F}_3$, which satisfies a simple algebraic equation of degree 4, introduced thirty years ago by D. Robbins.

7.1.3. A Lattice Basis Reduction Approach for the Design of Finite Wordlength FIR Filters

Many applications of finite impulse response (FIR) digital filters impose strict format constraints for the filter coefficients. Such requirements increase the complexity of determining optimal designs for the problem at hand. In [6], we introduce a fast and efficient method, based on the computation of good nodes for polynomial interpolation and Euclidean lattice basis reduction. Experiments show that it returns quasi-optimal finite wordlength FIR filters; compared to previous approaches it also scales remarkably well (length 125 filters are treated in < 9s). It also proves useful for accelerating the determination of optimal finite wordlength FIR filters.

7.1.4. Validated and numerically efficient Chebyshev spectral methods for linear ordinary differential equations

In [7], we develop a validated numerics method for the solution of linear ordinary differential equations (LODEs). A wide range of algorithms (i.e., Runge-Kutta, collocation, spectral methods) exist for numerically computing approximations of the solutions. Most of these come with proofs of asymptotic convergence, but usually, provided error bounds are non-constructive. However, in some domains like critical systems and computer-aided mathematical proofs, one needs validated effective error bounds. We focus on both the theoretical and practical complexity analysis of a so-called *a posteriori* quasi-Newton validation method, which mainly relies on a fixed-point argument of a contracting map. Specifically, given a polynomial approximation, obtained by some numerical algorithm and expressed in Chebyshev basis, our algorithm efficiently computes an accurate and rigorous error bound. For this, we study theoretical properties like compactness, convergence, invertibility of associated linear integral operators and their truncations in a suitable coefficient space of Chebyshev series. Then, we analyze the almost-banded matrix structure of these operators, which allows for very efficient numerical algorithms for both numerical solutions of LODEs and rigorous computation of the approximation error. Finally, several representative examples show the advantages of our algorithms as well as their theoretical and practical limits.

7.1.5. Validated semi-analytical transition matrices for linearized relative spacecraft dynamics via Chebyshev series appproximations

In [14], we provide an efficient generic algorithm to compute validated approximations of transition matrices of linear time-variant systems using Chebyshev expansions, and apply it to two different examples of relative motion of satellites (spacecraft rendezvous with Tschauner-Hempel equations and geostationary station keeping with J2 perturbation in the linearized Orange model).

7.1.6. A Newton-like Validation Method for Chebyshev Approximate Solutions of Linear Ordinary Differential Systems

In [22], we provide a new framework for *a posteriori* validation of vector-valued problems with componentwise tight error enclosures, and use it to design a symbolic-numeric Newton-like validation algorithm for Chebyshev approximate solutions of coupled systems of linear ordinary differential equations. More precisely, given a coupled differential system with polynomial coefficients over a compact interval (or continuous coefficients rigorously approximated by polynomials) and componentwise polynomial approximate solutions in Chebyshev basis, the algorithm outputs componentwise rigorous upper bounds for the approximation errors, with respect to the uniform norm over the interval under consideration.

A complexity analysis shows that the number of arithmetic operations needed by this algorithm (in floating-point or interval arithmetics) is proportional to the approximation degree when the differential equation is considered fixed. Finally, we illustrate the efficiency of this fully automated validation method on an example of a coupled Airy-like system.

7.1.7. Fuel-optimal impulsive fixed-time trajectories in the linearized circular restricted 3-body-problem

In [41], the problem of fixed-time fuel-optimal trajectories with high-thrust propulsion in the vicinity of a Lagrange point is tackled via the linear version of the primer vector theory. More precisely, the proximity to a Lagrange point i.e. any equilibrium point-stable or not-in the circular restricted three-body problem allows for a linearization of the dynamics. Furthermore, it is assumed that the spacecraft has ungimbaled thrusters, leading to a formulation of the cost function with the 1-norm for space coordinates, even though a generalization exists for steerable thrust and the 2-norm. In this context, the primer vector theory gives necessary and sufficient optimality conditions for admissible solutions to two-value boundary problems. Similarly to the case of rendezvous in the restricted two-body problem, the in-plane and out-of-plane trajectories being uncoupled, they can be treated independently. As a matter of fact, the out-of-plane dynamics is simple enough for the optimal control problem to be solved analytically via this indirect approach. As for the in-plane dynamics, the primer vector solution of the so-called primal problem is derived by solving a hierarchy of linear programs, as proposed recently for the aforementioned rendezvous. The optimal thrusting strategy is then numerically obtained from the necessary and sufficient conditions. Finally, in-plane and out-of-plane control laws are combined to form the complete 3-D fuel-optimal solution. Results are compared to the direct approach that consists in working on a discrete set of times in order to perform optimization in finite dimension. Examples are provided near various Lagrange points in the Sun-Earth and Earth-Moon systems, hinting at the extensive span of possible applications of this technique in station-keeping as well as mission analysis, for instance when connecting manifolds to achieve escape or capture.

7.2. Floating-point and Validated Numerics

7.2.1. Optimal bounds on relative errors of floating-point operations

Rounding error analyses of numerical algorithms are most often carried out via repeated applications of the so-called standard models of floating-point arithmetic. Given a round-to-nearest function fl and barring underflow and overflow, such models bound the relative errors $E_1(t) = |t - fl(t)|/|t|$ and $E_2(t) = |t - fl(t)|/|fl(t)|$ by the unit roundoff u. In [10] we investigate the possibility and the usefulness of refining these bounds, both in the case of an arbitrary real t and in the case where t is the exact result of an arithmetic operation on some floating-point numbers. We show that $E_1(t)$ and $E_2(t)$ are optimally bounded by $u/(1+u)$ and u, respectively, when t is real or, under mild assumptions on the base and the precision, when $t = x \pm y$ or $t = xy$ with x, y two floating-point numbers. We prove that while this remains true for division in base $\beta > 2$, smaller, attainable bounds can be derived for both division in base $\beta = 2$ and square root. This set of optimal bounds is then applied to the rounding error analysis of various numerical algorithms: in all cases, we obtain significantly shorter proofs of the best-known error bounds for such algorithms, and/or improvements on these bounds themselves.

7.2.2. On various ways to split a floating-point number

In [32] we review several ways to split a floating-point number, that is, to decompose it into the exact sum of two floating-point numbers of smaller precision. All the methods considered here involve only a few IEEE floating-point operations, with rounding to nearest and including possibly the fused multiply-add (FMA). Applications range from the implementation of integer functions such as `round` and `floor` to the computation of suitable scaling factors aimed, for example, at avoiding spurious underflows and overflows when implementing functions such as the hypotenuse.

7.2.3. Algorithms for triple-word arithmetic

Triple-word arithmetic consists in representing high-precision numbers as the unevaluated sum of three floating-point numbers. In [45], we introduce and analyze various algorithms for manipulating triple-word numbers. Our new algorithms are faster than what one would obtain by just using the usual floating-point expansion algorithms in the special case of expansions of length 3, for a comparable accuracy.

7.2.4. Error analysis of some operations involved in the Fast Fourier Transform

In [44], we are interested in obtaining error bounds for the classical FFT algorithm in floating-point arithmetic, for the 2-norm as well as for the infinity norm. For that purpose we also give some results on the relative error of the complex multiplication by a root of unity, and on the largest value that can take the real or imaginary part of one term of the FFT of a vector x, assuming that all terms of x have real and imaginary parts less than some value b.

7.3. Lattices: algorithms and cryptology

7.3.1. Reduction of orthogonal lattice bases

As a typical application, the LLL lattice basis reduction algorithm is applied to bases of the orthogonal lattice of a given integer matrix, via reducing lattice bases of a special type. With such bases in input, we have proposed in [26] a new technique for bounding from above the number of iterations required by the LLL algorithm. The main technical ingredient is a variant of the classical LLL potential, which could prove useful to understand the behavior of LLL for other families of input bases.

7.3.2. Lattice-Based Zero-Knowledge Arguments for Integer Relations

The paper [36] provides lattice-based protocols allowing to prove relations among committed integers. While the most general zero-knowledge proof techniques can handle arithmetic circuits in the lattice setting, adapting them to prove statements over the integers is non-trivial, at least if we want to handle exponentially large integers while working with a polynomial-size modulus q. For a polynomial L, the paper provides zero-knowledge arguments allowing a prover to convince a verifier that committed L-bit bitstrings x, y and z are the binary representations of integers X, Y and Z satisfying $Z = X + Y$ over $\mathbb{Z}$. The complexity of the new arguments is only linear in L. Using them, the paper constructs arguments allowing to prove inequalities $X < Z$ among committed integers, as well as arguments showing that a committed X belongs to a public interval $[\alpha, \beta]$, where α and β can be arbitrarily large. The new range arguments have logarithmic cost (i.e., linear in L) in the maximal range magnitude. Using these tools, the paper obtains zero-knowledge arguments showing that a committed element X does not belong to a public set S using $O(n \cdot \log |S|)$ bits of communication, where n is the security parameter. The paper finally gives a protocol allowing to argue that committed L-bit integers X, Y and Z satisfy multiplicative relations $Z = XY$ over the integers, with communication cost subquadratic in L. To this end, the paper uses its new protocol for integer addition to prove the correct recursive execution of Karatsuba's multiplication algorithm. The security of the new protocols relies on standard lattice assumptions with polynomial modulus and polynomial approximation factor.

7.3.3. Logarithmic-Size Ring Signatures With Tight Security from the DDH Assumption

Ring signatures make it possible for a signer to anonymously and, yet, convincingly leak a secret by signing a message while concealing his identity within a flexibly chosen ring of users. Unlike group signatures, they do not involve any setup phase or tracing authority. Despite a lot of research efforts in more than 15 years, most of their realizations require linear-size signatures in the cardinality of the ring. In the random oracle model, two recent constructions decreased the signature length to be only logarithmic in the number N of ring members. On the downside, their suffer from rather loose reductions incurred by the use of the Forking Lemma. This paper considers the problem of proving them tightly secure without affecting their space efficiency. Surprisingly, existing techniques for proving tight security in ordinary signature schemes do not trivially extend to the ring signature setting. The paper [37] overcomes these difficulties by combining the Groth-Kohlweiss Σ-protocol (Eurocrypt'15) with dual-mode encryption schemes. The main result is a fully tight construction based on the Decision Diffie-Hellman assumption in the random oracle model. By full tightness, we mean that the reduction's advantage is as large as the adversary's, up to a constant factor.

7.3.4. Adaptively Secure Distributed PRFs from LWE

In distributed pseudorandom functions (DPRFs), a PRF secret key SK is secret shared among N servers so that each server can locally compute a partial evaluation of the PRF on some input X. A combiner that collects t partial evaluations can then reconstruct the evaluation $F(SK, X)$ of the PRF under the initial secret key. So far, all non-interactive constructions in the standard model are based on lattice assumptions. One caveat is that they are only known to be secure in the static corruption setting, where the adversary chooses the servers to corrupt at the very beginning of the game, before any evaluation query. The paper [38] constructs the first fully non-interactive adaptively secure DPRF in the standard model. The construction is proved secure under the LWE assumption against adversaries that may adaptively decide which servers they want to corrupt. The new construction is also extended in order to achieve robustness against malicious adversaries.

7.3.5. Unbounded ABE via Bilinear Entropy Expansion, Revisited

This paper [24] presents simpler and improved constructions of unbounded attribute-based encryption (ABE) schemes with constant-size public parameters under static assumptions in bilinear groups. Concretely, we obtain: a simple and adaptively secure unbounded ABE scheme in composite-order groups, improving upon a previous construction of Lewko and Waters (Eurocrypt'11) which only achieves selective security; an improved adaptively secure unbounded ABE scheme based on the k-linear assumption in prime-order groups with shorter ciphertexts and secret keys than those of Okamoto and Takashima (Asiacrypt'12); the first adaptively secure unbounded ABE scheme for arithmetic branching programs under static assumptions. At the core of all of these constructions is a "bilinear entropy expansion" lemma that allows us to generate any polynomial amount of entropy starting from constant-size public parameters; the entropy can then be used to transform existing adaptively secure "bounded" ABE schemes into unbounded ones.

7.3.6. Improved Anonymous Broadcast Encryptions: Tight Security and Shorter Ciphertext

This paper [35] investigates anonymous broadcast encryptions (ANOBE) in which a ciphertext hides not only the message but also the target recipients associated with it. Following Libert et al.'s generic construction [PKC, 2012], we propose two concrete ANOBE schemes with tight reduction and better space efficiency.

- The IND-CCA security and anonymity of our two ANOBE schemes can be tightly reduced to standard k-Linear assumption (and the existence of other primitives). For a broadcast system with n users, Libert et al.'s security analysis suffers from $\mathcal{O}(n^3)$ loss while our security loss is constant.
- Our first ANOBE supports fast decryption and has a shorter ciphertext than the fast-decryption version of Libert et al.'s concrete ANOBE. Our second ANOBE is adapted from the first one. We sacrifice the fast decryption feature and achieve shorter ciphertexts than Libert et al.'s concrete ANOBE with the help of bilinear groups. Technically, we start from an instantiation of Libert et al.'s generic ANOBE [PKC, 2012], but we work out all our proofs from scratch instead of relying on their generic security result. This intuitively allows our optimizations in the concrete setting.

7.3.7. Compact IBBE and Fuzzy IBE from Simple Assumptions

This paper [29] proposes new constructions for identity-based broadcast encryption (IBBE) and fuzzy identity-based encryption (FIBE) in composite-order groups equipped with a bilinear pairing. Our starting point is the IBBE scheme of Delerablée (Asiacrypt 2007) and the FIBE scheme of Herranz et al. (PKC 2010) proven secure under parameterized assumptions called generalized decisional bilinear Diffie-Hellman (GDDHE) and augmented multi-sequence of exponents Diffie-Hellman (aMSE-DDH) respectively. The two schemes are described in the prime-order pairing group. We transform the schemes into the setting of (symmetric) composite-order groups and prove security from two static assumptions (subgroup decision). The Déjà Q framework of Chase et al. (Asiacrypt 2016) is known to cover a large class of parameterized assumptions (dubbed "Uber assumption"), that is, these assumptions, when defined in asymmetric composite-order groups, are implied by subgroup decision assumptions in the underlying composite-order groups. We argue that the GDDHE and aMSE-DDH assumptions are not covered by the Déjà Q uber assumption framework. We therefore work out direct security reductions for the two schemes based on subgroup decision assumptions. Furthermore, our proofs involve novel extensions of Déjà Q techniques of Wee (TCC 2016-A) and Chase et al. Our constructions have constant-size ciphertexts. The IBBE has constant-size keys as well and achieves a stronger security guarantee as compared to Delerablée's IBBE, thus making it the first compact IBBE known to be selectively secure without random oracles under simple assumptions. The fuzzy IBE scheme is the first to simultaneously feature constant-size ciphertexts and security under standard assumptions.

7.3.8. Improved Inner-product Encryption with Adaptive Security and Full Attribute-hiding

This paper [25] proposes two IPE schemes achieving both adaptive security and full attribute-hiding in the prime-order bilinear group, which improve upon the unique existing result satisfying both features from Okamoto and Takashima [Eurocrypt'12] in terms of efficiency.

- Our first IPE scheme is based on the standard k-Lin assumption and has shorter master public key and shorter secret keys than Okamoto and Takashima's IPE under weaker $DLIN$=2-lin assumption.
- Our second IPE scheme is adapted from the first one; the security is based on the XDLIN assumption (as Okamoto and Takashima's IPE) but now it also enjoys shorter ciphertexts.

Technically, instead of starting from composite-order IPE and applying existing transformation, we start from an IPE scheme in a very restricted setting but already in the prime-order group, and then gradually upgrade it to our full-fledged IPE scheme. This method allows us to integrate Chen et al.'s framework [Eurocrypt'15] with recent new techniques [TCC'17, Eurocrypt'18] in an optimized way.

7.3.9. Improved Security Proofs in Lattice-Based Cryptography: Using the Rényi Divergence Rather than the Statistical Distance

The Rényi divergence is a measure of closeness of two probability distributions. In this paper [5], we show that it can often be used as an alternative to the statistical distance in security proofs for lattice-based cryptography. Using the Rényi divergence is particularly suited for security proofs of primitives in which the attacker is required to solve a search problem (e.g., forging a signature). We show that it may also be used in the case of distinguishing problems (e.g., semantic security of encryption schemes), when they enjoy a public sampleability property. The techniques lead to security proofs for schemes with smaller parameters, and sometimes to simpler security proofs than the existing ones.

7.3.10. CRYSTALS-Dilithium: A Lattice-Based Digital Signature Scheme

This paper [8] presents Dilithium, a lattice-based signature scheme that is part of the CRYSTALS (Cryptographic Suite for Algebraic Lattices) package that will be submitted to the NIST call for post-quantum standards. The scheme is designed to be simple to securely implement against side-channel attacks and to have comparable efficiency to the currently best lattice-based signature schemes. Our implementation results show that Dilithium is competitive with lattice schemes of the same security level and outperforms digital signature schemes based on other post-quantum assumptions.

7.3.11. On the asymptotic complexity of solving LWE

In this paper [9], we provide for the first time an asymptotic comparison of all known algorithms for the search version of the Learning with Errors (LWE) problem. This includes an analysis of several lattice-based approaches as well as the combinatorial BKW algorithm. Our analysis of the lattice-based approaches defines a general framework, in which the algorithms of Babai, Lindner–Peikert and several pruning strategies appear as special cases. We show that within this framework, all lattice algorithms achieve the same asymptotic complexity. For the BKW algorithm, we present a refined analysis for the case of only a polynomial number of samples via amplification, which allows for a fair comparison with lattice-based approaches. Somewhat surprisingly, such a small number of samples does not make the asymptotic complexity significantly inferior, but only affects the constant in the exponent. As the main result we obtain that both, lattice-based techniques and BKW with a polynomial number of samples, achieve running time $2^{O(n)}$ for n-dimensional LWE, where we make the constant hidden in the big-O notion explicit as a simple and easy to handle function of all LWE-parameters. In the lattice case this function also depends on the time to compute a BKZ lattice basis with block size $\Theta(n)$. Thus, from a theoretical perspective our analysis reveals how LWE 's complexity changes as a function of the LWE-parameters, and from a practical perspective our analysis is a useful tool to choose LWE-parameters resistant to all currently known attacks.

7.3.12. Measuring, Simulating and Exploiting the Head Concavity Phenomenon in BKZ

The Blockwise-Korkine-Zolotarev (BKZ) lattice reduction algorithm is central in cryptanalysis, in particular for lattice-based cryptography. A precise understanding of its practical behavior in terms of run-time and output quality is necessary for parameter selection in cryptographic design. As the provable worst-case bounds poorly reflect the practical behavior, cryptanalysts rely instead on the heuristic BKZ simulator of Chen and Nguyen (Asiacrypt'11). It fits better with practical experiments, but not entirely. In particular, it over-estimates the norm of the first few vectors in the output basis. Put differently, BKZ performs better than its Chen-Nguyen simulation.

In this article [15], we first report experiments providing more insight on this shorter-than-expected phenomenon. We then propose a refined BKZ simulator by taking the distribution of short vectors in random lattices into consideration. We report experiments suggesting that this refined simulator more accurately predicts the concrete behavior of BKZ. Furthermore, we design a new BKZ variant that exploits the shorter-than-expected phenomenon. For the same cost assigned to the underlying SVP-solver, the new BKZ variant produces bases of better quality. We further illustrate its potential impact by testing it on the SVP-120 instance of the Darmstadt lattice challenge.

7.3.13. CRYSTALS - Kyber: A CCA-Secure Module-Lattice-Based KEM

Rapid advances in quantum computing, together with the announcement by the National Institute of Standards and Technology (NIST) to define new standards for digital signature, encryption, and key-establishment protocols, have created significant interest in post-quantum cryptographic schemes. This paper [17] introduces Kyber (part of CRYSTALS - Cryptographic Suite for Algebraic Lattices - a package submitted to NIST post-quantum standardization effort in November 2017), a portfolio of post-quantum cryptographic primitives built around a key-encapsulation mechanism (KEM), based on hardness assumptions over module lattices. Our KEM is most naturally seen as a successor to the NEWHOPE KEM (Usenix 2016). In particular, the key and ciphertext sizes of our new construction are about half the size, the KEM offers CCA instead of only passive security, the security is based on a more general (and flexible) lattice problem, and our optimized implementation results in essentially the same running time as the aforementioned scheme. We first introduce a CPA-secure public-key encryption scheme, apply a variant of the Fujisaki-Okamoto transform to create a CCA-secure KEM, and eventually construct, in a black-box manner, CCA-secure encryption, key exchange, and authenticated-key-exchange schemes. The security of our primitives is based on the hardness of Module-LWE in the classical and quantum random oracle models, and our concrete parameters conservatively target more than 128 bits of postquantum security.

7.3.14. Learning with Errors and Extrapolated Dihedral Cosets

The hardness of the learning with errors (LWE) problem is one of the most fruitful resources of modern cryptography. In particular, it is one of the most prominent candidates for secure post-quantum cryptography. Understanding its quantum complexity is therefore an important goal. In this paper [20], we show that under quantum polynomial time reductions, LWE is equivalent to a relaxed version of the dihedral coset problem (DCP), which we call extrapolated DCP (eDCP). The extent of extrapolation varies with the LWE noise rate. By considering different extents of extrapolation, our result generalizes Regev's famous proof that if DCP is in BQP (quantum poly-time) then so is LWE (FOCS'02). We also discuss a connection between eDCP and Childs and Van Dam's algorithm for generalized hidden shift problems (SODA'07). Our result implies that a BQP solution for LWE might not require the full power of solving DCP, but rather only a solution for its relaxed version, eDCP, which could be easier.

7.3.15. Pairing-friendly twisted Hessian curves

This paper [27] presents efficient formulas to compute Miller doubling and Miller addition utilizing degree-3 twists on curves with j-invariant 0 written in Hessian form. We give the formulas for both odd and even embedding degrees and for pairings on both $G_1 \times G_2$ and $G_2 \times G_1$. We propose the use of embedding degrees 15 and 21 for 128-bit and 192-bit security respectively in light of the NFS attacks and their variants. We give a comprehensive comparison with other curve models; our formulas give the fastest known pairing computation for embedding degrees 15, 21, and 24.

7.3.16. On the Statistical Leak of the GGH13 Multilinear Mapand some Variants

At EUROCRYPT 2013, Garg, Gentry and Halevi proposed a candidate construction (later referred as GGH13) of cryptographic multilinear map (MMap). Despite weaknesses uncovered by Hu and Jia (EUROCRYPT 2016), this candidate is still used for designing obfuscators. The naive version of the GGH13 scheme was deemed susceptible to averaging attacks, i.e., it could suffer from a statistical leak (yet no precise attack was described). A variant was therefore devised, but it remains heuristic. Recently, to obtain MMaps with low noise and modulus, two variants of this countermeasure were developed by Döttling et al. (EPRINT:2016/599). In this work [28], we propose a systematic study of this statistical leak for all these GGH13 variants. In particular, we confirm the weakness of the naive version of GGH13. We also show that, among the two variants proposed by Döttling et al., the so-called conservative method is not so effective: it leaks the same value as the unprotected method. Luckily, the leak is more noisy than in the unprotected method, making the straightforward attack unsuccessful. Additionally, we note that all the other methods also leak values correlated with secrets. As a conclusion, we propose yet another countermeasure, for which this leak is made unrelated to all secrets. On our way, we also make explicit and tighten the hidden exponents in the size of the parameters, as an effort to assess and improve the efficiency of MMaps.

7.3.17. Higher dimensional sieving for the number field sieve algorithms

Since 2016 and the introduction of the exTNFS (extended tower number field sieve) algorithm, the security of cryptosystems based on nonprime finite fields, mainly the pairing- and torus-based ones, is being reassessed. The feasibility of the relation collection, a crucial step of the NFS variants, is especially investigated. It usually involves polynomials of degree 1, i.e., a search space of dimension 2. However, exTNFS uses bivariate polynomials of at least four coefficients. If sieving in dimension 2 is well described in the literature, sieving in higher dimensions has received significantly less attention. In this work [30], we describe and analyze three different generic algorithms to sieve in any dimension for the NFS algorithms. Our implementation shows the practicability of dimension-4 sieving, but the hardness of dimension-6 sieving.

7.3.18. Speed-Ups and Time-Memory Trade-Offs for Tuple Lattice Sieving

In this work [31], we study speed-ups and time–space trade-offs for solving the shortest vector problem (SVP) on Euclidean lattices based on tuple lattice sieving. Our results extend and improve upon previous work of Bai–Laarhoven–Stehlé [ANTS'16] and Herold–Kirshanova [PKC'17], with better complexities for arbitrary tuple sizes and offering tunable time–memory tradeoffs. The trade-offs we obtain stem from the generalization and combination of two algorithmic techniques: the configuration framework introduced by Herold–Kirshanova, and the spherical locality-sensitive filters of Becker–Ducas–Gama–Laarhoven [SODA'16]. When the available memory scales quasi-linearly with the list size, we show that with triple sieving we can solve SVP in dimension n in time $2^{0.3588n+o(n)}$ and space $2^{0.1887n+o(n)}$, improving upon the previous best triple sieve time complexity of $2^{0.3717n+o(n)}$ of Herold–Kirshanova. Using more memory we obtain better asymptotic time complexities. For instance, we obtain a triple sieve requiring only $2^{0.3300n+o(n)}$ time and $2^{0.2075n+o(n)}$ memory to solve SVP in dimension n. This improves upon the best double Gauss sieve of Becker–Ducas–Gama–Laarhoven, which runs in $2^{0.3685n+o(n)}$ time when using the same amount of space.

7.3.19. Improved Quantum Information Set Decoding

In this paper [34], we present quantum information set decoding (ISD) algorithms for binary linear codes. First, we refine the analysis of the quantum walk based algorithms proposed by Kachigar and Tillich (PQCrypto'17). This refinement allows us to improve the running time of quantum decoding in the leading order term: for an n-dimensional binary linear code the complexity of May-Meurer-Thomae ISD algorithm (Asiacrypt'11) drops down from $2^{0.05904n+o(n)}$ to $2^{0.05806n+o(n)}$. Similar improvement is achieved for our quantum version of Becker-JeuxMay-Meurer (Eurocrypt'12) decoding algorithm. Second, we translate May-Ozerov Near Neighbour technique (Eurocrypt'15) to an 'updateand-query' language more common in a similarity search literature. This re-interpretation allows us to combine Near Neighbour search with the quantum walk framework and use both techniques to improve a quantum version of Dumer's ISD algorithm: the running time goes down from $2^{0.059962n+o(n)}$ to $2^{0.059450+o(n)}$.

7.3.20. Quantum Attacks against Indistinguishablility Obfuscators Proved Secure in the Weak Multilinear Map Model

We present in [39] a quantum polynomial time attack against the GMMSSZ branching program obfuscator of Garg et al. (TCC'16), when instantiated with the GGH13 multilinear map of Garg et al. (EUROCRYPT'13). This candidate obfuscator was proved secure in the weak multilinear map model introduced by Miles et al. (CRYPTO'16). Our attack uses the short principal ideal solver of Cramer et al. (EUROCRYPT'16), to recover a secret element of the GGH13 multilinear map in quantum polynomial time. We then use this secret element to mount a (classical) polynomial time mixed-input attack against the GMMSSZ obfuscator. The main result of this article can hence be seen as a classical reduction from the security of the GMMSSZ obfuscator to the short principal ideal problem (the quantum setting is then only used to solve this problem in polynomial time). As an additional contribution, we explain how the same ideas can be adapted to mount a quantum polynomial time attack against the DGGMM obfuscator of Döttling et al. (ePrint 2016), which was also proved secure in the weak multilinear map model.

7.3.21. On the Ring-LWE and Polynomial-LWE Problems

The Ring Learning With Errors problem (RLWE) comes in various forms. Vanilla RLWE is the decision dual-RLWE variant, consisting in distinguishing from uniform a distribution depending on a secret belonging to the dual $O_K^\vee$ of the ring of integers O_K of a specified number field K. In primal-RLWE, the secret instead belongs to O_K. Both decision dual-RLWE and primal-RLWE enjoy search counterparts. Also widely used is (search/decision) Polynomial Learning With Errors (PLWE), which is not defined using a ring of integers O_K of a number field K but a polynomial ring $Z[x]/f$ for a monic irreducible $f \in Z[x]$. We show that there exist reductions between all of these six problems that incur limited parameter losses. More precisely: we prove that the (decision/search) dual to primal reduction from Lyubashevsky et al. [EUROCRYPT 2010] and Peikert [SCN 2016] can be implemented with a small error rate growth for all rings (the resulting reduction is nonuniform polynomial time); we extend it to polynomial-time reductions between (decision/search) primal RLWE and PLWE that work for a family of polynomials f that is exponentially large as a function of deg(f) (the resulting reduction is also non-uniform polynomial time); and we exploit the recent technique from Peikert et al. [STOC 2017] to obtain a search to decision reduction for RLWE for arbitrary number fields. The reductions incur error rate increases that depend on intrinsic quantities related to K and f.

7.3.22. Non-Trivial Witness Encryption and Null-iO from Standard Assumptions

A *witness encryption (WE)* scheme can take any NP statement as a public-key and use it to encrypt a message. If the statement is true then it is possible to decrypt the message given a corresponding witness, but if the statement is false then the message is computationally hidden. Ideally, the encryption procedure should run in polynomial time, but it is also meaningful to define a weaker notion, which we call *non-trivially exponentially efficient* WE (XWE), where the encryption run-time is only required to be much smaller than the trivial 2^m bound for NP relations with witness size m. In [19], we show how to construct such XWE schemes for all of NP with encryption run-time $2^{m/2}$ under the sub-exponential learning with errors (LWE) assumption. For NP relations that can be verified in NC^1 (e.g., SAT) we can also construct such XWE schemes under the sub-exponential Decisional Bilinear Diffie-Hellman (DBDH) assumption. Although we find the result surprising, it follows via a very simple connection to *attribute-based encryption*.

We also show how to upgrade the above results to get non-trivially exponentially efficient *indistinguishability obfuscation for null circuits (niO)*, which guarantees that the obfuscations of any two circuits that always output 0 are indistinguishable. In particular, under the LWE assumptions we get a XniO scheme where the obfuscation time is $2^{n/2}$ for all circuits with input size n. It is known that in the case of indistinguishability obfuscation (iO) for all circuits, non-trivially efficient XiO schemes imply fully efficient iO schemes (Lin et al., PKC 2016) but it remains as a fascinating open problem whether any such connection exists for WE or niO.

Lastly, we explore a potential approach toward constructing fully efficient WE and niO schemes via multi-input ABE.

7.3.23. Function-Revealing Encryption

Multi-input functional encryption is a paradigm that allows an authorized user to compute a certain function—and nothing more—over multiple plaintexts given only their encryption. The particular case of two-input functional encryption has very exciting applications, including comparing the relative order of two plaintexts from their encrypted form (order-revealing encryption).

While being extensively studied, multi-input functional encryption is not ready for a practical deployment, mainly for two reasons. First, known constructions rely on heavy cryptographic tools such as multilinear maps. Second, their security is still very uncertain, as revealed by recent devastating attacks.

In [33], we investigate a simpler approach towards obtaining practical schemes for functions of particular interest. We introduce the notion of function-revealing encryption, a generalization of order-revealing encryption to any multi-input function as well as a relaxation of multi-input functional encryption. We then propose a simple construction of order-revealing encryption based on function-revealing encryption for simple functions, namely orthogonality testing and intersection cardinality. Our main result is an efficient order-revealing encryption scheme with limited leakage based on the standard DLin assumption.

7.3.24. Exploring Crypto Dark Matter: New Simple PRF Candidates and Their Applications

Pseudorandom functions (PRFs) are one of the fundamental building blocks in cryptography. We explore a new space of plausible PRF candidates that are obtained by mixing linear functions over different small moduli. Our candidates are motivated by the goals of maximizing simplicity and minimizing complexity measures that are relevant to cryptographic applications such as secure multiparty computation.

In [16], we present several concrete new PRF candidates that follow the above approach. Our main candidate is a *weak* PRF candidate (whose conjectured pseudorandomness only holds for uniformly random inputs) that first applies a secret mod-2 linear mapping to the input, and then a public mod-3 linear mapping to the result. This candidate can be implemented by depth-2 ACC^0 circuits. We also put forward a similar depth-3 *strong* PRF candidate. Finally, we present a different weak PRF candidate that can be viewed as a deterministic variant of "Learning Parity with Noise" (LPN) where the noise is obtained via a mod-3 inner product of the input and the key.

The advantage of our approach is twofold. On the theoretical side, the simplicity of our candidates enables us to draw natural connections between their hardness and questions in complexity theory or learning theory (e.g., learnability of depth-2 ACC^0 circuits and width-3 branching programs, interpolation and property testing for sparse polynomials, and natural proof barriers for showing super-linear circuit lower bounds). On the applied side, the "piecewise-linear" structure of our candidates lends itself nicely to applications in secure multiparty computation (MPC). Using our PRF candidates, we construct protocols for distributed PRF evaluation that achieve better round complexity and/or communication complexity (often both) compared to protocols obtained by combining standard MPC protocols with PRFs like AES, LowMC, or Rasta (the latter two are specialized MPC-friendly PRFs). Our advantage over competing approaches is maximized in the setting of MPC with an honest majority, or alternatively, MPC with preprocessing.

Finally, we introduce a new primitive we call an *encoded-input PRF*, which can be viewed as an interpolation between weak PRFs and standard (strong) PRFs. As we demonstrate, an encoded-input PRF can often be used as a drop-in replacement for a strong PRF, combining the efficiency benefits of weak PRFs and the security benefits of strong PRFs. We conclude by showing that our main weak PRF candidate can plausibly be boosted to an encoded-input PRF by leveraging error-correcting codes.

7.3.25. Related-Key Security for Pseudorandom Functions Beyond the Linear Barrier

Related-key attacks (RKAs) concern the security of cryptographic primitives in the situation where the key can be manipulated by the adversary. In the RKA setting, the adversary's power is expressed through the class of related-key deriving (RKD) functions which the adversary is restricted to using when modifying keys. Bellare and Kohno (Eurocrypt 2003) first formalised RKAs and pin-pointed the foundational problem of constructing RKA-secure pseudorandom functions (RKA-PRFs). To date there are few constructions for RKA-PRFs under standard assumptions, and it is a major open problem to construct RKA-PRFs for larger classes of RKD functions. We make significant progress on this problem. In [3], we first show how to repair the Bellare-Cash framework for constructing RKA-PRFs and extend it to handle the more challenging case of classes of RKD functions that contain claws. We apply this extension to show that a variant of the NaorReingold function already considered by Bellare and Cash is an RKA-PRF for a class of affine RKD functions under the DDH assumption, albeit with an exponential-time security reduction. We then develop a second extension of the Bellare-Cash framework, and use it to show that the same Naor-Reingold variant is actually an RKA-PRF for a class of degree d polynomial RKD functions under the stronger decisional d-Diffie-Hellman inversion assumption. As a significant technical contribution, our proof of this result avoids the exponential-time security reduction that was inherent in the work of Bellare and Cash and in our first result.

7.3.26. Practical Fully Secure Unrestricted Inner Product Functional Encryption modulo p

In [23], we provide adaptively secure functional encryption schemes for the inner product functionality which are both efficient and allow for the evaluation of unbounded inner products modulo a prime p. Our constructions rely on new natural cryptographic assumptions in a cyclic group containing a subgroup where the discrete logarithm (DL) problem is easy which extend Castagnos and Laguillaumie's assumption (RSA 2015) of a DDH group with an easy DL subgroup. Instantiating our generic construction using class groups of imaginary quadratic fields gives rise to the most efficient functional encryption for inner products modulo an arbitrary large prime p. One of our schemes outperforms the DCR variant of Agrawal et al.'s protocols in terms of size of keys and ciphertexts by factors varying between 2 and 20 for a 112-bit security.

7.4. Algebraic computing and high-performance kernels

7.4.1. Generalized Hermite Reduction, Creative Telescoping and Definite Integration of D-Finite Functions

Hermite reduction is a classical algorithmic tool in symbolic integration. It is used to decompose a given rational function as a sum of a function with simple poles and the derivative of another rational function. In [18], we extend Hermite reduction to arbitrary linear differential operators instead of the pure derivative, and develop efficient algorithms for this reduction. We then apply the generalized Hermite reduction to the computation of linear operators satisfied by single definite integrals of D-finite functions of several continuous or discrete parameters. The resulting algorithm is a generalization of reduction-based methods for creative telescoping.

7.4.2. Hermite-Padé approximant bases

In [46] we design fast algorithms for the computation of approximant bases in shifted Popov normal form. For K a commutative field, let F be a matrix in $\mathsf{K}[x]^{m \times n}$ (truncated power series) and $\overrightarrow{d}$ be a degree vector, the problem is to compute a basis $P \in \mathsf{K}[x]^{m \times m}$ of the $\mathsf{K}[x]$-module of the relations $p \in \mathsf{K}[x]^{1 \times m}$ such that $p(x) \cdot F(x) \equiv 0 \mod x^{\overrightarrow{d}}$. We obtain improved complexity bounds for handling arbitrary (possibly highly unbalanced) vectors $\overrightarrow{d}$. We also improve upon previously known algorithms for computing P in normalized shifted form for an arbitrary shift. Our approach combines a recent divide and conquer strategy which reduces the general case to the case where information on the output degree is available, and partial linearizations of the involved matrices.

7.4.3. Resultant of bivariate polynomials

We have proposed in [42] an algorithm for computing the resultant of two generic bivariate polynomials over a field K. For such p and q in $\mathsf{K}[x, y]$ both of degree d in x and n in y, the algorithm computes the resultant with respect to y using $(n^{2-1/\omega}d)^{1+o(1)}$ arithmetic operations, where ω is the exponent of matrix multiplication. Previous algorithms from the early 1970's required time $(n^2 d)^{1+o(1)}$. We have also described some extensions of the approach to the computation of generic Gröbner bases and of characteristic polynomials of generic structured matrices and in univariate quotient algebras.

7.4.4. Recursive Combinatorial Structures: Enumeration, Probabilistic Analysis and Random Generation

The probabilistic behaviour of many data-structures, like series-parallel graphs used as a running example is this tutorial [13], can be analysed very precisely, thanks to a set of high-level tools provided by Analytic Combinatorics, as described in the book by Flajolet and Sedgewick. In this framework, recursive combinatorial definitions lead to generating function equations from which efficient algorithms can be designed for enumeration, random generation and, to some extent, asymptotic analysis. With a focus on random generation, this tutorial given at STACS first covers the basics of Analytic Combinatorics and then describes the idea of Boltzmann sampling and its realisation. The tutorial addresses a broad TCS audience and no particular pre-knowledge on analytic combinatorics is expected.

7.4.5. Linear Differential Equations as a Data-Structure

A lot of information concerning solutions of linear differential equations can be computed directly from the equation. It is therefore natural to consider these equations as a data-structure, from which mathematical properties can be computed. A variety of algorithms has thus been designed in recent years that do not aim at "solving", but at computing with this representation. Many of these results are surveyed in [11].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Bosch (Germany) ordered from us some support for implementing complex numerical algorithms (participants: Claude-Pierre Jeannerod and Jean-Michel Muller).

8.2. Bilateral Grants with Industry

- Miruna Rosca and Radu Titiu are employees of BitDefender. Their PhD's are supervised by Damien Stehlé and Benoît Libert, respectively. Miruna Rosca works on the foundations of lattice-based cryptography, and Radu Titiu works on pseudo-random functions and functional encryption.
- Adel Hamdi is doing is PhD with Orange Labs and is supervised by Fabien Laguillaumie. He is working on advanced encryption protocols for the cloud.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR DYNA3S Project

Participants: Guillaume Hanrot, Gilles Villard.

Dyna3S has been a 2013-2018 ANR project headed by Valérie Berthé (IRIF, U. Paris 7). The Web page of the project is https://www.irif.fr/~dyna3s. The aim of Dyna3S was to study algorithms that compute the greatest common divisor (gcd) from the point of view of dynamical systems. A gcd algorithm is considered as a discrete dynamical system by focusing on integer input. In Lyon we have worked on the computation of the gcd of several integers, in link with integer relation algorithms based on lattice basis reduction. A main motivation of Dyna3S was also discrete geometry, a framework where the understanding of basic primitives, discrete lines and planes, relies on algorithms of the Euclidean type.

9.1.2. ANR FastRelax Project

Participants: Nicolas Brisebarre, Guillaume Hanrot, Vincent Lefèvre, Jean-Michel Muller, Bruno Salvy, Serge Torres.

FastRelax stands for "Fast and Reliable Approximation". It is a four year ANR project (started in October 2014 and extended till September 2019). The web page of the project is http://fastrelax.gforge.inria.fr/. It is headed by B. Salvy and involves AriC as well as members of the Marelle Team (Sophia), of the Mac group (LAAS, Toulouse), of the Specfun and Toccata Teams (Saclay), as well as of the Pequan group in UVSQ and a colleague in the Plume group of LIP.

The aim of this project is to develop computer-aided proofs of numerical values, with certified and reasonably tight error bounds, without sacrificing efficiency. Applications to zero-finding, numerical quadrature or global optimization can all benefit from using our results as building blocks. We expect our work to initiate a "fast and reliable" trend in the symbolic-numeric community. This will be achieved by developing interactions between our fields, designing and implementing prototype libraries and applying our results to concrete problems originating in optimal control theory.

9.1.3. ANR MetaLibm Project

Participants: Claude-Pierre Jeannerod, Jean-Michel Muller.

MetaLibm is a four-year project (started in October 2013 and extended till March 2018) focused on the design and implementation of code generators for mathematical functions and filters. The web page of the project is http://www.metalibm.org/ANRMetaLibm/. It is headed by Florent de Dinechin (INSA Lyon and Socrate team) and, besides Socrate and AriC, also involves teams from LIRMM (Perpignan), LIP6 (Paris), CERN (Geneva), and Kalray (Grenoble). The main goals of the project are to automate the development of mathematical libraries (libm), to extend it beyond standard functions, and to make it unified with similar approaches developed in or useful for signal processing (filter design). Within AriC, we are especially interested in studying the properties of fixed-point arithmetic and floating-point arithmetic that can help develop such a framework.

9.1.4. ANR ALAMBIC Project

Participants: Benoît Libert, Fabien Laguillaumie, Ida Tucker.

ALAMBIC is a four-year project (started in October 2016) focused on the applications of cryptographic primitives with homomorphic or malleability properties. The web page of the project is https://crypto.di.ens.fr/projects:alambic:description. It is headed by Damien Vergnaud (ENS Paris and CASCADE team) and, besides AriC, also involves teams from the XLIM laboratory (Université de Limoges) and the CASCADE team (ENS Paris). The main goals of the project are: (i) Leveraging the applications of malleable cryptographic primitives in the design of advanced cryptographic protocols which require computations on encrypted data; (ii) Enabling the secure delegation of expensive computations to remote servers in the cloud by using malleable cryptographic primitives; (iii) Designing more powerful zero-knowledge proof systems based on malleable cryptography.

9.1.5. RISQ Project

Participants: Chitchanok Chuengsatiansup, Fabien Laguillaumie, Benoît Libert, Damien Stehlé.

RISQ (Regroupement de l'Industrie française pour la Sécurité Post – Quantique) is a BPI-DGE four-year project (started in January 2017) focused on the transfer of post-quantum cryptography from academia to industrial poducts. The web page of the project is http://risq.fr. It is headed by Secure-IC and, besides AriC, also involves teams from ANSSI (Agence Nationale de la Sécurité des Systèmes d'Information), Airbus, C& S (Communication et Systèmes), CEA (CEA-List), CryptoExperts, Gemalto, Orange, Thales Communications & Security, Paris Center for Quantum Computing, the EMSEC team of IRISA, and the Cascade and Polsys Inria teams. The outcome of this project will include an exhaustive encryption and transaction signature product line, as well as an adaptation of the TLS protocol. Hardware and software cryptographic solutions meeting these constraints in terms of security and embedded integration will also be included. Furthermore, documents guiding industrials on the integration of these post-quantum technologies into complex systems (defense, cloud, identity and payment markets) will be produced, as well as reports on the activities of standardization committees.

9.2. European Initiatives

9.2.1. LattAC ERC grant

Participants: Shi Bai, Laurent Grémy, Gottfried Herold, Elena Kirshanova, Fabien Laguillaumie, Huyen Nguyen, Alice Pellet–Mary, Miruna Rosca, Damien Stehlé, Alexandre Wallet, Weiqiang Wen.

Damien Stehlé was awarded an ERC Starting Grant for his project *Euclidean lattices: algorithms and cryptography* (LattAC) in 2013 (1.4Meur for 5 years from January 2014). The LattAC project aims at studying all computational aspects of lattices, from algorithms for manipulating them to applications. The main objective is to enable the rise of lattice-based cryptography.

9.2.2. PROMETHEUS Project

Participants: Laurent Grémy, Fabien Laguillaumie, Benoît Libert, Damien Stehlé.

PROMETHEUS (Privacy-Preserving Systems from Advanced Cryptographic Mechanisms Using Lattices) is a 4-year European H2020 project (call H2020-DS-2016-2017, Cybersecurity PPP Cryptography, DS-06-2017) that started in January 2018. It gathers 8 academic partners (ENS de Lyon and Université de Rennes 1; CWI, Pays-Bas; IDC Herzliya, Israel; Royal Holloway University of London, United Kingdom; Universitat Politècnica de Catalunya, Spain; Ruhr-Universität Bochum, Germany; Weizmann Institute, Israel), 4 industrial partners (Orange, Thales, TNO, Scytl). The goal of this project is to develop a toolbox of privacy-preserving cryptographic algorithms and protocols (like group signatures, anonymous credentials, or digital cash systems) that resist quantum adversaries. Solutions will be mainly considered in the context of Euclidean lattices and they will be analyzed from a theoretical point of view (i.e., from a provable security aspect) and a practical angle (which covers the security of cryptographic implementations and side-channel leakages). The project is hosted by ENS de Lyon and Benoît Libert is the administrative coordinator while Orange is the scientific leader.

9.2.3. Other international projects

9.2.3.1. IFCPAR grant: "Computing on Encrypted Data: New Paradigms in Functional Encryption"

Participants: Benoît Libert, Damien Stehlé.

3-year project accepted in July 2018. Expected beginning on January 1, 2019. Benoît Libert is co-PI with Shweta Agrawal (IIT Madras, India). Budget on the French side amounts to 100k€.

Functional encryption is a paradigm that enables users to perform data mining and analysis on encrypted data. Users are provided cryptographic keys corresponding to particular functionalities which enable them to learn the output of the computation without learning anything about the input. Despite recent advances, efficient realizations of functional encryption are only available for restricted function families, which are typically represented by small-depth circuits: indeed, solutions for general functionalities are either way too inefficient for pratical use or they rely on uncertain security foundations like the existence of circuit obfuscators (or both). This project will explore constructions based on well-studied hardness assumptions and which are closer to being usable in real-life applications. To this end, we will notably consider solutions supporting other models of computation than Boolean circuits – like Turing machines – which support variable-size inputs. In the context of particular functionalities, the project will aim for more efficient realizations that satisfy stronger security notions.

9.3. International Initiatives

9.3.1. Participation in International Programs

Vincent Lefèvre actively participated in the revision of the IEEE Standard for Floating-Point Arithmetic (IEEE 754) for 2019.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Lloyd Nicholas Trefethen, from Oxford University (UK), is an expert in numerical analysis and notably the systematic use of Chebyshev approximation. He spent the academic year 2017-2018 with AriC.
- Warwick Tucker, from Uppsala University (Sweden), is an expert of certified computation for dynamical systems. He spent the academic year 2017-2018 with AriC.

9.4.2. Internships

Monosij Maitra, PhD student at IIT Madras (India) under the supervision of Shweta Agrawal, did a 2-month internship, in September and October 2018.

Joel Dahne did an internship with Bruno Salvy from May to July.

9.4.3. Visits to International Teams

- From November 15 to December 15, 2018, Benoît Libert visited the "Cryptography and Coding Research Group" of the Nanyang Technological University (Singapore).
- From July 1 to July 31, 2018, Damien Stehlé visited the cryptography group of Prof. Jung Hee Cheon, at Seoul National University (South Korea)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Member of Organizing Committees

- Claude-Pierre Jeannerod and Gilles Villard organized the workshop "Structured Matrix Days" (May 14–15, ENS de Lyon, France).
- Fabien Laguillaumie and Damien Stehlé organized the National Codes and Cryptography Days (Journées C2), in Aussois, France.
- Nathalie Revol co-organized the "École Jeunes Chercheurs et Jeunes Chercheuses en Programmation" (June 25–28, ENS de Lyon, France).
- Bruno Salvy is a co-chair of AofA'2019 (Analysis of Algorithms), in Luminy, France.

10.1.2. Scientific Events Selection

10.1.2.1. Member of Conference Program Committees

Chitchanok Chuengsatiansup was in the program committee of CRYPTO 2018.

Gottfried Herold was in the program committee of INDOCRYPT 2018.

Elena Kirshanova was in the program committee of INDOCRYPT 2018.

Benoît Libert was in the program committees of ACNS 2018, SCN 2018, Asiacrypt 2018, PKC 2019.

Jean-Michel Muller was in the program committee of Arith'25 and ASAP'2018.

Alain Passelègue was in the program committee of PKC 2018.

Nathalie Revol was in the program committee of Arith'25, of SCAN 2018 and of Correctness 2018.

Bruno Salvy was in the program committee for AofA'2018, is in the program committee of FPSAC 2019, in the steering committee of AofA and in the scientific committee of OPSFA 2019.

Damien Stehlé was in the program committees of Eurocrypt 2018, SCN 2018, PQCrypto 2018 and PQCrypto 2019. He is in the steering committee of the PQCrypto conference series.

Fabien Laguillaumie was in the program committee of ACISP 2018

10.1.3. Journals

Jean-Michel Muller is associate editor of the IEEE Transactions on Computers.

Nathalie Revol is a member of the editorial board of Reliable Computing.

Damien Stehlé is a member of the editorial board of the IACR Journal of Cryptology.

Bruno Salvy and Gilles Villard are members of the editorial board of Journal of Symbolic Computation.

Bruno Salvy is a member of the editorial board of the collection *Text and Monographs in Symbolic Computation* (Springer) and has been for 10 years in the editorial board of the *Journal of Algebra* (section Computational Algebra), which he left in March.

10.1.4. Invited Talks

- Claude-Pierre Jeannerod gave an invited talk *Recent results in fine-grained rounding error analysis* at the SCAN 2018 conference (Tokyo, September 10–15, 2018).
- Jean-Michel Muller gave an invited talk *Arithmétique et précision des calculs sur ordinateurs* at the conference *Tous mesureurs, tous mesurés*, organised by the INSHS and INP Institutes of CNRS, Paris, October 18-19, 2018.
- Benoît Libert gave an invited talk *New Applications of the Lossy Mode of LWE* at the *Chinacrypt 2018* conference, organised by the Chinese Association for Cryptologic Research (CACR) in Chengdu (China) on October 27-28, 2018.
- Damien Stehlé gave an invited talk *On algebraic variants of the LWE problem* at the ICERM workshop *Computational Challenges in the Theory of Lattices*, Providence (RI), on April 23-28, 2018. He also gave an invited talk on the same topic at the *Cryptography and Algorithmic Number Theory* workshop, held in Caen on June 20-22, 2018.

- Elena Kirshanova gave an invited talk *Sieving algorithms for the Shortest Vector Problem* at the *Joint Meeting of the Korean Mathematical Society and the German Mathematical Society*, held in Seoul, Korea, on October 3-6, 2018.
- Gottfried Herold gave an invited talk *Sieving in Practice* at the *Joint Meeting of the Korean Mathematical Society and the German Mathematical Society*, held in Seoul, Korea, on October 3-6, 2018.
- Jean-Michel Muller gave an invited talk *Make computer arithmetic great again* at a panel session on the future of computer arithmetic at Arith-25, 25-27 june 2018.
- Bruno Salvy gave an invited tutorial talk at STACS'2018 on random generation of combinatorial structures.

10.1.5. Leadership within the Scientific Community

Claude-Pierre Jeannerod was member of the scientific committee of JNCF (Journées Nationales de Calcul Formel). He was also a member of the recruitment committee for postdocs and sabbaticals at Inria Grenoble Rhône-Alpes.

Jean-Michel Muller is co-director of the *Groupement de Recherche Informatique Mathématique* (GDR IM) of CNRS; he chaired the HCERES evaluation committees of IRIF (UMR 8243, march 2018) and LIX (UMR 7161, october 2018); he is a member of the Scientific Concil of CERFACS; he participated to the jury of the *Prix La Recherche* award in 2018.

Alain Passelègue is a member of the steering committee of the *Groupe de Travail Codage et Cryptographie* (GT-C2) of the GDR-IM.

Bruno Salvy was a member of the HCERES evaluation committees of IRIF.

Damien Stehlé was a member of the jury for *prix de thèse SIF*.

10.1.6. Research Administration

Gilles Villard is a member of the *Section 6* of the *Comité national de la recherche scientifique*.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Claude-Pierre Jeannerod, Nathalie Revol, *Algorithmique numérique et fiabilité des calculs en arithmétique flottante* (24h), M2 ISFA (Institut de Science Financière et d'Assurances), Université Claude Bernard Lyon 1.

Master: Nicolas Brisebarre, Approximation Theory and Proof Assistants: Certified Computations, 18h, M2, ENS de Lyon, France

Master: Elena Kirshanova, Cryptanalysis, 18h, M2, ENS de Lyon, France

Master: Guillaume Hanrot, Cryptanalysis, 18h, M2, ENS de Lyon, France

Master: Damien Stehlé, Hard lattice problems, 36h, M2, ENS de Lyon, France

Post-graduate: Damien Stehlé, Hard lattice problems, 45h, Seoul National University, South Korea

Master: Elena Kirshanova, Computer Algebra, 10h, M1, ENS de Lyon, France

Master: Alexandre Wallet, Computer Algebra, 10h, M1, ENS de Lyon, France

Master: Guillaume Hanrot, Computer Algebra, 10h, M1, ENS de Lyon, France

Master: Bruno Salvy, Computer Algebra, 9h, MPRI, Paris, France

Master: Bruno Salvy, Logic and Complexity, 32h, École polytechnique, France

Master: Vincent Lefèvre, Computer arithmetic, 12h, M2 ISFA (Institut de Science Financière et d'Assurances), Université Claude Bernard Lyon 1.

Bachelor: Bruno Salvy, Design and Analysis of Algorithms, 15h, École polytechnique, France

Post-graduate: Bruno Salvy, Experimental Mathematics, 3h, Atelier jeunes chercheurs, St-Flour, France

Post-graduate: Bruno Salvy, Recent algorithms in symbolic summation and integration, 4h, Journées Louis Antoine, Rennes, France

Master: Fabien Laguillaumie, Cryptography, Security, Université Claude Bernard Lyon 1, 150h

Post-graduate : Fabien Laguillaumie, 2-party Computation and Homomorphic Encryption, 1h, École Cyber in Occitanie, France

10.2.2. Supervision

PhD: Fabrice Mouhartem, Privacy-preserving cryptography from pairings and lattices, ENS de Lyon (UdL), 18/10/2018, Benoît Libert

PhD in progress: Radu Titiu, Pseudo-random functions and functional encryption from lattices, ENS de Lyon (UdL), 01/01/2017, Benoît Libert

PhD in progress: Chen Qian, Additively homomorphic encryption and its applications, ENS de Lyon (UdL), 01/09/2016, Benoît Libert

PhD: Weiqiang Wen, Contributions to the hardness foundations of lattice-based cryptography, ENS de Lyon (UdL), 01/09/2015, Damien Stehlé

PhD in progress: Miruna Rosca, Algebraic variants of the LWE problem, ENS de Lyon (UdL), 01/01/2017, Damien Stehlé

PhD in progress: Alice Pellet–Mary, obfuscation cryptanalysis, ENS de Lyon (UdL), 01/09/2016, Damien Stehlé

PhD in progress: Huyen Nguyen, mathematical foundations of lattice-based cryptography, ENS de Lyon (UdL), 01/09/2018, Damien Stehlé

PhD in progress: Florent Bréhard, Outils pour un calcul numérique certifié -Applications aux systèmes dynamiques et à la théorie du contrôle, Ens de Lyon (UdL), 01/09/2016, Nicolas Brisebarre, Mioara Joldeş (CRNS, LAAS) et Damien Pous (CNRS, LIP, Plume)

PhD in progress: Adel Hamdi, Chiffrement fonctionnel pour le traitement de données externes en aveugle, UCBL (UdL) & Orange, 07/12/2017, Sébastien Canard (Orange), Fabien Laguillaumie

PhD in progress: Ida Tucker, Conception de systèmes cryptographiques avancés reposant sur des briques homomorphes, Ens de Lyon (UdL) et Université de Bordeaux, 17/10/2017, Guilhem Castagnos (IMB, Université de Bordeaux), Fabien Laguillaumie

10.2.3. Committees

Benoît Libert: reviewer for the PhD thesis of Pierre-Alain Dupont, ENS, 29/08/2018.

Damien Stehlé: reviewer for the PhD thesis of Thomas Ricosset, ENSEEIHT, 12/11/2018; reviewer for the PhD thesis of Ilaria Chillotti, UVSQ, 17/05/2018; examiner for the PhD thesis of Rachel Player, Royal Holloway University of London, 19/03/2018; president for the PhD thesis of Guillaume Bonnoron, Ecole nationale supérieure Mines-Télécom Atlantique Bretagne Pays de la Loire, 15/03/2018; jury member for the PhD thesis of Quentin Santos, ENS, 20/12/2018.

Bruno Salvy: member of the HdR committee of Guillaume Chapuy, IRIF, April and of Enrica Duchi, IRIF, November; reviewer for the PhD thesis of Pablo Rotondo, IRIF, September.

Fabien Laguillaumie: reviewer for the PhD thesis of Raphaël Bost, Université Rennes 1, 08/01/2018, Xavier Bultel, Université Clermont Auvergne, 17/05/2018, Vincent Zucca, Sorbonne Université, 25/06/2018, Quentin Santos, ENS, 20/12/2018

Nathalie Revol: examiner for the PhD thesis of Romain Picot, Université Paris 6, 27/03/2018

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- Nathalie Revol is a member of the editorial board of interstices; she belongs to the steering committee of MMI (Maison des Mathématiques et de l'Informatique, Lyon)
- Bruno Salvy is "référent chercheur" for the Inria Grenoble Center.

10.3.2. Articles and contents

Nathalie Revol belonged to the working group that elaborated the "7 families of computer science" playcards

10.3.3. Education

Nathalie Revol taught "Dissemination of Scientific Knowledge", 10h, to the 4th year students (between Master and PhD) of ENS de Lyon, France. She has been invited to a panel about "Flashmob" type activities, at ESOF 2018 (EuroScience Open Forum), July 9–14, 2018, Toulouse, France.

Nathalie Revol works with DANE (Délégation Académique au Numérique dans l'Éducation) of Rectorat de Lyon towards educating primary school teachers, by educating educators. She has been invited to present her past activities, using educational robots, at 3es Rencontres Nationales de la Robotique Éducative, October 2–3, Lyon, France.

10.3.4. Interventions

Laurent Grémy and Fabrice Mouhartem gave talks at *Fête de la Science* for a general audience. Nathalie Revol gave talks at *Fête de la Science* for 3 classes (9 years old, 11 years old and 13 years old).

As an incentive for high-school pupils, and especially girls, to choose scientific careers, Nathalie Revol gave talks at Lycée Ella Fitzgerald (Saint-Romain-en-Gal) and Mondial des Métiers (in February 2018). With Jérôme Germoni and Natacha Portier, she organized a day *Filles & Info* in March 2018, gathering about 100 high-school girls of 1e S. She was part of the panel discussing with the audience after the movie "Les figures de l'ombre - Hidden figures" at Comoedia cinema in Lyon in March 2018.

Damien Stehlé received at ENS de Lyon several winning teams of the Alkindi highschool competition. Alice Pellet–Mary and Fabrice Mouhartem gave talks at this event.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] F. MOUHARTEM.*Privacy-preserving cryptography from pairings and lattices*, Université de Lyon, October 2018, https://tel.archives-ouvertes.fr/tel-01913872

[2] W. WEN.*Contributions to the hardness foundations of lattice-based cryptography*, Université de Lyon, November 2018, https://tel.archives-ouvertes.fr/tel-01949339

Articles in International Peer-Reviewed Journal

[3] M. ABDALLA, F. BENHAMOUDA, A. PASSELÈGUE, K. PATERSON.*Related-Key Security for Pseudorandom Functions Beyond the Linear Barrier*, in "Journal of Cryptology", October 2018, vol. 31, n° 4, p. 917-964 [*DOI :* 10.1007/s00145-017-9274-8], https://hal.inria.fr/hal-01723012

[4] B. ALLOMBERT, N. BRISEBARRE, A. LASJAUNIAS.*On a two-valued sequence and related continued fractions in power series fields*, in "The Ramanujan Journal", 2018, vol. 45, n° 3, p. 859-871, https://arxiv.org/abs/1607.07235 [*DOI :* 10.1007/s11139-017-9892-7], https://hal.archives-ouvertes.fr/hal-01348576

[5] S. BAI, T. LEPOINT, A. ROUX-LANGLOIS, A. SAKZAD, D. STEHLÉ, R. STEINFELD.*Improved Security Proofs in Lattice-Based Cryptography: Using the Rényi Divergence Rather than the Statistical Distance*, in "Journal of Cryptology", April 2018, vol. 31, n^{o} 2, p. 610 - 640 [*DOI :* 10.1007/S00145-017-9265-9], https://hal.archives-ouvertes.fr/hal-01934177

[6] N. BRISEBARRE, S.-I. FILIP, G. HANROT.*A Lattice Basis Reduction Approach for the Design of Finite Wordlength FIR Filters*, in "IEEE Transactions on Signal Processing", 2018, vol. 66, n^{o} 10, p. 2673-2684 [*DOI :* 10.1109/TSP.2018.2812739], https://hal.inria.fr/hal-01308801

[7] F. BRÉHARD, N. BRISEBARRE, M. JOLDES.*Validated and numerically efficient Chebyshev spectral methods for linear ordinary differential equations*, in "ACM Transactions on Mathematical Software", July 2018, vol. 44, n^{o} 4, p. 44:1-44:42 [*DOI :* 10.1145/3208103], https://hal.archives-ouvertes.fr/hal-01526272

[8] L. DUCAS, E. KILTZ, T. LEPOINT, V. LYUBASHEVSKY, P. SCHWABE, G. SEILER, D. STEHLÉ.*CRYSTALS-Dilithium: A Lattice-Based Digital Signature Scheme*, in "IACR Transactions on Cryptographic Hardware and Embedded Systems", 2018, https://hal.archives-ouvertes.fr/hal-01934176

[9] G. HEROLD, E. KIRSHANOVA, A. MAY.*On the asymptotic complexity of solving LWE*, in "Designs, Codes and Cryptography", January 2018, vol. 86, n^{o} 1, p. 55 - 83 [*DOI :* 10.1007/S10623-016-0326-0], https://hal.archives-ouvertes.fr/hal-01934181

[10] C.-P. JEANNEROD, S. M. RUMP.*On relative errors of floating-point operations: optimal bounds and applications*, in "Mathematics of Computation", 2018, vol. 87, p. 803-819 [*DOI :* 10.1090/MCOM/3234], https://hal.inria.fr/hal-00934443

[11] B. SALVY.*Linear Differential Equations as a Data-Structure*, in "Foundations of Computational Mathematics", 2018, p. 1-35, https://arxiv.org/abs/1811.08616 - Based on an invited talk at FoCM'2017, https://hal.inria.fr/hal-01940078

[12] A. VOLKOVA, M. ISTOAN, F. DE DINECHIN, T. HILAIRE.*Towards Hardware IIR Filters Computing Just Right: Direct Form I Case Study*, in "IEEE Transactions on Computers", 2018 [*DOI :* 10.1109/TC.2018.2879432], https://hal.sorbonne-universite.fr/hal-01561052

Invited Conferences

[13] B. SALVY.*Recursive Combinatorial Structures: Enumeration, Probabilistic Analysis and Random Generation*, in "STACS 2018 - 35th Symposium on Theoretical Aspects of Computer Science", Caen, France, February 2018, Tutorial [*DOI :* 10.4230/LIPICS.STACS.2018.1], https://hal.inria.fr/hal-01926094

International Conferences with Proceedings

[14] P. R. ARANTES GILZ, F. BRÉHARD, C. GAZZINO.*Validated Semi-Analytical Transition Matrices for Linearized Relative Spacecraft Dynamics via Chebyshev Series Approximations*, in "SCITECH 2018 - AIAA Science and Technology Forum and Exposition, 28th Space Flight Mechanics Meeting", Kissimmee, United States, SCITECH 2018-AIAA Science and Technology Forum and Exposition, 28th Space Flight Mechanics Meeting, American Institute of Aeronautics and Astronautics, January 2018, p. 1-23 [*DOI :* 10.2514/6.2018-1960], https://hal.archives-ouvertes.fr/hal-01540170

[15] S. BAI, D. STEHLÉ, W. WEN.*Measuring, Simulating and Exploiting the Head Concavity Phenomenon in BKZ*, in "ASIACRYPT", Brisbane, Australia, 2018, https://hal.archives-ouvertes.fr/hal-01934174

[16] D. BONEH, Y. ISHAI, A. PASSELÈGUE, A. SAHAI, D. J. WU.*Exploring Crypto Dark Matter: New Simple PRF Candidates and Their Applications*, in "TCC 2018 - Theory of Cryptography Conference", Goa, India, LNCS, Springer, November 2018, vol. 11240, p. 699-729 [*DOI :* 10.1007/978-3-030-03810-6_25], https://hal.inria.fr/hal-01929288

[17] J. W. BOS, L. DUCAS, E. KILTZ, T. LEPOINT, V. LYUBASHEVSKY, J. SCHANCK, P. SCHWABE, G. SEILER, D. STEHLÉ.*CRYSTALS - Kyber: A CCA-Secure Module-Lattice-Based KEM*, in "EuroS&P 2018 - IEEE European Symposium on Security and Privacy", London, United Kingdom, IEEE, April 2018, p. 353-367 [*DOI :* 10.1109/EUROSP.2018.00032], https://hal.archives-ouvertes.fr/hal-01934169

[18] A. BOSTAN, F. CHYZAK, P. LAIREZ, B. SALVY.*Generalized Hermite Reduction, Creative Telescoping and Definite Integration of D-Finite Functions*, in "ISSAC 2018 - International Symposium on Symbolic and Algebraic Computation", New York, United States, July 2018, p. 1-8 [*DOI :* 10.1145/3208976.3208992], https://hal.inria.fr/hal-01788619

[19] Z. BRAKERSKI, A. JAIN, I. KOMARGODSKI, A. PASSELÈGUE, D. WICHS.*Non-Trivial Witness Encryption and Null-iO from Standard Assumptions*, in "SCN 2018 - International Conference on Security and Cryptography for Networks", Amalfi, Italy, LNCS, Springer, September 2018, vol. 11035, p. 425-441 [*DOI :* 10.1007/978-3-319-98113-0_23], https://hal.inria.fr/hal-01929279

[20] Z. BRAKERSKI, E. KIRSHANOVA, D. STEHLÉ, W. WEN.*Learning with Errors and Extrapolated Dihedral Cosets*, in "PKC 2018 - 21st International Conference on Practice and Theory of Public Key Cryptography", Rio de Janeiro, Brazil, March 2018, https://hal.archives-ouvertes.fr/hal-01934165

[21] N. BRISEBARRE, G. CONSTANTINIDES, M. ERCEGOVAC, S.-I. FILIP, M. ISTOAN, J.-M. MULLER.*A High Throughput Polynomial and Rational Function Approximations Evaluator*, in "ARITH 2018 - 25th IEEE Symposium on Computer Arithmetic", Amherst, MA, United States, IEEE, June 2018, p. 99-106 [*DOI :* 10.1109/ARITH.2018.8464778], https://hal.inria.fr/hal-01774364

[22] F. BRÉHARD.*A Newton-like Validation Method for Chebyshev Approximate Solutions of Linear Ordinary Differential Systems*, in "ISSAC 2018 - 43rd International Symposium on Symbolic and Algebraic Computation", New York, United States, ISSAC 2018-43rd International Symposium on Symbolic and Algebraic Computation, ACM, July 2018, p. 103-110 [*DOI :* 10.1145/3208976.3209000], https://hal.archives-ouvertes.fr/hal-01654396

[23] G. CASTAGNOS, F. LAGUILLAUMIE, I. TUCKER.*Practical Fully Secure Unrestricted Inner Product Functional Encryption modulo p*, in "ASIACRYPT 2018 - 24th International Conference on the Theory and Application of Cryptology and Information Security", Brisbane, Australia, T. PEYRIN, S. GALBRAITH (editors), Advances in Cryptology – ASIACRYPT 2018, December 2018, vol. LNCS, n^{o} 11273, p. 733-764, https://hal.archives-ouvertes.fr/hal-01934296

[24] J. CHEN, J. GONG, L. KOWALCZYK, H. WEE.*Unbounded ABE via Bilinear Entropy Expansion, Revisited*, in "EUROCRYPT 2018 - Annual International Conference on the Theory and Applications of Cryptographic Techniques", Tel Aviv, Israel, J. B. NIELSEN, V. RIJMEN (editors), LNCS - Lecture Notes in Computer Science, Springer, April 2018, vol. 10820, p. 503-534 [*DOI :* 10.1007/978-3-319-78381-9_19], https://hal.inria.fr/hal-01899901

[25] J. CHEN, J. GONG, H. WEE.*Improved Inner-product Encryption with Adaptive Security and Full Attribute-hiding*, in "ASIACRYPT 2018 - 24th Annual International Conference on the Theory and Application of Cryptology and Information Security", Brisbane, Australia, December 2018, https://hal.inria.fr/hal-01900153

[26] J. CHEN, D. STEHLÉ, G. VILLARD.*Computing an LLL-reduced Basis of the Orthogonal Lattice*, in "ISSAC 2018, 43rd International Symposium on Symbolic and Algebraic Computation (ISSAC 2018)", New York, United States, July 2018, https://arxiv.org/abs/1805.03418 [*DOI :* 10.1145/3208976.3209013], https://hal.archives-ouvertes.fr/hal-01921335

[27] C. CHUENGSATIANSUP, C. MARTINDALE.*Pairing-friendly twisted Hessian curves*, in "Indocrypt 2018 - 19th International Conference on Cryptology", New Delhi, India, December 2018, https://hal.archives-ouvertes.fr/hal-01934160

[28] L. DUCAS, A. PELLET–MARY.*On the Statistical Leak of the GGH13 Multilinear Map and some Variants*, in "Asiacrypt 2018", Brisbane, Australia, December 2018, p. 465-493, https://hal.archives-ouvertes.fr/hal-01895645

[29] J. GONG, B. LIBERT, S. C. RAMANNA.*Compact IBBE and Fuzzy IBE from Simple Assumptions*, in "SCN 2018 - 11th Conference on Security and Cryptography for Networks", Amalfi, Italy, Security and Cryptography for Networks (SCN) 2018, September 2018, p. 1-29, https://hal.inria.fr/hal-01686690

[30] L. GRÉMY.*Higher dimensional sieving for the number field sieve algorithms*, in "ANTS 2018 - Thirteenth Algorithmic Number Theory Symposium", Madison, United States, University of Wisconsin, July 2018, p. 1-16, https://hal.inria.fr/hal-01890731

[31] G. HEROLD, E. KIRSHANOVA, T. LAARHOVEN.*Speed-Ups and Time-Memory Trade-Offs for Tuple Lattice Sieving*, in "PKC 2018 - 21st International Conference on Practice and Theory of Public Key Cryptography", Rio de Janeiro, Brazil, March 2018, https://hal.archives-ouvertes.fr/hal-01934183

[32] C.-P. JEANNEROD, J.-M. MULLER, P. ZIMMERMANN.*On various ways to split a floating-point number*, in "ARITH 2018 - 25th IEEE Symposium on Computer Arithmetic", Amherst (MA), United States, IEEE, June 2018, p. 53-60 [*DOI :* 10.1109/ARITH.2018.8464793], https://hal.inria.fr/hal-01774587

[33] M. JOYE, A. PASSELÈGUE.*Function-Revealing Encryption: Definitions and Constructions*, in "SCN 2018 - International Conference on Security and Cryptography for Networks", Amalfi, Italy, LNCS, Springer, September 2018, vol. 11035, p. 527-543 [*DOI :* 10.1007/978-3-319-98113-0_28], https://hal.inria.fr/hal-01929272

[34] E. KIRSHANOVA.*Improved Quantum Information Set Decoding*, in "PQCrypto 2018 - The Ninth International Conference on Post-Quantum Cryptography", Fort Lauderdale, United States, April 2018, https://hal.archives-ouvertes.fr/hal-01934186

[35] J. LI, J. GONG.*Improved Anonymous Broadcast Encryptions: Tight Security and Shorter Ciphertext*, in "ACNS 2018 - 16th International Conference on Applied Cryptography and Network Security", Leuven, Belgium, Springer, July 2018, p. 497-515 [*DOI :* 10.1007/978-3-319-93387-0_26], https://hal.archives-ouvertes.fr/hal-01829132

[36] B. LIBERT, S. LING, K. NGUYEN, H. WANG.*Lattice-Based Zero-Knowledge Arguments for Integer Relations*, in "CRYPTO 2018 - Annual International Cryptology Conference", Santa Barbara, United States,

Springer, August 2018, vol. LNCS, n^o 10992, p. 700-732 [*DOI :* 10.1007/978-3-319-96881-0_24], https://hal.inria.fr/hal-01911886

[37] B. LIBERT, T. PETERS, C. QIAN.*Logarithmic-Size Ring Signatures With Tight Security from the DDH Assumption*, in "ESORICS 2018 - 23rd European Symposium on Research in Computer Security", Barcelone, Spain, LNCS, Springer, September 2018, vol. 11099, p. 288-308 [*DOI :* 10.1007/978-3-319-98989-1_15], https://hal.inria.fr/hal-01848134

[38] B. LIBERT, D. STEHLÉ, R. TITIU.*Adaptively Secure Distributed PRFs from LWE*, in "TCC 2018 - 16th International Conference on Theory of Cryptography", Panaji, India, LNCS, Springer, November 2018, vol. 11240, p. 391-421 [*DOI :* 10.1007/978-3-030-03810-6_15], https://hal.inria.fr/hal-01911887

[39] A. PELLET–MARY.*Quantum Attacks against Indistinguishablility Obfuscators Proved Secure in the Weak Multilinear Map Model*, in "Crypto 2018 - 38th International Cryptology Conference", Santa-Barbara, United States, Springer, August 2018, p. 153-183 [*DOI :* 10.1007/978-3-319-96878-0_6], https://hal.archives-ouvertes.fr/hal-01895639

[40] M. ROCA, D. STEHLÉ, A. WALLET.*On the Ring-LWE and Polynomial-LWE Problems*, in "EUROCRYPT 2018 - 37th Annual International Conference on the Theory and Applications", Tel Aviv, Israel, April 2018, https://hal.archives-ouvertes.fr/hal-01934170

[41] R. SERRA, D. ARZELIER, F. BRÉHARD, M. JOLDES.*Fuel-optimal impulsive fixed-time trajectories in the linearized circular restricted 3-body-problem*, in "IAC 2018 - 69th International Astronautical Congress; IAF Astrodynamics Symposium", Bremen, Germany, CSA/IAF Special issue IAF Astrodynamics Symposium (69TH international astronautical congress), International Astronautical Federation, October 2018, p. 1-9, https://hal.archives-ouvertes.fr/hal-01830253

[42] *Best Paper*
G. VILLARD.*On Computing the Resultant of Generic Bivariate Polynomials*, in "ISSAC 2018, 43rd International Symposium on Symbolic and Algebraic Computation, New York, USA, July 16-19, 2018", New York, United States, July 2018, https://hal.archives-ouvertes.fr/hal-01921369.

Scientific Books (or Scientific Book chapters)

[43] J.-M. MULLER, N. BRUNIE, F. DE DINECHIN, C.-P. JEANNEROD, M. JOLDES, V. LEFÈVRE, G. MELQUIOND, N. REVOL, S. TORRES.*Handbook of Floating-point Arithmetic (2nd edition)*, Birkhäuser Basel, July 2018, p. 1-627 [*DOI :* 10.1007/978-3-319-76526-6], https://hal.inria.fr/hal-01766584

Other Publications

[44] N. BRISEBARRE, M. JOLDES, J.-M. MULLER, A.-M. NANEŞ, J. PICOT.*Error analysis of some operations involved in the Fast Fourier Transform*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01949458

[45] N. FABIANO, J.-M. MULLER.*Algorithms for triple-word arithmetic*, September 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01869009

[46] C.-P. JEANNEROD, V. NEIGER, G. VILLARD.*Fast computation of approximant bases in canonical form*, January 2018, working paper or preprint, https://hal-unilim.archives-ouvertes.fr/hal-01683632

[47] F. QURESHI, A. VOLKOVA, T. HILAIRE, J. TAKALA.*Multiplierless Processing Element for Non-Power-of-Two FFTs*, January 2018, working paper or preprint, https://hal.inria.fr/hal-01690832

[48] A. VOLKOVA, T. HILAIRE, C. LAUTER.*Arithmetic approaches for rigorous design of reliable Fixed-Point LTI filters*, November 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01918650

[49] F. DE DINECHIN, L. FORGET, J.-M. MULLER, Y. UGUEN.*Posits: the good, the bad and the ugly*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01959581

Project-Team AVALON

Algorithms and Software Architectures for Distributed and HPC Platforms

IN COLLABORATION WITH: Laboratoire de l'Informatique du Parallélisme (LIP)

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Lyon

Université Claude Bernard (Lyon 1)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Distributed and High Performance Computing

Table of contents

Project-Team AVALON

Creation of the Team: 2012 February 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A1.1.1. - Multicore, Manycore
A1.1.4. - High performance computing
A1.1.5. - Exascale
A1.1.13. - Virtualization
A1.2.1. - Dynamic reconfiguration
A1.3. - Distributed Systems
A1.3.4. - Peer to peer
A1.3.5. - Cloud
A1.3.6. - Fog, Edge
A1.6. - Green Computing
A2.1.6. - Concurrent programming
A2.1.7. - Distributed programming
A2.1.10. - Domain-specific languages
A2.5.2. - Component-based Design
A2.6.2. - Middleware
A2.6.4. - Ressource management
A3.1.2. - Data management, quering and storage
A3.1.3. - Distributed data
A3.1.8. - Big data (production, storage, transfer)
A4.4. - Security of equipment and software
A6.2.7. - High performance computing
A7.1. - Algorithms
A7.1.1. - Distributed algorithms
A7.1.2. - Parallel algorithms
A8.2. - Optimization
A8.2.1. - Operations research
A8.2.2. - Evolutionary algorithms
A8.9. - Performance evaluation

Other Research Topics and Application Domains:

B1.1.7. - Bioinformatics
B3.2. - Climate and meteorology
B4.1. - Fossile energy production (oil, gas)
B4.2.2. - Fusion
B4.5. - Energy consumption
B4.5.1. - Green computing
B6.1.1. - Software engineering
B8.1.1. - Energy for smart buildings

B9.5.1. - Computer science
B9.7. - Knowledge dissemination
B9.7.1. - Open access
B9.7.2. - Open data
B9.8. - Reproducibility

1. Team, Visitors, External Collaborators

Research Scientists
Christian Perez [Team leader, Inria, Senior Researcher, HDR]
Marcos Dias de Assunção [Inria, Starting Research Position]
Thierry Gautier [Inria, Researcher]
Laurent Lefèvre [Inria, Researcher, HDR]
Frédéric Suter [CNRS, Researcher, HDR]

Faculty Members
Yves Caniou [Univ. Claude Bernard, Associate Professor]
Eddy Caron [École Normale Supérieure de Lyon, Associate Professor, HDR]
Jean-Patrick Gelas [Univ. Claude Bernard, Associate Professor]
Olivier Glück [Univ. Claude Bernard, Associate Professor]

Technical Staff
Mathilde Boutigny [Inria, Engineer, until Mar 2018, granted by BPIFRANCE FINANCEMENT SA]
Simon Delamare [CNRS, Engineer, 50%]
David Loup [Inria, Engineer]
Jean-Christophe Mignot [CNRS, Engineer, 20%]
Matthieu Imbert [Inria, Engineer, 40%]

PhD Students
Dorra Boughzala [Inria]
Maverick Chardet [Inria]
Arthur Chevalier [Orange Labs]
Hadrien Croubois [École Normale Supérieure de Lyon, until Aug 2018]
Idriss Daoudi [Inria, from Oct 2018]
Alexandre Da Silva Veith [Univ de Lyon]
Jad Darrous [Inria]
Felipe Rodrigo de Souza [École Normale Supérieure de Lyon]
Zeina Houmani [École Normale Supérieure de Lyon, from Nov 2018]
Aurélie Kong-Win-Chang [Univ. Claude Bernard]
Valentin Lorentz [École Normale Supérieure de Lyon, until Sep 2018]
Issam Raïs [Inria, until Sep 2018]
Philippe Virouleau [Inria, until Feb 2018]

Post-Doctoral Fellows
Jérôme Richard [Inria, until Mar 2018]
Cyril Seguin [Inria, from Apr 2018]

Administrative Assistants
Solene Audoux [Inria, Assistant, from Nov 2018]
Evelyne Blesle [Inria, Assistant]

2. Overall Objectives

2.1. Presentation

The fast evolution of hardware capabilities in terms of wide area communication, computation and machine virtualization leads to the requirement of another step in the abstraction of resources with respect to parallel and distributed applications. These large scale platforms based on the aggregation of large clusters (Grids), huge datacenters (Clouds) with IoT (Edge/Fog), collections of volunteer PCs (Desktop computing platforms), or high performance machines (Supercomputers) are now available to researchers of different fields of science as well as to private companies. This variety of platforms and the way they are accessed also have an important impact on how applications are designed (*i.e.,* the programming model used) as well as how applications are executed (*i.e.,* the runtime/middleware system used). The access to these platforms is driven through the use of multiple services providing mandatory features such as security, resource discovery, virtualization, load-balancing, monitoring, *etc.*

The goal of the AVALON team is to execute parallel and/or distributed applications on parallel and/or distributed resources while ensuring user and system objectives with respect to performance, cost, energy, security, *etc.* Users are generally not interested in the resources used during the execution. Instead, they are interested in how their application is going to be executed: the duration, its cost, the environmental footprint involved, *etc.* This vision of utility computing has been strengthened by the cloud concepts and by the short lifespan of supercomputers (around three years) compared to application lifespan (tens of years). Therefore a major issue is to design models, systems, and algorithms to execute applications on resources while ensuring user constraints (price, performance, *etc.*) as well as system administrator constraints (maximizing resource usage, minimizing energy consumption, *etc.*).

2.2. Objectives

To achieve the vision proposed in Section 2.1, the AVALON project aims at making progress to four complementary research axes: energy, data, component models, and application scheduling.

2.2.1. Energy Application Profiling and Modeling

AVALON will improve the profiling and modeling of scientific applications with respect to energy consumption. In particular, it will require to improve the tools that measure the energy consumption of applications, virtualized or not, at large scale, so as to build energy consumption models of applications.

2.2.2. Data-intensive Application Profiling, Modeling, and Management

AVALON will improve the profiling, modeling, and management of scientific applications with respect to CPU and data intensive applications. Challenges are to improve the performance prediction of parallel regular applications, to model and simulate (complex) intermediate storage components, and data-intensive applications, and last to deal with data management for hybrid computing infrastructures.

2.2.3. Resource-Agnostic Application Description Model

AVALON will design component-based models to capture the different facets of parallel and distributed applications while being resource agnostic, so that they can be optimized for a particular execution. In particular, the proposed component models will integrate energy and data modeling results. AVALON in particular targets OpenMP runtime as a specific use case.

2.2.4. Application Mapping and Scheduling

AVALON will propose multi-criteria mapping and scheduling algorithms to meet the challenge of automating the efficient utilization of resources taking into consideration criteria such as performance (CPU, network, and storage), energy consumption, and security. AVALON will in particular focus on application deployment, workflow applications, and security management in clouds.

All our theoretical results will be validated with software prototypes using applications from different fields of science such as bioinformatics, physics, cosmology, *etc.* The experimental testbeds GRID'5000 (cf Section 6.8) and Silecs (cf Section 6.10 will be our platforms of choice for experiments.

3. Research Program

3.1. Energy Application Profiling and Modeling

Despite recent improvements, there is still a long road to follow in order to obtain energy efficient, energy proportional and eco-responsible exascale systems by 2022. Energy efficiency is therefore a major challenge for building next generation large-scale platforms. The targeted platforms will gather hundreds of millions of cores, low power servers, or CPUs. Besides being very important, their power consumption will be dynamic and irregular.

Thus, to consume energy efficiently, we aim at investigating two research directions. First, we need to improve measurement, understanding, and analysis on how large-scale platforms consume energy. Unlike approaches [36] that mix the usage of internal and external wattmeters on a small set of resources, we target high frequency and precise internal and external energy measurements of each physical and virtual resource on large-scale distributed systems.

Secondly, we need to find new mechanisms that consume less and better on such platforms. Combined with hardware optimizations, several works based on shutdown or slowdown approaches aim at reducing energy consumption of distributed platforms and applications. To consume less, we first plan to explore the provision of accurate estimation of the energy consumed by applications without pre-executing and knowing them while most of the works try to do it based on in-depth application knowledge (code instrumentation [39], phase detection for specific HPC applications [44], *etc.*). As a second step, we aim at designing a framework model that allows interaction, dialogue and decisions taken in cooperation among the user/application, the administrator, the resource manager, and the energy supplier. While smart grid is one of the last killer scenarios for networks, electrical provisioning of next generation large IT infrastructures remains a challenge.

3.2. Data-intensive Application Profiling, Modeling, and Management

Recently, the term "Big Data" has emerged to design data sets or collections so large that they become intractable for classical tools. This term is most time implicitly linked to "analytics" to refer to issues such as data curation, storage, search, sharing, analysis, and visualization. However, the Big Data challenge is not limited to data-analytics, a field that is well covered by programming languages and run-time systems such as Map-Reduce. It also encompasses data-intensive applications. These applications can be sorted into two categories. In High Performance Computing (HPC), data-intensive applications leverage post-petascale infrastructures to perform highly parallel computations on large amount of data, while in High Throughput Computing (HTC), a large amount of independent and sequential computations are performed on huge data collections.

These two types of data-intensive applications (HTC and HPC) raise challenges related to profiling and modeling that the AVALON team proposes to address. While the characteristics of data-intensive applications are very different, our work will remain coherent and focused. Indeed, a common goal will be to acquire a better understanding of both the applications and the underlying infrastructures running them to propose the best match between application requirements and infrastructure capacities. To achieve this objective, we will extensively rely on logging and profiling in order to design sound, accurate, and validated models. Then, the proposed models will be integrated and consolidated within a single simulation framework (SIMGRID). This will allow us to explore various potential "what-if?" scenarios and offer objective indicators to select interesting infrastructure configurations that match application specificities.

Another challenge is the ability to mix several heterogeneous infrastructures that scientists have at their disposal (*e.g.,* Grids, Clouds, and Desktop Grids) to execute data-intensive applications. Leveraging the aforementioned results, we will design strategies for efficient data management service for hybrid computing infrastructures.

3.3. Resource-Agnostic Application Description Model

With parallel programming, users expect to obtain performance improvement, regardless its cost. For long, parallel machines have been simple enough to let a user program use them given a minimal abstraction of their hardware. For example, MPI [38] exposes the number of nodes but hides the complexity of network topology behind a set of collective operations; OpenMP [42] simplifies the management of threads on top of a shared memory machine while OpenACC [41] aims at simplifying the use of GPGPU.

However, machines and applications are getting more and more complex so that the cost of manually handling an application is becoming very high [37]. Hardware complexity also stems from the unclear path towards next generations of hardware coming from the frequency wall: multi-core CPU, many-core CPU, GPGPUs, deep memory hierarchy, *etc.* have a strong impact on parallel algorithms. Hence, even though an abstract enough parallel language (UPC, Fortress, X10, *etc.*) succeeds, it will still face the challenge of supporting distinct codes corresponding to different algorithms corresponding to distinct hardware capacities.

Therefore, the challenge we aim to address is to define a model, for describing the structure of parallel and distributed applications that enables code variations but also efficient executions on parallel and distributed infrastructures. Indeed, this issue appears for HPC applications but also for cloud oriented applications. The challenge is to adapt an application to user constraints such as performance, energy, security, *etc.*

Our approach is to consider component based models [45] as they offer the ability to manipulate the software architecture of an application. To achieve our goal, we consider a “compilation” approach that transforms a resource-agnostic application description into a resource-specific description. The challenge is thus to determine a component based model that enables to efficiently compute application mapping while being tractable. In particular, it has to provide an efficient support with respect to application and resource elasticity, energy consumption and data management. OpenMP runtime is a specific use case that we target.

3.4. Application Mapping and Scheduling

This research axis is at the crossroad of the AVALON team. In particular, it gathers results of the three other research axis. We plan to consider application mapping and scheduling addressing the following three issues.

3.4.1. Application Mapping and Software Deployment

Application mapping and software deployment consist in the process of assigning distributed pieces of software to a set of resources. Resources can be selected according to different criteria such as performance, cost, energy consumption, security management, *etc.* A first issue is to select resources at application launch time. With the wide adoption of elastic platforms, *i.e.,* platforms that let the number of resources allocated to an application to be increased or decreased during its execution, the issue is also to handle resource selection at runtime.

The challenge in this context corresponds to the mapping of applications onto distributed resources. It will consist in designing algorithms that in particular take into consideration application profiling, modeling, and description.

A particular facet of this challenge is to propose scheduling algorithms for dynamic and elastic platforms. As the number of elements can vary, some kind of control of the platforms must be used accordingly to the scheduling.

3.4.2. Non-Deterministic Workflow Scheduling

Many scientific applications are described through workflow structures. Due to the increasing level of parallelism offered by modern computing infrastructures, workflow applications now have to be composed not only of sequential programs, but also of parallel ones. New applications are now built upon workflows with conditionals and loops (also called non-deterministic workflows).

These workflows cannot be scheduled beforehand. Moreover cloud platforms bring on-demand resource provisioning and pay-as-you-go billing models. Therefore, there is a problem of resource allocation for non-deterministic workflows under budget constraints and using such an elastic management of resources.

Another important issue is data management. We need to schedule the data movements and replications while taking job scheduling into account. If possible, data management and job scheduling should be done at the same time in a closely coupled interaction.

3.4.3. Security Management in Cloud Infrastructure

Security has been proven to be sometimes difficult to obtain [43] and several issues have been raised in Clouds. Nowadays virtualization is used as the sole mechanism to allow multiple users to share resources on Clouds, but since not all components of Clouds (such as micro-architectural components) can be properly virtualized, data leak and modification can occur. Accordingly, next-generation protection mechanisms are required to enforce security on Clouds and provide a way to cope with the current limitation of virtualization mechanisms.

As we are dealing with parallel and distributed applications, security mechanisms must be able to cope with multiple machines. Our approach is to combine a set of existing and novel security mechanisms that are spread in the different layers and components of Clouds in order to provide an in-depth and end-to-end security on Clouds. To do it, our first challenge is to define a generic model to express security policies.

Our second challenge is to work on security-aware resource allocation algorithms. The goal of such algorithms is to find a good trade-off between security and unshared resources. Consequently, they can limit resources sharing to increase security. It leads to complex trade-off between infrastructure consolidation, performance, and security.

4. Application Domains

4.1. Overview

The AVALON team targets applications with large computing and/or data storage needs, which are still difficult to program, maintain, and deploy. Those applications can be parallel and/or distributed applications, such as large scale simulation applications or code coupling applications. Applications can also be workflow-based as commonly found in distributed systems such as grids or clouds.

The team aims at not being restricted to a particular application field, thus avoiding any spotlight. The team targets different HPC and distributed application fields, which brings use cases with different issues. This will be eased by our various collaborations: the team participates to the INRIA-Illinois Joint Laboratory for Petascale Computing, the Physics, Radiobiology, Medical Imaging, and Simulation French laboratory of excellence, the E-Biothon project, the INRIA large scale initiative Computer and Computational Sciences at Exascale (C2S@Exa), and to BioSyL, a federative research structure about Systems Biology of the University of Lyon. Moreover, the team members have a long tradition of cooperation with application developers such as CERFACS and EDF R&D. Last but not least, the team has a privileged connection with CC IN2P3 that opens up collaborations, in particular in the astrophysics field.

In the following, some examples of representative applications we are targeting are presented. In addition to highlighting some application needs, they also constitute some of the use cases we will use to valide our theoretical results.

4.2. Climatology

The world's climate is currently changing due to the increase of the greenhouse gases in the atmosphere. Climate fluctuations are forecasted for the years to come. For a proper study of the incoming changes, numerical simulations are needed, using general circulation models of a climate system. Simulations can be of different types: HPC applications (*e.g.,* the NEMO framework [40] for ocean modelization), code-coupling applications (*e.g.,* the OASIS coupler [46] for global climate modeling), or workflows (long term global climate modeling).

As for most applications the team is targeting, the challenge is to thoroughly analyze climate-forecasting applications to model their needs in terms of programing model, execution model, energy consumption, data access pattern, and computing needs. Once a proper model of an application has been set up, appropriate scheduling heuristics could be designed, tested, and compared. The team has a long tradition of working with CERFACS on this topic, for example in the LEGO (2006-09) and SPADES (2009-12) French ANR projects.

4.3. Astrophysics

Astrophysics is a major field to produce large volumes of data. For instance, the Large Synoptic Survey Telescope (https://www.lsst.org/lsst/) will produce 15 TB of data every night, with the goals of discovering thousands of exoplanets and of uncovering the nature of dark matter and dark energy in the universe. The Square Kilometer Array (http://www.skatelescope.org/) produces 9 Tbits/s of raw data. One of the scientific projects related to this instrument called Evolutionary Map of the Universe is working on more than 100 TB of images. The Euclid Imaging Consortium (https://www.euclid-ec.org/) will generate 1 PB data per year.

AVALON collaborates with the *Institut de Physique Nucléaire de Lyon* (IPNL) laboratory on large scale numerical simulations in astronomy and astrophysics. Contributions of the AVALON members have been related to algorithmic skeletons to demonstrate large scale connectivity, the development of procedures for the generation of realistic mock catalogs, and the development of a web interface to launch large cosmological simulations on GRID'5000.

This collaboration, that continues around the topics addressed by the CLUES project (http://www.clues-project.org), has been extended thanks to the tight links with the CC-IN2P3. Major astrophysics projects execute part of their computing, and store part of their data on the resources provided by the CC-IN2P3. Among them, we can mention SNFactory, Euclid, or LSST. These applications constitute typical use cases for the research developed in the AVALON team: they are generally structured as workflows and a huge amount of data (from TB to PB) is involved.

4.4. Bioinformatics

Large-scale data management is certainly one of the most important applications of distributed systems in the future. Bioinformatics is a field producing such kinds of applications. For example, DNA sequencing applications make use of MapReduce skeletons.

The AVALON team is a member of BioSyL (http://www.biosyl.org), a Federative Research Structure attached to University of Lyon. It gathers about 50 local research teams working on systems biology. Moreover, the team cooperated with the French Institute of Biology and Chemistry of Proteins (IBCP http://www.ibcp.fr) in particular through the ANR MapReduce project where the team focuses on a bio-chemistry application dealing with protein structure analysis. AVALON has also started working with the Inria Beagle team (https://team.inria.fr/beagle/) on artificial evolution and computational biology as the challenges are around high performance computation and data management.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Henri Casanova, Arnaud Legrand, Martin Quinson and Frédéric Suter. "SMPI Courseware: Teaching Distributed-Memory Computing with MPI in Simulation" received the "Best Paper Award" of the Workshop on Education for High-Performance Computing (EduHPC-18).
- Anchen Chai, Sorina Camarasu-Pop, Tristan Glatard, Hugues Benoit-Cattin and Frédéric Suter. "Evaluation through Realistic Simulations of File Replication Strategies for Large Heterogeneous Distributed Systems" received the "Best Workshop Paper on Heterogenous Systems" of the 24th International European Conference on Parallel and Distributed Computing (EuroPar'2018).

BEST PAPERS AWARDS :

[15]
H. CASANOVA, A. LEGRAND, M. QUINSON, F. SUTER.*SMPI Courseware: Teaching Distributed-Memory Computing with MPI in Simulation*, in "EduHPC-18 - Workshop on Education for High-Performance Computing", Dallas, United States, November 2018, p. 1-10, https://hal.inria.fr/hal-01891513

[17]
A. CHAI, S. CAMARASU-POP, T. GLATARD, H. BENOIT-CATTIN, F. SUTER.*Evaluation through Realistic Simulations of File Replication Strategies for Large Heterogeneous Distributed Systems*, in "Europar 2018 - 24th International European Conference on Parallel and Distributed Computing ; Workshop HeteroPar 2018", Turin, Italy, Lecture Notes in Computer Science (LNCS), August 2018, forthcoming, https://hal.archives-ouvertes.fr/hal-01887369

6. New Software and Platforms

6.1. MAD

Madeus Application Deployer

KEYWORDS: Automatic deployment - Distributed Software - Component models - Cloud computing

SCIENTIFIC DESCRIPTION: MAD is a Python implementation of the Madeus deployment model for multi-component distributed software. Precisely, it allows to: 1. describe the deployment process and the dependencies of distributed software components in accordance with the Madeus model, 2. describe an assembly of components, resulting in a functional distributed software, 3. automatically deploy the component assembly of distributed software following the operational semantics of Madeus.

RELEASE FUNCTIONAL DESCRIPTION: Initial submission with basic functionalities of MAD

NEWS OF THE YEAR: Operational prototype.

- Participants: Christian Pérez, Dimitri Pertin, Hélène Coullon and Maverick Chardet
- Partners: IMT Atlantique - LS2N - LIP
- Contact: Hélène Coullon
- Publications: Madeus: A formal deployment model - Behavioral interfaces for reconfiguration of component models

6.2. DIET

Distributed Interactive Engineering Toolbox

KEYWORDS: Scheduling - Clusters - Grid - Cloud - HPC - Middleware - Data management.

FUNCTIONAL DESCRIPTION: Middleware for grids and clouds. Toolbox for the use and porting of intensive computing applications on heterogeneous architectures.

RELEASE FUNCTIONAL DESCRIPTION: - Upgrade to support Cmake 3.3 and later - Update workflow unit tests to take the results of the execution into account - DIET workflow engine was improved

NEWS OF THE YEAR: Work on the next DIET release (DIET 2.11) New DIET Webboard based on Angular Two news biological application platform based on DIET (Aevol and Wasabi) Rutgers University (NJ, USA) Collaboration

- Participants: Joel Faubert, Hadrien Croubois, Abdelkader Amar, Arnaud Lefray, Aurélien Bouteiller, Benjamin Isnard, Daniel Balouek, Eddy Caron, Eric Bois, Frédéric Desprez, Frédéric Lombart, Gaël Le Mahec, Guillaume Verger, Huaxi Zhang, Jean-Marc Nicod, Jonathan Rouzaud-Cornabas, Lamiel Toch, Maurice Faye, Peter Frauenkron, Philippe Combes, Philippe Laurent, Raphaël Bolze and Yves Caniou
- Partners: CNRS - ENS Lyon - UCBL Lyon 1 - Sysfera
- Contact: Eddy Caron
- URL: http://graal.ens-lyon.fr/diet/

6.3. SimGrid

KEYWORDS: Large-scale Emulators - Grid Computing - Distributed Applications
SCIENTIFIC DESCRIPTION: SimGrid is a toolkit that provides core functionalities for the simulation of distributed applications in heterogeneous distributed environments. The simulation engine uses algorithmic and implementation techniques toward the fast simulation of large systems on a single machine. The models are theoretically grounded and experimentally validated. The results are reproducible, enabling better scientific practices.

Its models of networks, cpus and disks are adapted to (Data)Grids, P2P, Clouds, Clusters and HPC, allowing multi-domain studies. It can be used either to simulate algorithms and prototypes of applications, or to emulate real MPI applications through the virtualization of their communication, or to formally assess algorithms and applications that can run in the framework.

The formal verification module explores all possible message interleavings in the application, searching for states violating the provided properties. We recently added the ability to assess liveness properties over arbitrary and legacy codes, thanks to a system-level introspection tool that provides a finely detailed view of the running application to the model checker. This can for example be leveraged to verify both safety or liveness properties, on arbitrary MPI code written in C/C++/Fortran.
NEWS OF THE YEAR: There were 3 major releases in 2018: The public API was sanitized (with compatibility wrappers in place). Th documentation was completely overhauled. Our continuous integration was greatly improved (45 Proxy Apps + BigDFT + StarPU + BatSim now tested nightly). Some kernel headers are now installed, allowing external plugins. Allow dynamic replay of MPI apps, controlled by S4U actors. Port the MPI trace replay engine to C++, fix visualization (+ the classical bug fixes and doc improvement).

- Participants: Adrien Lèbre, Arnaud Legrand, Augustin Degomme, Florence Perronnin, Frédéric Suter, Jean-Marc Vincent, Jonathan Pastor, Luka Stanisic and Martin Quinson
- Partners: CNRS - ENS Rennes
- Contact: Martin Quinson
- URL: https://simgrid.org/

6.4. Kwapi

KiloWatt API
KEYWORD: Power monitoring
FUNCTIONAL DESCRIPTION: Kwapi is a software framework dealing with energy monitoring of large scale infrastructures through heterogeneous energy sensors. Kwapi has been designed inside the FSN XLCloud project for Openstack infrastructures. Through the support of Hemera Inria project, kwapi has been extended and deployed in production mode to support easy and large scale energy profiling of the Grid5000 resources. Kwapi now supports high frequency powermeters of the Grid5000 Lyon platform.

RELEASE FUNCTIONAL DESCRIPTION: - many bugfixes - multiprocessing instead of multithreading - many optimizations

- Participants: François Rossigneux, Jean-Patrick Gelas, Laurent Lefèvre, Laurent Pouilloux, Simon Delamare and Matthieu Imbert
- Contact: Laurent Lefèvre
- URL: https://launchpad.net/kwapi

6.5. execo

KEYWORDS: Toolbox - Deployment - Orchestration - Python
FUNCTIONAL DESCRIPTION: Execo offers a Python API for asynchronous control of local or remote, standalone or parallel, unix processes. It is especially well suited for quickly and easily scripting workflows of parallel/distributed operations on local or remote hosts: automate a scientific workflow, conduct computer science experiments, perform automated tests, etc. The core python package is execo. The execo_g5k package provides a set of tools and extensions for the Grid5000 testbed. The execo_engine package provides tools to ease the development of computer sciences experiments.

RELEASE FUNCTIONAL DESCRIPTION: - misc python3 support fixes - basic documentation for wheezy compatible package build - remove some debug outputs - fix crash in processes conductor in some situations - improve/fix process stdout/stderr handlers - fix ge_cluste_networ_equipments - add a FAQ

- Participants: Florent Chuffart, Laurent Pouilloux and Matthieu Imbert
- Contact: Matthieu Imbert
- URL: http://execo.gforge.inria.fr

6.6. SeeDep

Seed based Deployment
KEYWORDS: Reproducibility - Deployment - Cloud
SCIENTIFIC DESCRIPTION: SeeDep aims at devising a new way where researchers can communicate in a comprehensive and accurate way the experimentation set-up used in their work. It lies on two components: (i) a public algorithm that generates experimentation networks, and (ii) a generation key (i.e. a seed) that can be shared which specifies the said network. Therefore, researchers only need to share (in their paper for instance) the "generation key" that corresponds to their experimentation network. With such key, any other researcher/professional will be able to re-generate a comprehensive and accurate model of the same network.
FUNCTIONAL DESCRIPTION: SeeDep is a framework aiming at generating, reproducing and deploying experiments set-up on different Cloud platforms.

- Participants: Cyril Seguin and Eddy Caron
- Partner: Nokia Bell Labs
- Contact: Eddy Caron

6.7. libkomp

Runtime system libkomp
KEYWORDS: HPC - Multicore - OpenMP
FUNCTIONAL DESCRIPTION: libKOMP is a runtime support for OpenMP compatible with différent compiler: GNU gcc/gfortran, Intel icc/ifort or clang/llvm. It is based on source code initially developed by Intel for its own OpenMP runtime, with extensions from Kaapi software (task representation, task scheduling). Moreover it contains an OMPT module for recording trace of execution.

RELEASE FUNCTIONAL DESCRIPTION: Initial version

- Contact: Thierry Gautier
- URL: http://gitlab.inria.fr/openmp/libkomp

6.8. Platform: Grid'5000

Participants: Laurent Lefèvre, Simon Delamare, David Loup, Christian Perez.

FUNCTIONAL DESCRIPTION

The Grid'5000 experimental platform is a scientific instrument to support computer science research related to distributed systems, including parallel processing, high performance computing, cloud computing, operating systems, peer-to-peer systems and networks. It is distributed on 10 sites in France and Luxembourg, including Lyon. Grid'5000 is a unique platform as it offers to researchers many and varied hardware resources and a complete software stack to conduct complex experiments, ensure reproducibility and ease understanding of results. In 2018, a new generation of high speed wattmeters has been deployed on the Lyon site. They allow energy monitoring with up to 50 measurements per second. In parallel, a new version of kwapi (software stack for energy monitoring) has been proposed and redesigned.

- Contact: Laurent Lefèvre
- URL: https://www.grid5000.fr/

6.9. Platform: Leco

Participants: Thierry Gautier, Laurent Lefèvre, Christian Perez.

FUNCTIONAL DESCRIPTION

The LECO experimental platform is a new medium size scientific instrument funded by DRRT to investigate research related to BigData and HPC. It is located in Grenoble as part of the the HPCDA computer managed by UMS GRICAD. The platform has been deployed in 2018 and was available for experiment since the summer. All the nodes of the platform are instrumented to capture the energy consumption and data are available through the Kwapi software.

- Contact: Thierry Gautier

6.10. Platform: SILECS

Participants: Laurent Lefèvre, Simon Delamare, Christian Perez.

The SILECS infrastructure (IR ministère) aims at providing an experimental platform for experimental computer Science (Internet of things, clouds, hpc, big data, *etc.*). This new infrastructure is based on two existing infrastructures, Grid'5000 and FIT.

- Contact: Christian Perez
- URL: https://www.silecs.net/

7. New Results

7.1. Energy Efficiency in HPC and Large Scale Distributed Systems

Participants: Laurent Lefèvre, Dorra Boughzala, Christian Perez, Issam Raïs, Mathilde Boutigny.

7.1.1. Building and Exploiting the Table of Leverages in Large Scale HPC Systems

Large scale distributed systems and supercomputers consume huge amounts of energy. To address this issue, an heterogeneous set of capabilities and techniques that we call leverages exist to modify power and energy consumption in large scale systems. This includes hardware related leverages (such as Dynamic Voltage and Frequency Scaling), middleware (such as scheduling policies) and application (such as the precision of computation) energy leverages. Discovering such leverages, benchmarking and orchestrating them, remains a real challenge for most of the users. We have formally defined energy leverages, and we proposed a solution to automatically build the table of leverages associated with a large set of independent computing resources. We have shown that the construction of the table can be parallelized at very large scale with a set of independent nodes in order to reduce its execution time while maintaining precision of observed knowledge [22], [25].

7.1.2. Automatic Energy Efficient HPC Programming: A Case Study

Energy consumption is one of the major challenges of modern datacenters and supercomputers. By applying Green Programming techniques, developers have to iteratively implement and test new versions of their software, thus evaluating the impact of each code version on their energy, power and performance objectives. This approach is manual and can be long, challenging and complicated, especially for High Performance Computing applications. In [24], we formally introduces the definition of the Code Version Variability (CVV) leverage and present a first approach to automate Green Programming (*i.e.,* CVV usage) by studying the specific use-case of an HPC stencil-based numerical code, used in production. This approach is based on the automatic generation of code versions thanks to a Domain Specific Language (DSL), and on the automatic choice of code version through a set of actors. Moreover, a real case study is introduced and evaluated though a set of benchmarks to show that several trade-offs are introduced by CVV 1. Finally, different kinds of production scenarios are evaluated through simulation to illustrate possible benefits of applying various actors on top of the CVV automation.

7.1.3. Performance and Energy Analysis of OpenMP Runtime Systems with Dense Linear Algebra Algorithms

In the article [9], we analyze performance and energy consumption of five OpenMP runtime systems over a non-uniform memory access (NUMA) platform. We also selected three CPU-level optimizations or techniques to evaluate their impact on the runtime systems: processors features Turbo Boost and C-States, and CPU Dynamic Voltage and Frequency Scaling through Linux CPUFreq governors. We present an experimental study to characterize OpenMP runtime systems on the three main kernels in dense linear algebra algorithms (Cholesky, LU, and QR) in terms of performance and energy consumption. Our experimental results suggest that OpenMP runtime systems can be considered as a new energy leverage, and Turbo Boost, as well as C-States, impacted significantly performance and energy. CPUFreq governors had more impact with Turbo Boost disabled, since both optimizations reduced performance due to CPU thermal limits. An LU factorization with concurrent-write extension from libKOMP achieved up to 63% of performance gain and 29% of energy decrease over original PLASMA algorithm using GNU C compiler (GCC) libGOMP runtime.

7.1.4. Energy Simulation of GPU based Infrastructures

Through the IPL Hac-Specis and the PhD of Dorra Boughzala we begin to explore the modeling and calibrating of energy consumption of GPU architectures. We use the SimGrid simulation framework for the integration and validation on large scale systems.

7.2. HPC Component Models and Runtimes

Participants: Thierry Gautier, Christian Perez, Jérôme Richard.

7.2.1. On the Impact of OpenMP Task Granularity

Tasks are a good support for composition. During the development of a high-level component model for HPC, we have experimented to manage parallelism from components using OpenMP tasks. Since version 4-0, the standard proposes a model with dependent tasks that seems very attractive because it enables the description of dependencies between tasks generated by different components without breaking maintainability constraints such as separation of concerns. In [20], we present our feedback on using OpenMP in our context. We discover that our main issues are a too coarse task granularity for our expected performance on classical OpenMP runtimes, and a harmful task throttling heuristic counter-productive for our applications. We present a completion time breakdown of task management in the Intel OpenMP runtime and propose extensions evaluated on a testbed application coming from the Gysela application in plasma physics.

7.2.2. Building and Auto-Tuning Computing Kernels: Experimenting with BOAST and StarPU in the GYSELA Code

Modeling turbulent transport is a major goal in order to predict confinement performance in a tokamak plasma. The gyrokinetic framework considers a computational domain in five dimensions to look at kinetic issues in a plasma; this leads to huge computational needs. Therefore, optimization of the code is an especially important aspect, especially since coprocessors and complex manycore architectures are foreseen as building blocks for Exascale systems. This project [6] aims to evaluate the applicability of two auto-tuning approaches with the BOAST and StarPU tools on the gysela code in order to circumvent performance portability issues. A specific computation intensive kernel is considered in order to evaluate the benefit of these methods. StarPU enables to match the performance and even sometimes outperform the hand-optimized version of the code while leaving scheduling choices to an automated process. BOAST on the other hand reveals to be well suited to get a gain in terms of execution time on four architectures. Speedups in-between 1.9 and 5.7 are obtained on a cornerstone computation intensive kernel.

7.3. Modeling and Simulation of Parallel Applications and Distributed Infrastructures

Participants: Eddy Caron, Zeina Houmani, Frédéric Suter.

7.3.1. SMPI Courseware: Teaching Distributed-Memory Computing with MPI in Simulation

It is typical in High Performance Computing (HPC) courses to give students access to HPC platforms so that they can benefit from hands-on learning opportunities. Using such platforms, however, comes with logistical and pedagogical challenges. For instance, a logistical challenge is that access to representative platforms must be granted to students, which can be difficult for some institutions or course modalities; and a pedagogical challenge is that hands-on learning opportunities are constrained by the configurations of these platforms. A way to address these challenges is to instead simulate program executions on arbitrary HPC platform configurations. In [15] we focus on simulation in the specific context of distributed-memory computing and MPI programming education. While using simulation in this context has been explored in previous works, our approach offers two crucial advantages. First, students write standard MPI programs and can both debug and analyze the performance of their programs in simulation mode. Second, large-scale executions can be simulated in short amounts of time on a single standard laptop computer. This is possible thanks to SMPI, an MPI simulator provided as part of SimGrid. After detailing the challenges involved when using HPC platforms for HPC education and providing background information about SMPI, we present SMPI Courseware. SMPI Courseware is a set of in-simulation assignments that can be incorporated into HPC courses to provide students with hands-on experience for distributed-memory computing and MPI programming learning objectives. We describe some these assignments, highlighting how simulation with SMPI enhances the student learning experience.

7.3.2. Evaluation through Realistic Simulations of File Replication Strategies for Large Heterogeneous Distributed Systems

File replication is widely used to reduce file transfer times and improve data availability in large distributed systems. Replication techniques are often evaluated through simulations, however, most simulation platform models are oversimplified, which questions the applicability of the findings to real systems. In [17], we investigate how platform models influence the performance of file replication strategies on large heterogeneous distributed systems, based on common existing techniques such as prestaging and dynamic replication. The novelty of our study resides in our evaluation using a realistic simulator. We consider two platform models: a simple hierarchical model and a detailed model built from execution traces. Our results show that conclusions depend on the modeling of the platform and its capacity to capture the characteristics of the targeted production infrastructure. We also derive recommendations for the implementation of an optimized data management strategy in a scientific gateway for medical image analysis.

7.3.3. WRENCH: Workflow Management System Simulation Workbench

Scientific workflows are used routinely in numerous scientific domains, and Workflow Management Systems (WMSs) have been developed to orchestrate and optimize workflow executions on distributed platforms. WMSs are complex software systems that interact with complex software infrastructures. Most WMS research and development activities rely on empirical experiments conducted with full-fledged software stacks on actual hardware platforms. Such experiments, however, are limited to hardware and software infrastructures at hand and can be labor- and/or time-intensive. As a result, relying solely on real-world experiments impedes WMS research and development. An alternative is to conduct experiments in simulation.

In [16] we presented WRENCH, a WMS simulation framework, whose objectives are (i) accurate and scalable simulations; and (ii) easy simulation software development. WRENCH achieves its first objective by building on the SimGrid framework. While SimGrid is recognized for the accuracy and scalability of its simulation models, it only provides low-level simulation abstractions and thus large software development efforts are required when implementing simulators of complex systems. WRENCH thus achieves its second objective by providing high- level and directly re-usable simulation abstractions on top of SimGrid. After describing and giving rationales for WRENCH's software architecture and APIs, we present a case study in which we apply WRENCH to simulate the Pegasus production WMS. We report on ease of implementation, simulation accuracy, and simulation scalability so as to determine to which extent WRENCH achieves its two above objectives. We also draw both qualitative and quantitative comparisons with a previously proposed workflow simulator.

7.3.4. A Microservices Architectures for Data-Driven Service Discovery

Usual microservices discovery mechanisms are normally based on a specific user need *(Goal-based approaches)*. However, in today's evolving architectures, users need to discover what features they can take advantage of before looking for the available microservices. In collaboration with RDI2 (Rutgers University) we developped a data-driven microservices architecture that allows users to discover, from specific objects, the features that can be exerted on these objects as well as all the microservices dedicated to them [28]. This architecture, based on the main components of the usual microservices architectures, adopts a particular communication strategy between clients and registry to achieve the goal. This article contains a representation of a microservice data model and a P2P model that transforms our architecture into a robust and scalable system. Also, we designed a prototype to validate our approach using Istio library.

7.4. Cloud Resource Management

Participants: Eddy Caron, Hadrien Croubois, Jad Darrous, Christian Perez.

7.4.1. Nitro: Network-Aware Virtual Machine Image Management in Geo-Distributed Clouds

Recently, most large cloud providers, like Amazon and Microsoft, replicate their Virtual Machine Images (VMIs) on multiple geographically distributed data centers to offer fast service provisioning. Provisioning a service may require to transfer a VMI over the wide-area network (WAN) and therefore is dictated by the distribution of VMIs and the network bandwidth in-between sites. Nevertheless, existing methods to facilitate VMI management (*i.e.,* retrieving VMIs) overlook network heterogeneity in geo-distributed clouds. In [19], we design, implement and evaluate Nitro, a novel VMI management system that helps to minimize the transfer time of VMIs over a heterogeneous WAN. To achieve this goal, Nitro incorporates two complementary features. First, it makes use of deduplication to reduce the amount of data which will be transferred due to the high similarities within an image and in-between images. Second, Nitro is equipped with a network-aware data transfer strategy to effectively exploit links with high bandwidth when acquiring data and thus expedites the provisioning time. Experimental results show that our network-aware data transfer strategy offers the optimal solution when acquiring VMIs while introducing minimal overhead. Moreover, Nitro outperforms state-of-the-art VMI storage systems (*e.g.,* OpenStack Swift) by up to 77%.

7.4.2. Toward an Autonomic Engine for Scientific Workflows and Elastic Cloud Infrastructure

The constant development of scientific and industrial computation infrastructures requires the concurrent development of scheduling and deployment mechanisms to manage such infrastructures. Throughout the last decade, the emergence of the Cloud paradigm raised many hopes, but achieving full platform autonomicity is still an ongoing challenge. We built a workflow engine that integrated the logic needed to manage workflow execution and Cloud deployment on its own. More precisely, we focus on Cloud solutions with a dedicated Data as a Service (DaaS) data management component. Our objective was to automate the execution of workflows submitted by many users on elastic Cloud resources. This contribution proposes a modular middleware infrastructure and details the implementation of the underlying modules:

- A workflow clustering algorithm that optimises data locality in the context of DaaS-centered communications;
- A dynamic scheduler that executes clustered workflows on Cloud resources;
- A deployment manager that handles the allocation and deallocation of Cloud resources according to the workload characteristics and users' requirements.

All these modules have been implemented in a simulator to analyse their behaviour and measure their effectiveness when running both synthetic and real scientific workflows. We also implemented these modules in the Diet middleware to give it new features and prove the versatility of this approach. Simulation running the WASABI workflow (waves analysis based inference, a framework for the reconstruction of gene regulatory networks) showed that our approach can decrease the deployment cost by up to 44% while meeting the required deadlines [13].

7.4.3. Madeus: A Formal Deployment Model

Distributed software architecture is composed of multiple interacting modules, or components. Deploying such software consists in installing them on a given infrastructure and leading them to a functional state. However, since each module has its own life cycle and might have various dependencies with other modules, deploying such software is a very tedious task, particularly on massively distributed and heterogeneous infrastructures. To address this problem, many solutions have been designed to automate the deployment process. In [18], we introduce Madeus, a component-based deployment model for complex distributed software. Madeus accurately describes the life cycle of each component by a Petri net structure, and is able to finely express the dependencies between components. The overall dependency graph it produces is then used to reduce deployment time by parallelizing deployment actions. While this increases the precision and performance of the model, it also increases its complexity. For this reason, the operational semantics need to be clearly defined to prove results such as the termination of a deployment. In this paper, we formally describe the operational semantics of Madeus, and show how it can be used in a use-case: the deployment of a real and large distributed software (*i.e.,* , OpenStack).

In [18], we have proposed an extension based on component behavioral interfaces to the Aeolus component model to better separate the concerns of component users (*e.g.,* application architect) from component developers.

7.5. Data Stream Processing on Edge Computing

Participants: Eddy Caron, Felipe Rodrigo de Souza, Marcos Dias de Assunção, Laurent Lefèvre, Alexandre Da Silva Veith.

7.5.1. Latency-Aware Placement of Data Stream Analytics on Edge Computing

The interest in processing data events under stringent time constraints as they arrive has led to the emergence of architecture and engines for data stream processing. Edge computing, initially designed to minimize the latency of content delivered to mobile devices, can be used for executing certain stream processing operations. Moving opera- tors from cloud to edge, however, is challenging as operator-placement decisions must consider the application requirements and the network capabilities. We introduce strategies to create placement configurations for data stream processing applications whose operator topologies follow series parallel graphs[35]. We consider the operator characteristics and requirements to improve the response time of such appli- cations. Results show that our strategies can improve the response time in up to 50% for application graphs comprising multiple forks and joins while transferring less data and better using the resources.

7.5.2. Estimating Throughput of Stream Processing Applications in FoG Computing

Recent trends exploit decentralized infrastructures (e.g.. Fog computing) to deploy DSP (Data Stream Processing) applications and leverage the computational power. Fog computing overlaps some features of Cloud computing and includes others, for instance, location awareness. The operator placement problem consists of determining, within a set of distributed computing resources, the computing resources that should host and execute each operator of the DSP application, with the goal of optimizing QoS requirements of the application. The QoS requirements of the application refer to processing time, costs, throughput, etc. We propose a model to estimate the application throughput at each layer of Fog computing (Devices, Edge and Cloud) by considering a given placement solution. The estimated throughput provides a useful insight to determine the amount of physical resources to meet the QoS requirements. The model allows to identify the application bottleneck, when facing data rate variations, and provides information to self-scale in or out the DSP application.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. IFPEN

We have a collaboration with IFPEN (http://ifpenergiesnouvelles.com/). IFPEN develops numerical codes to solve PDE with specific adaption of the preconditioning step to fit the requirement of their problems. With a PhD student (Adrien Roussel) we have studied the parallel implementation of multi-level decomposition domains on many-core architecture and KNL processor.

8.1.2. Nokia Bell Labs

AVALON has been actively collaborating with Nokia, formerly Alcatel-Lucent Bell Labs, in the framework of the Nokia/Alcatel-Lucent Inria Joint Laboratory. We was involved in the following Research Actions (Actions de Recherche (ADR) in French) of this laboratory. ADR Nokia Bell Labs /Inria: Procedural Generation of Networks for Security Research & Experimentations. The objective of this project is to address such challenge. We aim at devising a new way where researchers can communicate in a comprehensive and accurate way the experimentation set-up used in their work. The main objective would be to research and develop the procedural generation of credible network topologies and test beds resembling real operational infrastructures of various kinds (e.g. classical ICT, virtualized Cloud or SDN based, SCADA infrastructures etc.), as a method of creating data algorithmically as opposed to manually. This work is done with a postdoc position: Cyril Seguin.

8.2. Bilateral Grants with Industry

8.2.1. Orange

We have a collaboration with Orange. This collaboration is sealed through a CIFRE Phd grant. The research of the Phd student (Arthur Chevalier) focuses on placement and compliance aspects of software licenses in a Cloud architecture. Today, the use of software is regulated by licenses, whether they are free, paid for and with or without access to its sources. The number of licenses required for specific software can be calculated with several metrics, each defined by the software vendor. Our goal is to propose a deployment algorithm that takes into account different metrics.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. CPER

Participants: Thierry Gautier, Laurent Lefèvre, Christian Perez.

The LECO experimental platform is a new medium size scientific instrument deployed in Grenoble in 2018. It was funded by the CPER 2015-2020 LECO++ to investigate research related to BigData and HPC.

9.2. National Initiatives

9.2.1. PIA

9.2.1.1. PIA ELCI, Environnement Logiciel pour le Calcul Intensif, 2014-2018

Participants: Mathilde Boutigny, Thierry Gautier, Laurent Lefèvre, Christian Perez, Issam Raïs, Jérôme Richard, Philippe Virouleau.

The ELCI PIA project is coordinated by BULL with several partners: CEA, Inria, SAFRAB, UVSQ.

This project aims to improve the support for numerical simulations and High Performance Computing (HPC) by providing a new generation software stack to control supercomputers, to improve numerical solvers, and pre- and post computing software, as well as programming and execution environment. It also aims to validate the relevance of these developments by demonstrating their capacity to deliver better scalability, resilience, modularity, abstraction, and interaction on some application use-cases. AVALON is involved in WP1 and WP3 ELCI Work Packages through the PhD of Issam Raïs and the postdoc of Hélène Coullon. Laurent Lefèvre is the Inria representative in the ELCI technical committee.

9.2.2. MRSEI

9.2.2.1. Fennec, FastEr NaNo-Characterisation, 24 months, 2018-2021

Participants: Eddy Caron, Christian Perez.

The goal of the ANR-MRSEI FENNEC project is to support the submission of a project to the European call DT-NMBP-08-2019 entitled "Real-time nano-characterisation technologies (RIA)".

9.2.3. Inria Large Scale Initiative

9.2.3.1. DISCOVERY, DIStributed and COoperative management of Virtual EnviRonments autonomouslY, 4 years, 2015-2019

Participants: Maverick Chardet, Jad Darrous, Christian Perez.

To accommodate the ever-increasing demand for Utility Computing (UC) resources, while taking into account both energy and economical issues, the current trend consists in building larger and larger Data Centers in a few strategic locations. Although such an approach enables UC providers to cope with the actual demand while continuing to operate UC resources through centralized software system, it is far from delivering sustainable and efficient UC infrastructures for future needs.

The DISCOVERY initiative aims at exploring a new way of operating Utility Computing (UC) resources by leveraging any facilities available through the Internet in order to deliver widely distributed platforms that can better match the geographical dispersal of users as well as the ever increasing demand. Critical to the emergence of such locality-based UC (LUC) platforms is the availability of appropriate operating mechanisms. The main objective of DISCOVERY is to design, implement, demonstrate and promote the LUC Operating System (OS), a unified system in charge of turning a complex, extremely large-scale and widely distributed infrastructure into a collection of abstracted computing resources which is efficient, reliable, secure and at the same time friendly to operate and use.

To achieve this, the consortium is composed of experts in research areas such as large-scale infrastructure management systems, network and P2P algorithms. Moreover two key network operators, namely Orange and RENATER, are involved in the project.

By deploying and using such a LUC Operating System on backbones, our ultimate vision is to make possible to host/operate a large part of the Internet by its internal structure itself: A scalable set of resources delivered by any computing facilities forming the Internet, starting from the larger hubs operated by ISPs, government and academic institutions, to any idle resources that may be provided by end-users.

9.2.3.2. HAC SPECIS, High-performance Application and Computers, Studying PErformance and Correctness In Simulation, 4 years, 2016-2020

Participants: Dorra Boughzala, Idriss Daoudi, Thierry Gautier, Laurent Lefèvre, Frédéric Suter.

Over the last decades, both hardware and software of modern computers have become increasingly complex. Multi-core architectures comprising several accelerators (GPUs or the Intel Xeon Phi) and interconnected by high-speed networks have become mainstream in HPC. Obtaining the maximum performance of such heterogeneous machines requires to break the traditional uniform programming paradigm. To scale, application developers have to make their code as adaptive as possible and to release synchronizations as much as possible. They also have to resort to sophisticated and dynamic data management, load balancing, and scheduling strategies. This evolution has several consequences:

First, this increasing complexity and the release of synchronizations are even more error-prone than before. The resulting bugs may almost never occur at small scale but systematically occur at large scale and in a non deterministic way, which makes them particularly difficult to identify and eliminate.

Second, the dozen of software stacks and their interactions have become so complex that predicting the performance (in terms of time, resource usage, and energy) of the system as a whole is extremely difficult. Understanding and configuring such systems therefore becomes a key challenge.

These two challenges related to correctness and performance can be answered by gathering the skills from experts of formal verification, performance evaluation and high performance computing. The goal of the HAC SPECIS Inria Project Laboratory is to answer the methodological needs raised by the recent evolution of HPC architectures by allowing application and runtime developers to study such systems both from the correctness and performance point of view.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. POP

Participant: Frédéric Suter.

Program: H2020 Center of Excellence

Project acronym: POP

Project title: Performance Optimisation and Productivity

Duration: 2015-2018

Coordinator: Barcelona Supercomputing Centre (BSC)

Other partners: High Performance Computing Center Stuttgart of the University of Stuttgart (HLRS), Jülich Supercomputing Centre (JSC), Numerical Algorithm Group (NAG), Rheinisch-Westfälische Technische Hochschule Aachen (RWTH), TERATEC (TERATEC).

Abstract: The Center of Excellence for Performance Optimisation and Productivity provides performance optimisation and productivity services for academic and industrial codes. European's leading experts from the High Performance Computing field will help application developers getting a precise understanding of application and system behaviour. This project is supported by the European Commission under H2020 Grant Agreement No. 676553.

Established codes, but especially codes never undergone any analysis or performance tuning, may profit from the expertise of the POP services which use latest state-of-the-art tools to detect and locate bottlenecks in applications, suggest possible code improvements, and may even help by Proof-of-Concept experiments and mock-up test for customer codes on their own platforms.

9.3.2. Collaborations in European Programs, Except FP7 & H2020

9.3.2.1. COST IC1305 : Nesus

Participants: Marcos Dias de Assunção, Laurent Lefèvre.

Program: COST

Project acronym: IC1305

Project title: Network for Sustainable Ultrascale Computing (NESUS)

Duration: 2014-2018

Coordinator: Jesus Carretero (Univ. Madrid)

Abstract: Ultrascale systems are envisioned as large-scale complex systems joining parallel and distributed computing systems that will be two to three orders of magnitude larger that today's systems. The EU is already funding large scale computing systems research, but it is not coordinated across researchers, leading to duplications and inefficiencies. The goal of the NESUS Action is to establish an open European research network targeting sustainable solutions for ultrascale computing aiming at cross fertilization among HPC, large scale distributed systems, and big data management. The network will contribute to glue disparate researchers working across different areas and provide a meeting ground for researchers in these separate areas to exchange ideas, to identify synergies, and to pursue common activities in research topics such as sustainable software solutions (applications and system software stack), data management, energy efficiency, and resilience. In Nesus, Laurent Lefèvre is co-chairing the Working on Energy Efficiency (WG5).

9.4. International Initiatives

9.4.1. Inria International Labs

9.4.1.1. Joint Laboratory for Extreme Scale Computing (JLESC) (2014-2018)

Participants: Thierry Gautier, Christian Perez, Jérôme Richard.

Partners: NCSA (US), ANL (US), Inria (FR), Jülich Supercomputing Centre (DE), BSC (SP), Riken (JP).
The purpose of the Joint Laboratory for Extreme Scale Computing (JLESC) is to be an international, virtual organization whose goal is to enhance the ability of member organizations and investigators to make the bridge between Petascale and Extreme computing. The founding partners of the JLESC are Inria and UIUC. Further members are ANL, BSC, JSC and RIKEN-AICS.

JLESC involves computer scientists, engineers and scientists from other disciplines as well as from industry, to ensure that the research facilitated by the Laboratory addresses science and engineering's most critical needs and takes advantage of the continuing evolution of computing technologies.

9.4.1.2. Associate Team DALHIS – Data Analysis on Large-scale Heterogeneous Infrastructures for Science (2013-2018)

Participant: Frédéric Suter.

Partners: EPC Myriads (Rennes, Bretagne Atlantique), AVALON (Grenoble, Rhône-Alpes), Data Science and Technology Department (LBNL,USA).

The goal of the Inria-LBL collaboration is to create a collaborative distributed software ecosystem to manage data lifecycle and enable data analytics on distributed data sets and resources. Specifically, our goal is to build a dynamic software stack that is user-friendly, scalable, energy-efficient and fault tolerant. We plan to approach the problem from two dimensions: (i) Research to determine appropriate execution environments that allow users to seamlessly execute their end-to-end dynamic data analysis workflows in various resource environments and scales while meeting energy-efficiency, performance and fault tolerance goals; (ii) Engagement in deep partnerships with scientific teams and use a mix of user research with system software R&D to address specific challenges that these communities face, and inform future research directions from acquired experience.

9.4.2. Inria Associate Teams Not Involved in an Inria International Labs

9.4.2.1. Associate Team SUSTAM – Sustainable Ultra Scale compuTing, dAta and energy Management (2017-2020)

Participants: Eddy Caron, Hadrien Croubois, Marcos Dias de Assunção, Alexandre Da Silva Veith, Jean-Patrick Gelas, Olivier Glück, Laurent Lefèvre, Valentin Lorentz, Christian Perez, Issam Raïs.

International Partners: Rutgers University (United States) - RDI2 - Manish Parashar

The SUSTAM associate team will focus on the joint design of a multi-criteria orchestration framework dealing with resources, data and energy management in an sustainable way. The SUSTAM associated team will enable a long-term collaboration between the Inria AVALON team and the Rutgers Discovery Informatics Institute (RDI2) from Rutgers University (USA). The SUSTAM associated team is leaded by Laurent Lefèvre.

9.4.3. Participation in Other International Programs

9.4.3.1. Joint Project CNRS/University of Melbourne – Algorithms for Placement and Reconfiguration of Data Stream Processing Applications (2017-2018)

Participants: Marcos Dias de Assunção, Alexandre Da Silva Veith, Laurent Lefèvre.

Partner: Clouds Lab (The University of Melbourne, Australia).

Much of the “big data” produced today is created as continuous data streams that are most valuable when processed quickly. Several data stream processing frameworks have been designed for running on clusters of homogeneous computers. Under most frameworks, an application is a Direct Acyclic Graph (DAG) whose vertices are operators that execute transformations over the incoming data and edges that define how the data flows between operators. While cloud computing is a key infrastructure for deploying such frameworks, more modern solutions leverage the edges of the Internet (e.g. edge computing) to offload some of the processing from the cloud and hence reduce the end-to-end latency. The placement and reconfiguration of stream processing DAGs onto highly distributed and heterogeneous infrastructure are, however, challenging endeavours. This project aims to investigate algorithms for the placement and dynamic reconfiguration of stream processing components considering multiple criteria.

9.5. International Research Visitors

9.5.1. Research Stays Abroad

Participant: Olivier Glück.

Olivier Glück has been invited professor by Concordia University (Faculty of Engineering & Computer Science, Department of Computer Science & Software Engineering). He has conducted researches with Dr. Brigitte Jaumard, Professor & Research Chair, Tier 1, on the Optimization of Communication Networks. He has worked on the following Virtual Machine (VM) migration optimization problem: find a scheduling of VM migration that minimizes the makespan *i.e.,* total duration of the migration assuming that the current VM placement and the target one are given. He has proposed a new sequence-based optimization model with a Mixed Integer Linear Program (MILP), which not only guarantees the finding of the best VM migration scheduling but also the migration of the largest possible number of VMs in the case of deadlocks. He also worked on the design of heuristic algorithms for VM migration and a generator of real VM migration instances to evaluate the models and heuristics proposed. He has also worked on the task offloading problem in edge computing.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Laurent Lefèvre was co-General chair of the conference SBAC-PAD : The International Symposium on Computer Architecture and High Performance Computing, Lyon, France (September 24-27, 2018). He was co-organizer of colloquium : "Digital society vs eco-responsability: impact on billions of connected objects, networks and clouds", with Centre Jacques Cartier, Concordia University, Inria, GDS CNRS EcoInfo, Ecole Normale Superieure of Lyon (November 12-13, 2018). He co-organized the E3-RSD school on Energy Efficiency in Networks and Distributed Systems, Dinard (October 1-4, 2018) and the GreenDays@Toulouse v2.0 (2018 Edition) "From IoT to Exascale, what about energy efficiency and carbon emission reduction ?", Toulouse (July 2-3, 2018)

10.1.1.2. Member of the Organizing Committees

Laurent Lefèvre was co Special Session Organizer of Special Session on High Performance Computing Benchmarking and Optimization (HPBench 2018), during HPCS conference, Orleans (July 16-20, 2018).

Christian Perez served on the Organizing Committee of the 2018 International Conference on High Performance Computing & Simulation (HPCS 2017) as Conference Awards and Recognitions Co-Chair, on the Organizing Comittee of the French Journées Calculs et Données (Lyon, October 24-26) and of the 1st Grid'5000-FIT school (Sophia Antipolis, April 3-6).

Frédéric Suter served in the Organization commitee of the session on Exascale Computing for High Energy Physics of the 14th IEEE International Conference on eScience and chaired the "Parallelism" track of the Conférence d'informatique en Parallélisme, Architecture et Système (Compas 2018).

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

Christian Perez was member of the program committees of the IEEE International Conference on Cluster Computing 2018 (Cluster 2018), the 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid (CCGRID 2018), the 26th annual High Performance Computing Symposium 2018 (HPC '18), the 2018 Supercomputing Asia conference, and of the 4th International Workshop on Autonomic High Performance Computing. He was a member of the Birds of Feather Committee of SuperComputing 2018 and of the Tutorial Committee of ISC High Performance 2018.

Olivier Glück was a member of the program committees of PDP 2018 (26th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing) and ICA3PP-2018 (18th International Conference on Algorithms and Architectures for Parallel Processing).

Frédéric Suter was member of program committee of SBAC-PAD 2018 (30th International Symposium on Computer Architecture and High Performance Computing), VECPAR 2018 (13th International Meeting High Performance Computing for Computational Science) and ICA3PP-2018 (18th International Conference on Algorithms and Architectures for Parallel Processing)

10.1.3. Journal

10.1.3.1. Reviewer - Reviewing Activities

Christian Perez reviewed articles for the IEEE's Transactions on Big Data journal and for Oil & Gas Science and Technology journal of IFP Energies nouvelles.

Thierry Gautier reviewed articles for the ACM Transaction on Architecture and Code Optimization.

Frédéric Suter reviewed articles for the Journal of Parallel and Distributed Computing and Concurrency and Computation: Practice and Experience.

10.1.4. Invited Talks

Christian Perez has been invited to give the following talks:

- "Towards Reconfigurable HPC Component Models", Invited talk, 4th International Workshop on Autonomic High Performance Computing, Orléans, France, July 17.
- "HPC Component Models", Keynote talk, HPCS, Orléans, France, July 19, 2018.

Thierry Gautier has been invited to give the following talk:

- "Tasks' management in OpenMP: what about performance guarantee?", Maison de la Simulation, November 20, 2018.

Frédéric Suter has been invited to give the following talk:

- "Simulating MPI applications : the SMPI Approach" at the "Modeling and Simulation of HPC Architectures and Applications" mini-symposium at the SIAM Conference on Parallel Processing for Scientific Applications.

10.1.5. Leadership within the Scientific Community

Eddy Caron is animator and co-chair for the FIL (Fédération Informatique de Lyon) on the theme IDCHP (Informatique Distribuée et Calcul Haute Performance).

Laurent Lefèvre is animator and chair of the transversal action on "Energy" of the French GDR RSD ("Réseaux et Systèmes Distribués").

Christian Perez is co-leader of the pole Distributed Systems of the French GDR RSD ("Réseaux et Systèmes Distribués").

10.1.6. Scientific Expertise

Christian Perez reviewed two projects for the Regional Program STIC-AmSud.

Olivier Glück is member of the CNU (Conseil National des Universités) section 27 (Computer Science). He participated to the 2018 "Qualifications" session and "Suivi de carriere" session.

10.1.7. Research Administration

Eddy Caron is Deputy Director in charge of call for projects, research transfert and international affairs for the LIP. He is a member of ”Conseil Technologique Logiciel MINALOGIC". He is member and reviewer for InriaHub.

Olivier Glück is member of the "Conseil Académique" of Lyon 1 University and Lyon University.

Christian Perez represents INRIA in the overview board of the France Grilles Scientific Interest Group. He is a member of the executive board and the sites committee of the Grid'5000 Scientific Interest Group and member of the executive board of the Silecs testbed. He is a member of the Inria Grenoble Rhône-Alpes Strategic Orientation Committee. He is in charge of exploring potential scientific collaborations between INRIA and SKA France.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence: Yves Caniou, Algorithmique programmation impérative initiation, 25h, niveau L1, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Pratique d'Unix, 3h, niveau L1, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Programmation Concurrente, 35h et responsabilité d'UE, niveau L3, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Projet Informatique, 3h, niveau L3, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Réseaux, 45h, niveau L3, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Système d'Exploitation, 32h and responsible of UE, niveau L2, Université d'Ho Chi Minh Ville, Vietnam.

Licence: Yves Caniou, Système d'Exploitation, 39h et co-responsible of UE, niveau L3, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Projet pour l'Orientation en Master, 3h, niveau M1, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Responsible of alternance students, 30h, niveau M1, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Responsible of Master SRIV (Systèmes, Réseaux et Infrastructures Virtuelles), 30h, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Projet Bibliographie et Certifications, 4h and responsible of UE, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Sécurité, 36h et responsible of UE, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Sécurité, 20h, niveau M2, IGA Casablanca, Maroc.

Master: Yves Caniou, Systèmes Avancés, 4.5h, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Approfondissement Scientifique, 6h, niveau M2, Université Claude Bernard Lyon 1, France.

Licence: Eddy Caron, Projet 1, 48h, L3, ENS de Lyon. France.

Master: Eddy Caron, Projet Intégré, 42h, M1, ENS de Lyon. France.

Master: Eddy Caron, Système distribués, 30h, M1, ENS de Lyon. France.

Master: Eddy Caron, Advanced Topics in Scalable Data Management, 20h, M2, ENS de Lyon. France.

Licence: Olivier Glück, Licence pedagogical advisor, 30h, niveaux L1, L2, L3, Université Lyon 1, France.

Licence: Olivier Glück, Introduction Réseaux et Web, 32h, niveau L1, Université Lyon 1, France.

Licence: Olivier Glück, Bases de l'architecture pour la programmation, 62h, niveau L1, Université Lyon 1, France.

Licence: Olivier Glück, Algorithmique programmation impérative initiation, 50h, niveau L1, Université Lyon 1, France.

Licence: Olivier Glück, Réseaux, 2x70h, niveau L3, Université Lyon 1, France.

Master: Olivier Glück, Réseaux par la pratique, 8h, niveau M1, Université Lyon 1, France.

Master: Olivier Glück, Responsible of Master SRIV (Systèmes, Réseaux et Infrastructures Virtuelles) located at IGA Casablanca, 20h, niveau M2, IGA Casablanca, Maroc.

Master: Olivier Glück, Applications systèmes et réseaux, 30h, niveau M2, Université Lyon 1, France.

Master: Olivier Glück, Applications systèmes et réseaux, 24h, niveau M2, IGA Casablanca, Maroc.

Master: Olivier Glück, Administration des Systèmes et des Réseaux, 16h, niveau M2, Université Lyon 1, France.

Master: Laurent Lefèvre, Parallélisme, 12h, niveau M1, Université Lyon 1, France.

Master: Laurent Lefèvre, Réseaux avancés, 24h, niveau M2, IGA Casablanca, Maroc.

Licence: Frédéric Suter, Programmation Concurrente, 32h, niveau L3, Université Lyon 1, France.

Master: Jean-Patrick Gelas, Programmation embarquée et mobile des objets, 30h, niveau M1, Université Lyon 1, France

Master: Jean-Patrick Gelas, Introduction au Cloud Computing, 21h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Système d'exploitation, 45h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Projet en Informatique en Anglais, 15h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Réseaux Avancés, 24h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Sécurité et Admin des infra résx, 30h, niveau M2 (CCI), Université Lyon 1, France

Master: Jean-Patrick Gelas, Technologies embarquées, 18h, niveau M2 (Image), Université Lyon 1, France

Master: Jean-Patrick Gelas, Routage (BGP), Routeurs et IPv6, 12.5h, niveau M2, Université Lyon 1, France

Master: Jean-Patrick Gelas, Systèmes embarqués (GNU/Linux, Android, ARM, Arduino), 24h, niveau M2, Université Lyon 1, France

Master: Jean-Patrick Gelas, Analyse de performance, 3h, niveau M2 (TIW), Université Lyon 1, France

Master: Jean-Patrick Gelas, Cloud Computing, 15h, niveau M2 (TIW), Université Lyon 1, France

Master: Jean-Patrick Gelas, Stockage, Cloud et Virtualisation, 9.5h, niveau M2 (SRIV), Université Lyon 1, France

Master: Jean-Patrick Gelas, Développement informatique, 23h, niveau M2 (Data science), Université Lyon 1, France

10.2.2. Supervision

PhD: Issam Raïs, *Discover, model and combine energy leverages for large scale energy efficient infrastructures*, 29 sept. 2018, Laurent Lefèvre (dir), Anne Benoit (Roma Team, LIP, ENS Lyon, co-dir) and Anne-Cécile Orgerie (CNRS, Myriads team, Irisa Rennes, co-dir).

PhD: Hadrien Croubois, *Toward an autonomic engine for scientific workflows and elastic Cloud infrastructure*, 16 oct. 2018, Eddy Caron (dir).

PhD: Philippe Virouleau,*Studying and improving the use of NUMA architectures through runtime systems*, 5 jun. 2018, Fabrice Rastello (Inria, Corse team, dir), Thierry Gautier (AVALON team, LIP, ENS Lyon, co-dir) and François Broquedis (UGA, Corse team, Grenoble, co-dir).

PhD: Adrien Roussel, *Parallelization of iterative methods to solve sparse linear systems using task based runtime systems on multi and many-core architectures: application to Multi-Level Domain Decomposition methods*, 6 fev. 2018, Thierry Gautier (dir), Jean-Marc Gratien (IFPen).

PhD in progress: Dorra Boughzala, *Simulating Energy Consumption of Continuum Computing between Heterogeneous Numerical Infrastructures in HPC*, IPL Hac-Specis Inria, Laurent Lefèvre (dir) , Martin Quinson and Anne-Cécile Orgerie (Myriads, Rennes, co-dir) (since december 2017).

PhD in progress: Anchen Chai: *Simulation of the Distributed Execution of a Medical Imaging Simulator*, Hugues Benoit-Cattin (co-dir, CREATIS, INSA Lyon), Frédéric Suter (co-dir).

PhD in progress: Arthur Chevalier, *Optimisation du placement des licences logiciel des fonctions réseau dans le Cloud pour un déploiement économique et efficace*, Eddy Caron (dir), Noëlle Baillon (co-dir, Orange) (since october 2017).

PhD in progress: Jad Darrous, *Geo-distributed storage for distributed Cloud*, Gilles Fedak (dir) until Aug. 2017 then Christian Perez (dir), Shadi Ibrahim (co-dir).

PhD in progress: Idris Daoudi, *Simulation of OpenMP programs*, Olivier Aumage (dir, Storm team, Bordeaux), Thierry Gautier (co-dir) (since oct. 2018).

PhD in progress: Zeina Houmani, *A Data-driven microservices architecture for Deep Learning applications*, Eddy Caron (dir), Daniel Balouek-Thomert (Rutgers University) (since oct. 2018).

PhD pended: Aurélie Kong-Win-Chang: *Techniques de résilience pour l'ordonnancement de workflows sur plates-formes décentralisées (cloud computing) avec contraintes de sécurité*, Yves Robert (dir, ROMA, ÉNS-Lyon), Eddy Caron (co-dir) et Yves Caniou (co-dir).

PhD in progress: Felipe Rodrigo De Souza, *Networking Provisioning Algorithms for Highly Distributed Data Stream Processing*, École Doctorale, Eddy Caron (dir), Marcos Dias de Assunção (co-dir) (2017-2020).

PhD in progress: Alexandre Da Silva Veith: *Elastic Mechanisms for Big-Data Stream Analytics*, Labex MiLyon, Laurent Lefèvre (dir), Marcos Dias de Assunção (co-dir) (2016-2019).

PhD cancelled: Valentin Lorentz : *Energy traceability of data*, Gilles Fedak (dir), Laurent Lefèvre (co-dir) (2016-2018).

10.2.3. Juries

Eddy Caron was reviewer and examiner of the PhD defense committe of Bassirou Ngom. UPMC (Co-tutelle avec l'UCAD). (July 13, 2018). Hajer Salhi., University of Tunis El Manar (Phd defense will be in 2019). and examiner of the PhD defense committe of Anne-Lucie Vion, Université de Grenoble and Orange (March 29, 2018).

Laurent Lefèvre was PhD Opponent and Jury Member of the PhD defense committee of Selome Kostentinos Tesfatsion, Umea University, Sweden (April 16, 2018), reviewer and examiner of the PhD defense committee of Wilfried Yoro, Telecom SudParis (March 8, 2018), reviewer of the PhD of Jungmin Son, University of Melbourne, Australia (May 2018) and examiner of the PhD defense committee of David Guyon, IRISA, University of Rennes 1 (December 7, 2018).

Christian Perez was examiner of the HdR defense committe of Samuel Thibault, University of Bordeaux (December 12, 2018), reviewer and examiner of the PhD defense committe of Pierre Matri, University Polytenics of Madrid, Spain (June 10, 2018), Hugo Taboada, University of Bordeaux (December 11, 2018), Mohamed Abderrahim, University of Bretagne Loire (December 19, 2018), Stéphanie Challita, University of Lille (December 20, 2018), and examiner of the PhD defense committe of Yacine Taleb, European University of Brittany (October 2, 2018).

10.3. Popularization

10.3.1. Articles and contents

Laurent Lefèvre has been interviewed for the articles:

- "Le Groupe Casino remplace ses radiateurs par des data centers", Le Point Journal (December 23, 2018)
- "Comment cliquer sans trop polluer", Laurent Lefèvre, Le Progrès Journal, November 21, 2018
- "Google utilise-t-il réellement 100% d'énergie verte pour alimenter ses services ?", Libération, July 13, 2018
- "J'ai le (dé)clic écolo pour réduire mon empreinte", Journal La Vie, February 15, 2018

He was interviewed and filmed for "Pollution numérique/Digital pollution", Journal 19/20, France3 National Channel, December 9, 2018.

10.3.2. Education

Yves Caniou is responsible of the LPI Certification at Université Claude Bernard Lyon 1. Further discussions are in progress concerning the proposed tools and redaction of courses/questions.

In the contexte of the CHEL[s] (Collège des Hautes Etudes Lyon Science[s]) Eddy Caron is mentor for the MedTech project building the start-up VETARIA.

Jean-Patrick Gelas manage and build several modules of education about blockchain technologies for an erasmus+ european project. The BLISS project (Blockchain skills for ICT professionals), aims to improve the skills and competencies of ICT professionals by developing and making available educational resources and materials to address existing occupational needs and mismatches, resulting from the dynamic penetration of blockchain technology across all sectors of the EU economy (including banking, accounting, auditing and government services).

10.3.3. Interventions

Laurent Lefèvre has given several invited talks for :

- "Green-It and Video Games infrastructures - Jeu vidéo et collapsologie #2 : Meetup - Le dématérialisé, ça sauve les pandas ?", Meet-up Game Impact, Paris, November 29, 2018
- "Vers des Data Centers éco-responsables", Data Center World, Paris, November 28, 2018
- "Agile (d'accord !) mais Green (d'abord !)", Agile Tour 2018, Epitech, Bordeaux, November 2-3, 2018
- "Numérique et Environnement", with Françoise Berthoud, Groupe de travail développement durable, Centre Inria Bordeaux, October 19, 2018
- "CES 2018 : l'humain numérique de demain...", Journée Réseau Aramis, University Lyon2, Lyon, May 24, 2018
- "Le futur peut-il être numérique ?", with Françoise Berthoud and Jonathan Schaeffer, Journée Réseau Aramis, University Lyon2, Lyon, May 24, 2018
- "EE@G5K : the story continues...", Grid5000 Scientific Advisory Board, Sophia Antipolis, April 4, 2018
- "Rendre l'informatique plus verte", Conférence de formation des professeurs du secondaire en science informatique, Inria Montbonnot, March 21, 2018
- "WG5 : Energy efficiency final results", Network for Sustainable Computing (Nesus), COST European Action IC1305, Final Action Meeting, with Ariel Oleksiak, Madrid, Spain, March 15-16, 2018
- "CES 2018 - so little Green", Rencontres Ecoinfo, Lyon, January 30, 2018

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] H. CROUBOIS.*Toward an autonomic engine for scientific workflows and elastic Cloud infrastructure*, Université de Lyon, October 2018, https://tel.archives-ouvertes.fr/tel-01988995

[2] I. RAÏS.*Discover, model and combine energy leverages for large scale energy efficient infrastructures*, Université de Lyon, September 2018, https://tel.archives-ouvertes.fr/tel-01892387

[3] A. ROUSSEL.*Parallelization of iterative methods to solve sparse linear systems using task based runtime systems on multi and many-core architectures: application to Multi-Level Domain Decomposition methods*, Université Grenoble Alpes, February 2018, https://tel.archives-ouvertes.fr/tel-01753992

Articles in International Peer-Reviewed Journal

[4] F. ALMEIDA, M. DIAS DE ASSUNCAO, J. BARBOSA, V. BLANCO, I. BRANDIC, G. DA COSTA, M. F. DOLZ, A. C. ELSTER, M. JARUS, H. KARATZA, L. LEFÈVRE, I. MAVRIDIS, A. OLEKSIAK, A.-C. ORGERIE, J.-M. PIERSON.*Energy Monitoring as an Essential Building Block Towards Sustainable Ultrascale Systems*, in "Sustainable Computing : Informatics and Systems", March 2018, vol. 17, p. 27-42 [*DOI :* 10.1016/J.SUSCOM.2017.10.013], https://hal.inria.fr/hal-01627757

[5] A. BENOIT, L. LEFÈVRE, A.-C. ORGERIE, I. RAÏS.*Reducing the energy consumption of large scale computing systems through combined shutdown policies with multiple constraints*, in "International Journal of High Performance Computing Applications", January 2018, vol. 32, n^{o} 1, p. 176-188 [*DOI :* 10.1177/1094342017714530], https://hal.inria.fr/hal-01557025

[6] J. BIGOT, V. GRANDGIRARD, G. LATU, J.-F. MÉHAUT, L. F. MILLANI, C. PASSERON, S. Q. MASNADA, J. RICHARD, B. VIDEAU.*Building and Auto-Tuning Computing Kernels: Experimenting with BOAST and StarPU in the GYSELA Code*, in "ESAIM: Proceedings and Surveys", October 2018, vol. 63 (2018), p. 152 - 178 [*DOI :* 10.1051/PROC/201863152], https://hal.inria.fr/hal-01909325

[7] M. DIAS DE ASSUNCAO, A. DA SILVA VEITH, R. BUYYA.*Distributed Data Stream Processing and Edge Computing: A Survey on Resource Elasticity and Future Directions*, in "Journal of Network and Computer Applications", February 2018, vol. 103, p. 1–17 [*DOI :* 10.1016/J.JNCA.2017.12.001], https://hal.inria.fr/hal-01653842

[8] M. DIAS DE ASSUNCAO, L. LEFÈVRE.*Bare-Metal Reservation for Cloud: an Analysis of the Trade Off between Reactivity and Energy Efficiency*, in "Cluster Computing", 2018, p. 1-12 [*DOI :* 10.1007/S10586-017-1094-Y], https://hal.inria.fr/hal-01571288

[9] J. V. FERREIRA LIMA, I. RAÏS, L. LEFÈVRE, T. GAUTIER.*Performance and Energy Analysis of OpenMP Runtime Systems with Dense Linear Algebra Algorithms*, in "International Journal of High Performance Computing Applications", August 2018, p. 1-17 [*DOI :* 10.1177//1094342018792079], https://hal.inria.fr/hal-01957220

[10] I. RAÏS, A.-C. ORGERIE, M. QUINSON, L. LEFÈVRE.*Quantifying the Impact of Shutdown Techniques for Energy-Efficient Data Centers*, in "Concurrency and Computation: Practice and Experience", 2018, vol. 30, n^{o} 17, p. 1-13 [*DOI :* 10.1002/CPE.4471], https://hal.archives-ouvertes.fr/hal-01711812

International Conferences with Proceedings

[11] H. ACAR, H. BENFENATKI, J.-P. GELAS, C. FERREIRA DA SILVA, G. I. ALPTEKIN, A.-N. BENHARKAT, P. GHODOUS.*Software Greenability: A Case Study of Cloud-Based Business Applications Provisioning*, in "CLOUD 2018 - IEEE 11th International Conference on Cloud Computing", San Francisco, CA, United States, IEEE, July 2018, p. 875-878 [*DOI :* 10.1109/CLOUD.2018.00125], https://hal.archives-ouvertes.fr/hal-01887065

[12] J. ANJOS, K. J. MATTEUSSI, P. R. R. DE SOUZA, C. GEYER, A. DA SILVA VEITH, G. FEDAK, J. L. VICTORIA BARBOSA.*Enabling Strategies for Big Data Analytics in Hybrid Infrastructures*, in "2018 International Conference on High Performance Computing Simulation (HPCS)", Orléans, France, July 2018, p. 869-876, https://hal.inria.fr/hal-01875952

[13] A. BONNAFFOUX, E. CARON, H. CROUBOIS, O. GANDRILLON.*A Cloud-aware autonomous workflow engine and its application to Gene Regulatory Networks inference*, in "CLOSER 2018 - 8th International conference on Cloud computing and Service Science", Funchal, Madeira, Portugal, March 2018, p. 1-8, https://hal.inria.fr/hal-01808764

[14] Y. CANIOU, E. CARON, A. KONG WIN CHANG, Y. ROBERT.*Budget-aware scheduling algorithms for scientific workflows with stochastic task weights on heterogeneous IaaS Cloud platforms*, in "IPDPSW 2018 - IEEE International Parallel and Distributed Processing Symposium Workshops", Vancouver, Canada, IEEE, May 2018, p. 15-26 [*DOI :* 10.1109/IPDPSW.2018.00014], https://hal.inria.fr/hal-01808831

[15] *Best Paper*
H. CASANOVA, A. LEGRAND, M. QUINSON, F. SUTER.*SMPI Courseware: Teaching Distributed-Memory Computing with MPI in Simulation*, in "EduHPC-18 - Workshop on Education for High-Performance Computing", Dallas, United States, November 2018, p. 1-10, https://hal.inria.fr/hal-01891513.

[16] H. CASANOVA, S. PANDEY, J. OETH, R. TANAKA, F. SUTER, R. FERREIRA DA SILVA.*WRENCH: A Framework for Simulating Workflow Management Systems*, in "WORKS 2018 - 13th Workshop on Workflows in Support of Large-Scale Science", Dallas, United States, November 2018, p. 1-12, https://hal.inria.fr/hal-01948162

[17] *Best Paper*
A. CHAI, S. CAMARASU-POP, T. GLATARD, H. BENOIT-CATTIN, F. SUTER.*Evaluation through Realistic Simulations of File Replication Strategies for Large Heterogeneous Distributed Systems*, in "Europar 2018 - 24th International European Conference on Parallel and Distributed Computing ; Workshop HeteroPar 2018", Turin, Italy, Lecture Notes in Computer Science (LNCS), August 2018, forthcoming, https://hal.archives-ouvertes.fr/hal-01887369.

[18] M. CHARDET, H. COULLON, D. PERTIN, C. PÉREZ.*Madeus: A formal deployment model*, in "4PAD 2018 - 5th International Symposium on Formal Approaches to Parallel and Distributed Systems (hosted at HPCS 2018)", Orléans, France, July 2018, p. 1-8, https://hal.inria.fr/hal-01858150

[19] J. DARROUS, S. IBRAHIM, A. C. ZHOU, C. PÉREZ.*Nitro: Network-Aware Virtual Machine Image Management in Geo-Distributed Clouds*, in "CCGrid 2018 - 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing", Washington D.C., United States, IEEE, May 2018, p. 553-562 [*DOI :* 10.1109/CCGRID.2018.00082], https://hal.inria.fr/hal-01745405

[20] T. GAUTIER, C. PÉREZ, J. RICHARD.*On the Impact of OpenMP Task Granularity*, in "IWOMP 2018 - 14th International Workshop on OpenMP for Evolving Architectures", Barcelone, Spain, Springer, September 2018, p. 205-221 [*DOI :* 10.1007/978-3-319-98521-3_14], https://hal.inria.fr/hal-01901806

[21] B. NOËLLE, E. CARON, A. CHEVALIER, A.-L. VION.*Towards economic and compliant deployment of licenses in a Cloud architecture*, in "Workshop: Cloud Management and Operations. In conjunction with CLOUD 2018 - IEEE International Conference on Cloud Computing", San Francisco, United States, July 2018, https://hal.inria.fr/hal-01808751

[22] I. RAÏS, M. BOUTIGNY, L. LEFÈVRE, A.-C. ORGERIE, A. BENOIT.*Building the Table of Energy and Power Leverages for Energy Efficient Large Scale Systems*, in "HPCS: International Conference on High Performance Computing & Simulation", Orléans, France, July 2018, p. 284-291 [*DOI :* 10.1109/HPCS.2018.00056], https://hal.archives-ouvertes.fr/hal-01845970

[23] I. RAÏS, E. CARON, L. LEFÈVRE.*Thermal aware scheduling on distributed computing water heaters*, in "CCNC 2018 - IEEE Consumer Communications & Networking Conference", Las Vegas, United States, January 2018, p. 1-4 [*DOI :* 10.1109/CCNC.2018.8319253], https://hal.inria.fr/hal-01671728

[24] I. RAÏS, H. COULLON, L. LEFÈVRE, C. PÉREZ.*Automatic Energy Efficient HPC Programming: A Case Study*, in "ISPA 2018", Melbourne, Australia, December 2018, https://hal.inria.fr/hal-01922301

[25] I. RAÏS, L. LEFÈVRE, A.-C. ORGERIE, A. BENOIT.*Exploiting the Table of Energy and Power Leverages*, in "ICA3PP 2018 - 18th International Conference on Algorithms and Architectures for Parallel Processing", Guangzhou, China, November 2018, p. 1-10, https://hal.archives-ouvertes.fr/hal-01927829

[26] A. DA SILVA VEITH, M. DIAS DE ASSUNCAO, L. LEFÈVRE.*Latency-Aware Placement of Data Stream Analytics on Edge Computing*, in "Service-Oriented Computing", Hangzhou, Zhejiang, China, Springer International Publishing, November 2018, p. 215-229, https://hal.inria.fr/hal-01875936

[27] P. DE SOUZA JUNIOR, K. J. MATTEUSSI, J. C. S. DOS ANJOS, J. D. D. D. SANTOS, A. DA SILVA VEITH, C. R. GEYER.*Aten: A Dispatcher for Big Data Applications in Heterogeneous Systems*, in "2018 International Conference on High Performance Computing Simulation (HPCS)", Orléans, France, July 2018, p. 585-592 [*DOI :* 10.1109/HPCS.2018.00098], https://hal.inria.fr/hal-01876973

National Conferences with Proceeding

[28] E. CARON, Z. HOUMANI.*Architecture Microservices pilotée par les données*, in "CAL 2018 - Conférence francophone sur les Architectures Logicielles", Grenoble, France, June 2018, p. 1-6, https://hal.inria.fr/hal-01808758

Conferences without Proceedings

[29] M. CHARDET, H. COULLON, C. PÉREZ.*Behavioral interfaces for reconfiguration of component models*, in "Compas 2018 - Conférence d'informatique en Parallélisme, Architecture et Système", Toulouse, France, July 2018, p. 1-8, https://hal.inria.fr/hal-01897803

[30] A. CHEVALIER.*Vers un déploiement conforme et économique des licences dans une architecture Cloud*, in "COMPAS 2018", Toulouse, France, July 2018, https://hal.archives-ouvertes.fr/hal-01905417

[31] T. PELLISSIER TANON, M. DIAS DE ASSUNCAO, E. CARON, F. M. SUCHANEK.*Demoing Platypus – A Multilingual Question Answering Platform for Wikidata*, in "ESWC 2018 - Extended Semantic Web Conference", Heracklion, Greece, June 2018, p. 1-5 [*DOI :* 10.1007/978-3-319-98192-5_21], https://hal.archives-ouvertes.fr/hal-01824972

Research Reports

[32] T. P. TANON, M. DIAS DE ASSUNCAO, E. CARON, F. SUCHANEK.*Platypus – A Multilingual Question Answering Platform for Wikidata*, LIP - ENS Lyon, 2018, https://hal.archives-ouvertes.fr/hal-01730479

Scientific Popularization

[33] G. AMARASINGHE, M. DIAS DE ASSUNCAO, A. HARWOOD, S. KARUNASEKERA.*A Data Stream Processing Optimisation Framework for Edge Computing Applications*, in "ISORC 2018 - IEEE 21st International Symposium on Real-Time Distributed Computing", Singapore, Singapore, IEEE, May 2018, p. 91-98 [*DOI :* 10.1109/ISORC.2018.00020], https://hal.inria.fr/hal-01862063

[34] L. L. NESI, M. PILLON, M. DIAS DE ASSUNCAO, G. P. KOSLOVSKI.*GPU-Accelerated Algorithms for Allocating Virtual Infrastructure in Cloud Data Centers*, in "CCGRID 2018 - 18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing", Washington, France, IEEE, May 2018, p. 364-365 [*DOI :* 10.1109/CCGRID.2018.00057], https://hal.inria.fr/hal-01862065

[35] A. DA SILVA VEITH, M. DIAS DE ASSUNCAO, L. LEFÈVRE.*Latency-Aware Strategies for Placing Data Stream Analytics onto Edge Computing*, July 2018, 1, HotEdge'18 - USENIX Workshop on Hot Topics in Edge Computing, Poster, https://hal.inria.fr/hal-01875999

References in notes

[36] R. GE, X. FENG, S. SONG, H.-C. CHANG, D. LI, K. W. CAMERON.*PowerPack: Energy Profiling and Analysis of High-Performance Systems and Applications*, in "IEEE Trans. Parallel Distrib. Syst.", May 2010, vol. 21, n^{o} 5, p. 658–671 [*DOI :* 10.1109/TPDS.2009.76], http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4906989

[37] A. GEIST, S. DOSANJH.*IESP Exascale Challenge: Co-Design of Architectures and Algorithms*, in "Int. J. High Perform. Comput. Appl.", November 2009, vol. 23, n^{o} 4, p. 401–402, http://dx.doi.org/10.1177/1094342009347766

[38] W. GROPP, S. HUSS-LEDERMAN, A. LUMSDAINE, E. LUSK, B. NITZBERG, W. SAPHIR, M. SNIR.*MPI: The Complete Reference – The MPI-2 Extensions*, 2, The MIT Press, September 1998, vol. 2, ISBN 0-262-57123-4

[39] H. KIMURA, T. IMADA, M. SATO.*Runtime Energy Adaptation with Low-Impact Instrumented Code in a Power-Scalable Cluster System*, in "Proceedings of the 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing", Washington, DC, USA, CCGRID '10, IEEE Computer Society, 2010, p. 378–387

[40] G. MADEC.*NEMO ocean engine*, Institut Pierre-Simon Laplace (IPSL), France, 2008, n^{o} 27, ISSN No 1288-1619

[41] OPENACC.*The OpenACC Application Programming Interface*, November 2011, Version 1.0, http://www.openacc-standard.org

[42] OPENMP ARCHITECTURE REVIEW BOARD.*OpenMP Application Program Interface*, July 2011, Version 3.1, http://www.openmp.org

[43] T. RISTENPART, E. TROMER, H. SHACHAM, S. SAVAGE.*Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds*, in "Proceedings of the 16th ACM conference on Computer and communications security", New York, NY, USA, CCS '09, ACM, 2009, p. 199–212, http://doi.acm.org/10.1145/1653662.1653687

[44] B. ROUNTREE, D. K. LOWNENTHAL, B. R. DE SUPINSKI, M. SCHULZ, V. W. FREEH, T. BLETSCH.*Adagio: Making DVS Practical for Complex HPC Applications*, in "Proceedings of the 23rd international conference on Supercomputing", New York, NY, USA, ICS '09, ACM, 2009, p. 460–469

[45] C. SZYPERSKI.*Component Software - Beyond Object-Oriented Programming*, 2, Addison-Wesley / ACM Press, 2002, 608

[46] S. VALCKE.*The OASIS3 coupler: a European climate modelling community software*, in "Geoscientific Model Development", 2013, vol. 6, p. 373-388, doi:10.5194/gmd-6-373-2013

Project-Team BEAGLE

Artificial Evolution and Computational Biology

IN COLLABORATION WITH: Laboratoire d'InfoRmatique en Image et Systèmes d'information (LIRIS)

IN PARTNERSHIP WITH:
CNRS

Institut national des sciences appliquées de Lyon

Université Claude Bernard (Lyon 1)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Computational Biology

Table of contents

Project-Team BEAGLE

Creation of the Team: 2011 June 17, updated into Project-Team: 2013 January 01

Keywords:

Computer Science and Digital Science:
A3.3.2. - Data mining
A5.1.5. - Body-based interfaces
A5.7.2. - Music
A5.11.1. - Human activity analysis and recognition
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A6.1.4. - Multiscale modeling
A6.2.7. - High performance computing
A8.1. - Discrete mathematics, combinatorics

Other Research Topics and Application Domains:
B1. - Life sciences
B1.1.2. - Molecular and cellular biology
B1.1.6. - Evolutionnary biology
B1.1.7. - Bioinformatics
B1.1.10. - Systems and synthetic biology
B1.2.1. - Understanding and simulation of the brain and the nervous system
B9.2.1. - Music, sound
B9.2.4. - Theater

1. Team, Visitors, External Collaborators

Research Scientists
Hugues Berry [Inria, Senior Researcher, HDR]
Anton Crombach [Inria, Researcher, from Mar 2018]
Eric Tannier [Inria, Researcher, HDR]

Faculty Members
Guillaume Beslon [Team leader, INSA Lyon, Professor, HDR]
Carole Knibbe [INSA, Associate Professor]
Christophe Rigotti [INSA Lyon, Associate Professor, HDR]
Jonathan Rouzaud-Cornabas [INSA Lyon, Associate Professor]

External Collaborator
Hedi Soula [Univ Pierre et Marie Curie]

Technical Staff
Nicolas Comte [Inria, until Sep 2018]
David Parsons [Inria, from Apr 2018]

PhD Students
Audrey Denizot [INSA Lyon]
Marie Fernandez [Inria, until Jun 2018]
Alexandre Foncelle [Inria, until Mar 2018]

Vincent Liard [INSA Lyon]

Post-Doctoral Fellow

Priscila Do Nascimento Biller [Inria, until Feb 2018]

Administrative Assistants

Florence Maillard [Inria, from Feb 2018]

Gaelle Tworkowski [Inria, until Feb 2018]

2. Overall Objectives

2.1. An interface between biology and computer science

The expanded name for the BEAGLE research group is “Artificial Evolution and Computational Biology”. Our aim is to position our research at the interface between biology and computer science and to contribute new results in biology by modeling biological systems. In other words we are making artifacts – from the Latin *artis factum* (an entity made by human art rather than by Nature) – and we explore them in order to understand Nature. The team is an Inria Project-Team since January, 2014. It gathers researchers from Inria, INSA, UCBL, who are members of three different labs, the LIRIS [0], the LBBE [0], and CARMEN [0]. It is led by Prof. Guillaume Beslon (INSA-Lyon, LIRIS, Computer Science Dept.).

Our research program requires the team members to have skills in computer science but also in life sciences: they must have or develop a strong knowledge in biosciences to interact efficiently with biologists or, ideally, to directly interpret the results given by the models they develop. A direct consequence of this claim is that it is mandatory to restrict the domain of expertise in life sciences. This is why we focus on a specific scale, central in biology: the cellular scale. Indeed, we restrict our investigations on the cell, viewed as a dynamical system made of molecular elements. This specific scale is rich in open questions that deserve modeling and simulation approaches. We also focus on two different kinds of constraints that structure the cellular level: biophysical constraints and historical constraints. The cell is a system composed of molecules that physically interact and the spatio-temporal nature of these interactions is likely to strongly influence its dynamics. But the cell is also the result of an evolutionary process that imposes its own limits on what can evolve (or is the most likely to evolve) and what cannot (or is the less likely to evolve). A better understanding of what kind of systems evolution is the most likely to lead to in a given context could give us important clues for the analysis of extant biological systems.

2.2. An organization into two tools and four main axes

To study these two kinds of constraints we mainly rely on two specific tools: computational cellular biochemistry and evolution models. We use these tools to develop our “artifacts” and we compare their output with real data, either direct measurements collected by experimentalists or ancestral properties computationally inferred from their extant descendants. The team research is currently organized in four main research axes. The first two ones are methodologically-oriented: we develop general formalisms and tools for computational cellular biochemistry (research axis 1) and families of models to study the evolutionary process (research axis 2). The third “NeuroCell” axis (research axis 3) is the one in which biochemical models are specifically applied on brain cells (neurons and glia). Eventually the last axis aims at integrating the two tools, computational biochemistry and evolution, in what we call "Evolutionary Systems Biology" (research axis 4). The next four sections describe these four axes in more details. The biological questions described are not the sole topics tackled by the team. They are the ones that mobilize a substantial fraction of the researchers on the long run.

[0] Laboratoire d’Informatique en Image et Systèmes d’Information: UMR 5205 CNRS, INSA-Lyon, Univ. Claude Bernard Lyon 1, Univ. Louis Lumière Lyon 2, École Centrale de Lyon

[0] Laboratoire de Biometrie et Biologie Evolutive: UMR CNRS 5558, Univ. Claude Bernard Lyon 1.

[0] Laboratoire de Recherche en Cardiovasculaire, Métabolisme, Diabétologie et Nutrition: UMR U1060 INSERM, INSA-Lyon, INRA 1235, Univ. Claude Bernard Lyon 1.

Many other questions are tackled by individual researchers or even small groups. In the following these ones will be briefly described in their methodological context, *i.e.* in the two sections devoted to research axes 1 and 2.

2.3. A strategy

The scientific objective of the BEAGLE team is to develop a consistent set of concepts and tools – mainly based on computational science – to *in fine* contribute to knowledge discovery in systems biology. Our strategy is to develop strong interactions with life science researchers to become active partners of the biological discovery process. Thus, our aim as a team is not to be a computer science team interacting with biologists, nor to be a team of biologists using computer science tools, but rather to stay in the middle and to become a *trading zone* [34] between biology and computer science. Our very scientific identity is thus fuzzy, melting components from both sciences. Indeed, one of the central claims of the team is that interdisciplinarity involves permanent exchanges between the disciplines. Such exchanges can hardly be maintained between distant teams. That's why the BEAGLE team tries to develop local collaborations with local scientists. That's also why BEAGLE also tries to organize itself as an intrinsically interdisciplinary group, gathering different sensitivities between biology and computer science inside the group. Our ultimate objective is to develop interdisciplinarity at the individual level, all members of the team being able to interact efficiently with specialists from both fields.

3. Research Program

3.1. Introduction

As stated above, the research topics of the BEAGLE Team are centered on the modelization and simulation of cellular processes. More specifically, we focus on two specific processes that govern cell dynamics and behavior: Biophysics and Evolution. We are strongly engaged into the integration of these level of biological understanding.

3.2. Research axis 1: Computational cellular biochemistry

Biochemical kinetics developed as an extension of chemical kinetics in the early 20th century and inherited the main hypotheses underlying Van't Hoff's law of mass action : a perfectly-stirred homogeneous medium with deterministic kinetics. This classical view is however challenged by recent experimental results regarding both the movement and the metabolic fate of biomolecules. First, it is now known that the diffusive motion of many proteins in cellular media exhibits deviations from the ideal case of Brownian motion, in the form of position-dependent diffusion or anomalous diffusion, a hallmark of poorly mixing media. Second, several lines of evidence indicate that the metabolic fate of molecules in the organism not only depends on their chemical nature, but also on their spatial organisation – for example, the fate of dietary lipids depends on whether they are organized into many small or a few large droplets (see e.g. [36]). In this modern-day framework, cellular media appear as heterogeneous collections of contiguous spatial domains with different characteristics, thus providing spatial organization of the reactants. Moreover, the number of implicated reactants is often small enough that stochasticity cannot be ignored. To improve our understanding of intracellular biochemistry, we study spatiotemporal biochemical kinetics using computer simulations (particle-based spatially explicit stochastic simulations) and mathematical models (age-structured PDEs).

3.3. Research axis 2: Models for Molecular Evolution

We study the processes of genome evolution, with a focus on large-scale genomic events (rearrangements, duplications, transfers). We are interested in deciphering general laws which explain the organization of the genomes we observe today, as well as using the knowledge of these processes to reconstruct some aspects of the history of life. To do so, we construct mathematical models and apply them either in a "forward" way, *i.e.* observing the course of evolution from known ancestors and parameters, by simulation (*in silico experimental* *evolution*) or mathematical analysis (*theoretical biology*), or in a “backward” way, *i.e.* reconstructing ancestral states and parameters from known extant states (*phylogeny, comparative genomics*). Moreover we often mix the two approaches either by validating backwards reconstruction methods on forward simulations, or by using the forward method to test evolutionary hypotheses on biological data.

3.4. Research axis 3: Computational systems biology of neurons and astrocytes

Brain cells are rarely considered by computational systems biologists, though they are especially well suited for the field: their major signaling pathways are well characterized, the cellular properties they support are well identified (e.g. synaptic plasticity) and eventually give rise to well known functions at the organ scale (learning, memory). Moreover, electro-physiology measurements provide us with an experimental monitoring of signaling at the single cell level (sometimes at the sub-cellular scale) with unrivaled temporal resolution (milliseconds) over durations up to an hour. In this research axis, we develop modeling approaches for systems biology of both neuronal cells and glial cells, in particular astrocytes. We are mostly interested in understanding how the pathways implicated in the signaling between neurons, astrocytes and neurons-astrocytes interactions implement and regulate synaptic plasticity.

3.5. Research axis 4: Evolutionary Systems Biology

This axis, consisting in integrating the two main biological levels we study, is a long-standing and long-term objective in the team. These last years we did not make significant advances in this direction and we even removed this objective from last year’s report. However the evolution of the team staff and projects allows us to give it back its central place. We now have the forces and ideas to progress. We have several short and middle term projects to integrate biochemical data and evolution. In particular we are analysing with an evolutionary perspective the 3D conformation of chromosomes, the regulatory landscape of genomes, the chromatin-associated proteins.

4. Application Domains

4.1. Domain 1

Applications concern Functional and Evolutionary Biology, plant, animal and human health. They are not detailed here because the project itself is oriented by its applications, so the description of applications is described along the project in the previous sections.

5. Highlights of the Year

5.1. Highlights of the Year

We had several remarkable publications in 2018, including 3 in the highest standard journals and 2 best paper awards.

Cui Y, Yang Y, Ni Z, Dong Y, Cai G, Foncelle A, Ma S, Sang K, Tang S, Li Y, Shen Y, Berry H, Wu S and Hu H (2018). Astroglial-Kir4.1 in Lateral Habenula Drives Neuronal Bursts to Mediate Depression. *Nature* 554:323-327 [15]

Davin AA, Tannier E, Williams TA, Boussau B, Daubin V, Szollosi GJ (2018) Gene transfers can date the tree of life, *Nature ecology and evolution*, vol. 2 pp.904-909. [16]

Berta Verd, Erik Clark, Karl R Wotton, Hilde Janssens, Eva Jiménez-Guri, Anton Crombach, Johannes Jaeger (2018) A damped oscillator imposes temporal order on posterior gap gene expression in Drosophila *PLoS biology* 16 (2), e2003174 [35]

5.1.1. Awards

BEST PAPERS AWARDS :

[28]

V. F. LIARD, D. P. PARSONS, J. ROUZAUD-CORNABAS, G. BESLON.*The Complexity Ratchet: Stronger than selection, weaker than robustness*, in "ALIFE 2018 - the 2018 conference on artificial Life", Tokyo, Japan, July 2018, p. 1-8 [*DOI :* 10.1162/ISAL_A_00051], https://hal.archives-ouvertes.fr/hal-01882628

[26]

S. PEIGNIER, C. RIGOTTI, A. ROSSI, G. BESLON.*Weight-based search to find clusters around medians in subspaces*, in "SAC 2018 - ACM Symposium On Applied Computing", Pau, France, Proceedings of the 33rd ACM Symposium On Applied Computing, April 2018, p. 1-10, https://hal.archives-ouvertes.fr/hal-01869974

6. New Software and Platforms

6.1. aevol

Artificial Evolution

KEYWORDS: Bioinformatics - Genomics - Evolution

FUNCTIONAL DESCRIPTION: Aevol is a digital genetics model: populations of digital organisms are subjected to a process of selection and variation, which creates a Darwinian dynamics. By modifying the characteristics of selection (e.g. population size, type of environment, environmental variations) or variation (e.g. mutation rates, chromosomal rearrangement rates, types of rearrangements, horizontal transfer), one can study experimentally the impact of these parameters on the structure of the evolved organisms. In particular, since Aevol integrates a precise and realistic model of the genome, it allows for the study of structural variations of the genome (e.g. number of genes, synteny, proportion of coding sequences).

The simulation platform comes along with a set of tools for analysing phylogenies and measuring many characteristics of the organisms and populations along evolution.

An extension of the model (R-Aevol), integrates an explicit model of the regulation of gene expression, thus allowing for the study of the evolution of gene regulation networks.

RELEASE FUNCTIONAL DESCRIPTION: Fix compilation error on Mac (tr1 included in std). The new mac compiler includes the tr1 directly in std which caused a compilation error. This issue was specific to aevol-4.4.1

- Participants: Antoine Frénoy, Bérénice Batut, Carole Knibbe, David Parsons, Dusan Misevic, Guillaume Beslon, Jonathan Rouzaud-Cornabas and Vincent Liard
- Partners: UCBL Lyon 1 - INSERM - Université Paris-Descartes - Insa de Lyon
- Contact: Guillaume Beslon
- URL: http://www.aevol.fr/

6.2. DeCoSTAR

KEYWORDS: Bioinformatics - Evolution

FUNCTIONAL DESCRIPTION: DeCoSTAR reconstructs ancestral genomes and improves the assembly of extant genomes. It takes as input a set of gene trees, a species tree and adjacency relations between extant genes. It outputs ancestral genes, adjacencies between extant and ancestral genes, and a statistical support associated to each inferred adjacency.

NEWS OF THE YEAR: Publication of the software with several test sets in Genome Biology and Evolution

- Participants: Eric Tannier and Wandrille Duchemin
- Contact: Eric Tannier
- Publication: hal-01503766
- URL: http://pbil.univ-lyon1.fr/software/DeCoSTAR/

6.3. Evo2Sim

Evolution of Evolution Simulator
KEYWORDS: Bioinformatics - Biology - Evolution
FUNCTIONAL DESCRIPTION: In the context of the EvoEvo european project we developed an integrated model of microorganisms evolution. This model extends the evolutionary models developped in the team (Aevol and R-Aevol) by adding a metabolic level and an ecosystem level. It includes the genomic, genetic and metabolic levels.

- Participants: Carole Knibbe, Charles Rocabert and Guillaume Beslon
- Contact: Guillaume Beslon
- URL: http://www.evoevo.eu/

6.4. evowave

KEYWORDS: Data stream - Clustering - Evolution - Wireless network
FUNCTIONAL DESCRIPTION: This package is a toolbox to analyse signal strength in wifi activity logfiles. It includes three main modules. The first is a preprocessing module to agregate logfile contents. The second one is a subspace clustering module, based on an evolutionary algorithm, to identify similar wifi activity contexts. This similarity is defined on signal strength of wifi devices and the clusters can change over time. The third module is a visualisation tool to display the cluster modifications over time.

- Participants: Anthony Rossi, Christophe Rigotti, Guillaume Beslon, Jonas Abernot, Leo Lefebvre and Sergio Peignier
- Contact: Christophe Rigotti
- URL: http://evoevo.liris.cnrs.fr/download/4_-_deliverables/wp5/Deliverable_D5.1_software_archive.zip

6.5. FluoBacTracker

KEYWORDS: Bioinformatics - Biology - Biomedical imaging
SCIENTIFIC DESCRIPTION: FluoBacTracker is an ImageJ plugin allowing the segementation and tracking of growing bacterial cells from time-lapse microscopy movies. The segmentation and tracking algorithms used by FluoBacTracker have been developed by Lionel Moisan and colleagues at Université Paris Descartes.
FUNCTIONAL DESCRIPTION: FluoBacTracker has the following functionalities: 1) Select regions of interest in images of microcolonies 2) Denoise and renormalize the images 3) Identify each cells in each image (segmentation) 4) Follow cells through the whole movie (tracking), including the detection of cells washed out from a microfluidics channel 5) Detect divisions and construct cell lineage of the population
NEWS OF THE YEAR: Version 2 of FluoBacTracker also allows the analysis of microscopy of bacteria growing in a microfluidics device called "mother machine".

- Participants: Hugues Berry, Cyril Dutrieux, Hidde De Jong, Charles Kervrann, David Parsons and Magali Vangkeosay
- Partners: Université Descartes - UGA
- Contact: Hugues Berry
- URL: http://fluobactracker.inrialpes.fr

6.6. Tewep

Simulator of the dynamics of Transposable Elements Within Expanding Populations
KEYWORDS: Simulator - Transposable elements - Population genetics - Geographic expansion

FUNCTIONAL DESCRIPTION: Transposable elements, found in the genomes of most living organisms (including humans), are pieces of DNA able to replicate themselves and to proliferate. Their presence is a source of mutations which are, most of the time, detrimental to their host. As a consequence, natural selection usually limits their spread. There are, however, some conditions where natural selection cannot be efficient enough to remove them, for example when the population size is small. It is also hypothesized that when a population geographically expands, the efficiency of natural selection could be reduced at the expansion front. TEWEP is an individual-based simulator designed to test whether transposable elements could proliferate in large expanding populations. It combines several population genetics models to simulate the evolution of the number of transposable elements in each individual of an expanding population.

- Partner: Laboratoire de Biométrie et Biologie Evolutive (LBBE) - UMR CNRS 5558
- Contact: Carole Knibbe
- URL: https://gforge.inria.fr/projects/tewep/

6.7. Treerecs

KEYWORDS: Bioinformatics - Biology - Computational biology
SCIENTIFIC DESCRIPTION: The reconciliation between gene trees and species trees is a modern method of molecular phylogeny, which does not yet have its standard software, as for example phylogeny from DNA or amino acid sequences. Treerecs has this ambition, incorporating the classic functionalities of reconciliation: annotating the vertices of a gene tree with the tops of a species tree, rooting and correcting the gene tree. Rooting and correction are calculated to minimize the number of duplications and losses in reconciliation. Medium-sized solutions are randomly sampled according to a uniform law. A likelihood can then be calculated using probabilistic methods. In addition, Treerecs is integrated into a standard software ecosystem of phylogeny, bio ++, ALE, Seaview, and has a graphical interface. Some original features are implemented, such as the possibility of combining two types of likelihoods, the one calculated from the sequences and the one calculated from the reconciliations, the possibility of estimating the costs of the evolutionary events, the possibility of exploring the space of trees according to a joined likelihood.
FUNCTIONAL DESCRIPTION: Treerecs takes as minimum input a gene tree and a species tree. It "reconciles" them, that is, it annotates gene tree nodes with events and assign them to species tree nodes. Biologically, it is a reconstruction of the gene history, given the species history, in terms of duplications, speciations, losses.

With the appropriate options Treerecs can root and correct the gene tree.
NEWS OF THE YEAR: Release of a 0.1 stable version

- Participants: Nicolas Comte, David Parsons, Eric Tannier and Benoît Morel
- Partner: Laboratoire de Biométrie et Biologie Evolutive (LBBE) - UMR CNRS 5558
- Contact: Eric Tannier

6.8. EvoMove

KEYWORDS: Music - Improvisation - Clustering - Evolution - Evolutionary Algorithms
FUNCTIONAL DESCRIPTION: EvoMove uses data from Inertial Measurement Units carried by dancers. It classifies these data in a non-supervised way to recognise "moves" and from these triggers music samples that accompany the dancers.

- Contact: Guillaume Beslon

7. New Results

7.1. Dopamine interacts with endocannabinoids to regulate spike timing dependent plasticity

participants: H. Berry, I. Prokin

Dopamine modulates striatal synaptic plasticity, a key substrate for action selection and procedural learning. Thus, characterizing the repertoire of activity-dependent plasticity in striatum and its dependence on dopamine is of crucial importance. In collaboration with L. Venance Lab (CIRB, Collège de France) we recently unraveled a striatal spike-timing-dependent long-term potentiation (tLTP) mediated by endocannabinoids (eCBs) and induced with few spikes (5-15). Whether this eCB-tLTP interacts with the dopaminergic system remains to be investigated. We found that eCB-tLTP is impaired in a rodent model of Parkinson's disease and rescued by L-DOPA. Dopamine controls eCB-tLTP via dopamine type-2 receptors (D2R) located presynaptically in cortical terminals. Dopamine-endocannabinoid interactions via D2R are required for the emergence of tLTP in response to few coincident pre-and post-synaptic spikes and control eCB-plasticity by modulating the long-term potentiation (LTP)/depression (LTD) thresholds. While usually considered as a depressing synaptic function, our results show that eCBs in the presence of dopamine constitute a versatile system underlying bidirectional plasticity implicated in basal ganglia pathophysiology. These results have been published in Nature Communications [23]

7.2. Estimating the robustness of spike timing dependent plasticity to timing jitter

participants: H. Berry, I. Prokin

In Hebbian plasticity, neural circuits adjust their synaptic weights depending on patterned firing. Spike-timing-dependent plasticity (STDP), a synaptic Hebbian learning rule, relies on the order and timing of the paired activities in pre- and postsynaptic neurons. Classically, in ex vivo experiments, STDP is assessed with deterministic (constant) spike timings and time intervals between successive pairings, thus exhibiting a regularity that differs from biological variability. Hence, STDP emergence from noisy inputs as occurring in in vivo-like firing remains unresolved. In collaboration with the laboratories of L. Venance (CIRB, Collège de France) and A. De Kerchove d'Exaerde (Univ. Libre Bruxelles), we used noisy STDP pairings where the spike timing and/or interval between pairings were jittered. We explored with electrophysiology and mathematical modeling, the impact of jitter on three forms of STDP at corticostriatal synapses: NMDAR-LTP, endocannabinoid-LTD and endocannabinoid-LTP. We found that NMDAR-LTP was highly fragile to jitter, whereas endocannabinoid-plasticity appeared more resistant. When the frequency or number of pairings was increased, NMDAR-LTP became more robust and could be expressed despite strong jittering. Our results identify endocannabinoid-plasticity as a robust form of STDP, whereas the sensitivity to jitter of NMDAR-LTP varies with activity frequency. This provides new insights into the mechanisms at play during the different phases of learning and memory and the emergence of Hebbian plasticity in in vivo-like activity. These results have been published in Scientific Reports [14]

7.3. A new method to monitor gap junctional comunication in astrocytes

participants: H. Berry

Intercellular communication through gap junction channels plays a key role in cellular homeostasis and in synchronizing physiological functions, a feature that is modified in number of pathological situations. In the brain, astrocytes are the cell population that expresses the highest amount of gap junction proteins, named connexins. Several techniques have been used to assess the level of gap junctional communication in astrocytes, but so far they remain very difficult to apply in adult brain tissue. Using specific loading of astrocytes with sulforhodamine 101, we adapted in collaboration with C. Giaume's laboratory (CIRB, Collège de France) the gap-FRAP (Fluorescence Recovery After Photobleaching) to acute hippocampal slices from 9 month-old adult mice. We show that gap junctional communication monitored in astrocytes with this technique was inhibited either by pharmacological treatment with a gap junctional blocker or in mice lacking the two main astroglial connexins, while a partial inhibition was measured when only one connexin was knocked-out. We validate this approach using a mathematical model of sulforhodamine 101 diffusion in an elementary astroglial network and a quantitative analysis of the exponential fits to the fluorescence recovery curves. Consequently, we consider that the adaptation of the gap-FRAP technique to acute brain slices from adult mice provides an easy going and valuable approach that allows overpassing this age-dependent obstacle and will facilitate the investigation of gap junctional communication in adult healthy or pathological brain. These results have been published in J. Neuroscience Methods [24].

7.4. Kir4.1 upregulation in astrocytes of the lateral habenula is involved in depression

participants: H. Berry, A. Foncelle

Enhanced bursting activity of neurons in the lateral habenula (LHb) is essential in driving depression-like behaviours, but the cause of this increase has been unknown. In collaboration with H. Hu's laboratory (Zhejiang University, China), using a high-throughput quantitative proteomic screen, we show that an astroglial potassium channel (Kir4.1) is upregulated in the LHb in rat models of depression. Kir4.1 in the LHb shows a distinct pattern of expression on astrocytic membrane processes that wrap tightly around the neuronal soma. Electrophysiology and modelling data show that the level of Kir4.1 on astrocytes tightly regulates the degree of membrane hyperpolarization and the amount of bursting activity of LHb neurons. Astrocyte-specific gain and loss of Kir4.1 in the LHb bidirectionally regulates neuronal bursting and depression-like symptoms. Together, these results show that a glia–neuron interaction at the perisomatic space of LHb is involved in setting the neuronal firing mode in models of a major psychiatric disease. Kir4.1 in the LHb might have potential as a target for treating clinical depression. These results have been published in Nature [15] and were commented in the "News and views" section of the journal: Howe WM and Kenny PJ (2018). Burst firing sets the stage for depression.

7.5. The evolutionary complexity ratchet

participants: G Beslon, V Liard, D Parsons, Jonathan Rouzaud-Cornabas

Using the *in silico* experimental evolution platform Aevol, we evolved populations of digital organisms in conditions where a simple functional structure is best.

Strikingly, we observed that in a large fraction of the simulations, organisms evolved a complex functional structure and that their complexity increased during evolution despite being a lot less fit than simple organisms in other populations. However, when submitted to a harsh mutational pressure, we observed that a significant proportion of complex individuals ended up with a simple functional structure.

Our results suggest the existence of a complexity ratchet that is powered by epistasis and that cannot be beaten by selection. They also show that this ratchet can be overthrown by robustness because of the strong constraints it imposes on the coding capacity of the genome.

This result has been published in the International conference ALife in Tokyo (July 2018) where it received the best paper award [28]

7.6. Weight-based search to find clusters around medians in subspaces

participants: C Rigotti, G Beslon

There exist several clustering paradigms, leading to different techniques that are complementary in the analyst toolbox, each having its own merits and interests. Among these techniques, the K-medians approach is recognized as being robust to noise and outliers, and is an important optimization task with many different applications (e.g., facility location). In the context of subspace clustering, several paradigms have been investigated (e.g., centroid-based, cell-based), while the median-based approach has received less attention. Moreover, using standard subspace clustering outputs (e.g., centroids, medoids) there is no straightforward procedure to compute the cluster membership that optimizes the dispersion around medians. We advocated for the use of median-based subspace clustering as a complementary tool. Indeed, we showed that such an approach exhibits satisfactory quality clusters when compared to well-established paradigms, while medians have still their own interests depending on the user application (robustness to noise/outliers and location optimality). We showed that a weight-based hill climbing algorithm using a stochastic local exploration step can be sufficient to produce the clusters.

This research has been published in the proceedings of the ACM-SAC conference (Pau, March 2018) where it received the best paper award [26].

7.7. The surprising creativity of digital evolution

participants: C Knibbe, G Beslon

Natural evolution is a creative fount of complex adaptations that often surprise the scientists who discover them. However, the creativity of evolution is not limited to the natural world; artificial organisms evolving in computational environments are also able to elicit a similar degree of surprise and wonder from the researchers studying them. The process of evolution has proven to be an algorithmic process that transcends the substrate to which it is applied. Indeed, most digital evolution researchers can relate anecdotes highlighting how common it is for their algorithms to creatively subvert their expectations or intentions, expose unrecognized bugs in their code, produce unexpectedly potent adaptations, or engage in behaviors and outcomes uncannily convergent with ones found in nature. Such stories routinely reveal surprise and creativity by evolution in these digital worlds, but they rarely fit into the standard scientific narrative and are thus often treated as obstacles to be overcome rather than results that warrant publication in their own right. Bugs are fixed, experiments are refocused, one-off surprises are collapsed 1into a single data point. The stories themselves are traded among researchers through oral tradition, but that mode of information transmission is lossy, inefficient and error-prone. Moreover, the very fact that these stories tend to be confined to practitioners means that many natural scientists do not recognize how lifelike digital organisms are and how natural their evolution can be. We actively participated to a crowd-sourced research in which evolutionary computation researchers providing first-hand reports of such cases, and thus functions as a written, fact-checked collection of entertaining and important stories.

7.8. HPC support for Aevol

participants: Jonathan Rouzaud-Cornabas, David Parsons, Guillaume Beslon

During the year, we had three internships that focus around HPC. The three of them were founded through the Federation Informatique de Lyon (FIL FR2000) and were common between the Inria Beagle team (LIRIS) and the Inria Avalon team (LIP).

The first one (Lukas Schmidt - M2) was working on component-based software engineering and HPC with Aevol as use-case. The goal was to see if and how the COMET [1] task-based parallel component model (and its implementation Halley) can fit the parallelization requirement of Aevol. An extension of the model was proposed to support hierarchical data structure and a prototype implementation has been done. In the future, we will work on the formalization of the extension and an efficient implementation on it. The goal is to ease the development and replacement of core components of the Aevol software (e.g. be able to easily replace the 2-base DNA code by a 4-base one).

The second internship (Valentin Huguet - M2) was evolving around Aevol and how to ease the distribution of the computation. To do so, an extension of the DIET software [2] was proposed and a fully functional webboard was implemented. We have a first prototype that support the execution of a large set of distributed computing resources and the control of its execution through a webboard. Moreover, basic visualization of the simulation results can be done through the same webboard. A following internship (starting Feb. 2019) will continue the work. The goal is to support workflow composed of multiple execution of Aevol and its pre/post treatments to automate the execution of large campaign that are done manually at the moment.

The goal of the third internship (Nathan Payre - L3) was to propose a prototype of a bitset for Aevol and its efficient implementation on modern hardware (Intel Skylake and Intel Xeon Phi). Indeed, the current implementation of Aevol DNA (2 base) uses a char type (8bit) to store a bit value (0 or 1). Accordingly, working at the bitset level could save up to 8 time memory space and speed up the computation (as Aevol is memory bound, reducing the memory transfer by 8 could dramatically speed up the global execution). Moreover, modern processors have vectorization extension that are perfectly fitting our requirements (we could process 512bit per cycle with AVX512 extension). During the internship, the bitset and the different operation we use in Aevol model (e.g. Hamming distance) were formalized and implemented. The preliminary results show a speed up of 140x on these operations. A full evaluation on the impact of the performance of Aevol and how different modern processor react to such implementation will be done in the future.

Last, a part of the Beagle team (Guillaume Beslon, David Parsons, Jonathan Rouzaud-Cornabas) were selected and participate to the EuroHack 2018 GPU Programming Hackathon in Lugano (Switzerland) organized by CSCS (Swiss National Supercomputing Centre) and NVidia. The goal was to port Aevol to modern GPU and thus to the CUDA programming language. In order to be able to do so in a week, we propose a mini-application (mini-Aevol) of Aevol [3] that is representative of the computation and memory pattern of the full Aevol. This prototype will be reuse in our collaboration with team focusing on HPC research. At the end of the week, we had a full implementation of mini-Aevol on GPU. New core algorithms of Aevol have been proposed to support massively parallel processors such as GPU. The prototype will be transfer to the full Aevol code in the future to be able to support GPU. It is worth noting that this mini-apps is also used in teaching context (INSA Lyon - Computer Science M2) to learn how to parallelize and optimize code with OpenMP and CUDA.

[1] Olivier Aumage, Julien Bigot, Hélène Coullon, Christian Pérez, Jérôme Richard. Combining Both a Component Model and a Task-based Model for HPC Applications: a Feasibility Study on GYSELA. 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)., May 2017, Madrid, Spain.

[2] https://graal.ens-lyon.fr/diet/

[3] https://gitlab.inria.fr/rouzaudc/mini-aevol

[4] https://github.com/fomics/EuroHack18/blob/master/final/beagle_aevol.pdf

7.9. Exploring the evolution of chromatin-associated proteins

participants: A Crombach

Eukaryotic gene regulation depends strongly on chromatin state. High-throughput studies in the fruit fly Drosophila melanogaster have shown that instead of the canonical two types of chromatin, hetero- and euchromatin, one can subdivide chromatin into five states. These states are each characterized by a unique combination of chromatin-associated proteins (CAPs). We were interested in the evolution of CAPs and studied them by means of phylogenomic methods. We found three evolutionary trends. One type of heterochomatin, called GREEN, is specific to centromeres and some of its proteins are found to be under a Red Queen type evolution, where they rapidly accumulate amino acid changes. The second type of heterochromatin, BLUE, is tightly linked to Polycomb Group proteins. These proteins are important regulators in developmental processes and our findings confirm their origin in multicellular organisms. Finally, the two euchromatic types, YELLOW and RED, have strong lineage-specific characteristics. Their origins seem to date back to the start of eukaryotic life.

7.10. Evolutionary interplay of genome content and 3D spatial structure

participants: A Crombach

Genomes are hierarchically folded, which involves transposable elements (TEs). The most prominently observed folding domains are conserved between cell types and across species, yet their building blocks, TEs, are powerful mutagens. This paradox raises the question why we observe evolutionary stable folding domains. Using in silico evolution of polymer genomes, the aim is to elucidate the interplay between mutations and folding structure. We have built the software and are in the process of generating data. First results indicate that due to accessibility in 3D (some parts of the genome are more tightly compacted than others), a positive feedback is created between (1) where mutations happen, (2) how genome content is changed, and (3) how genomes fold in 3D.

7.11. Network inference for mammalian cortex development

participants: A Crombach

The mammalian cortex divides into two major regions, neocortex (NCx) and the structurally simpler allocortex. Whereas NCx is well-characterized, the allocortex is much less studied. Its best known region is the olfactory (piriform, PCx) cortex. The regions have a laminar structure, with distinct neuronal cell types in each of the layers: NCx has 6 layers, PCx has 3. The differentiation of precursor cells into various neuronal cell types determines to which layer these cells will migrate. This process is mostly studied in NCx and depends on the activity of 10–20 developmental genes. In PCx the same genes are used, yet they appear in other combinations and may indicate diverse target layers, sometimes violating rules-of-thumb derived from NCx. Current understanding is rather incomplete with respect to how cortical neurons are specified. We propose that, despite apparent contradictions, a single gene network can explain the development of distinct cortical regions.

In collaboration with Dr. A. Fleischmann at Brown University (USA), we are measuring the expression of genes involved in neurodevelopment at cellular resolution using light-sheet microscopy. These data will form the basis for the inference of a regulatory network describing neuronal differentiation in NCx and PCx. Inference is done by fitting mathematical models of gene regulation to the data using global optimization methods. Currently, we are processing the image data. Moreover, single cell RNA sequencing will allow the study of the temporal dynamics of the expression of these genes and many others. We are completing an in-depth statistical analysis of the resulting genome-wide expression data.

7.12. Gene transfers can date the tree of life

participants: E Tannier
Biodiversity has always been predominantly microbial, and the scarcity of fossils from bacteria, archaea and microbial eukaryotes has prevented a comprehensive dating of the tree of life. Here, we show that patterns of lateral gene transfer deduced from an analysis of modern genomes encode a novel and abundant source of information about the temporal coexistence of lineages throughout the history of life. We use state-of-the-art species tree-aware phylogenetic methods to reconstruct the history of thousands of gene families and demonstrate that dates implied by gene transfers are consistent with estimates from relaxed molecular clocks in Bacteria, Archaea and Eukarya. We present the order of speciations according to lateral gene transfer data calibrated to geological time for three datasets comprising 40 genomes for Cyanobacteria, 60 genomes for Archaea and 60 genomes for Fungi. An inspection of discrepancies between transfers and clocks and a comparison with mammalian fossils show that gene transfer in microbes is potentially as informative for dating the tree of life as the geological record in macroorganisms. [16]

7.13. The devil in the details of evolvability

Participants: E Tannier, P Biller, V Liard, G Beslon
The theory of Evolvability consists in studying the evolution of living organisms as a computational learning process. It defines the possibilities of a population under Darwinian selection, to evolve in a certain direction, in a reasonable amount of time. While its robustness to certain parameters has been theoretically assessed, this theory has not been experimentally tested. We use a standard *in silico* experimental evolution tool to compare some predictions of the theory and the behavior of digital populations designed to resemble biological organisms. We obtain that the evolvability of monotone conjunctions under the uniform distribution of environmental conditions, presented as a major result of the theory, is not reproduced by the experiments. We show that this is due to different mutation algorithms, by a proof of exponential expectation time to target under theoretical conditions closer to the experiments. We examine into detail the choices of mutation algorithms. In the Evolvability theory it is any Turing machine, while much more restricted in the experimental design. This definition allows a wider range of conditions and in a certain way is conform to biological reality, where mutators evolve and can be selected. However it also allows, if it is misused, for the inclusion of oracles that are incompatible with the principles of a Darwinian evolution. Unfortunately these oracles are extensively used in the current evolvability proofs.

8. Partnerships and Cooperations

8.1. Regional Initiatives

- Lipuscale (2018-2019): Hybrid simulation of lipid digestion and absorption, a two-year project funded by the Rhône-Alpes Institute for Complex Systems (IXXI). With Marie-Caroline Michalski (CarMeN, INSERM U1060, INRA U1397) and Samuel Bernard (Institut Camille Jordan and Inria Dracula team). Participant: Carole Knibbe.

8.2. National Initiatives

8.2.1. ANR

Dopaciumcity (2014-2018): Dopamine modulation of calcium influx underlying synaptic plasticity, a 4-year project funded by a grant from the ANR-NSF-NIH Call for French-US Projects in Computational Neuroscience. With L. Venance, College de France, CIRB, CNRS/UMR 7241 - INSERM U1050, Paris, France and K Blackwell, Krasnow Institute of Advanced Studies, George Mason University, Fairfax, VA, USA. Supervisor: L Venance (for France) and K.L. Blackwell (for US). Participants: H Berry, I Prokin, A Foncelle

Dallish (2016-2020): Data Assimilation and Lattice LIght SHeet imaging for endocytosis/exocytosis pathway modeling in the whole cell, Call AAPG ANR 2016. With C. Kervrann (Inria Rennes), J. Salamero (Institute Curie, Paris), B. Laroche (INRA, Jouy-en-Josas). Participants: H. Berry.

Storiz (2018-2020): Horizontal transfers as documents from extinct or unknown species. Call ANR JCJC 2018. Leaded by Damien de Vienne (LBBE, Lyon) Participant: Eric Tannier

LncEvoSys (2017-2019): An evolutionary systems approach to understand long non-coding RNA functionality, Call ANR JCJC 2017. Leaded by Anamaria Necsulea (LBBE, Lyon). Participant: Eric Tannier

8.2.2. Inria

ADT Phylophile (2016-2018). Participants: E Tannier, in collaboration with D Parsons, Inria, V Daubin, B Boussau, CNRS, Université de Lyon 1. This project aims at producing an easy to use software integrating modern algorithmic methods to build gene trees. It has been funded by Inria by a 24 month software engineer.

Naviscope (Inria Project Lab, 2018-2022): image-guided Navigation and VIsualization of large data sets in live cell imaging and microSCOPy. Nowadays, the detection and visualization of important localized events and process in multidimensional and multi-valued images, especially in cell and tissue imaging, is tedious and inefficient. Specialized scientists can miss key events due to complexity of the data and the lack of computer guidance. In Naviscope we develop original and cutting-edge visualization and navigation methods to assist scientists, enabling semi-automatic analysis, manipulation, and investigation of temporal series of multi-valued volumetric images, with a strong focus on live cell imaging and microscopy application domains. We build Naviscope upon the strength of scientific visualization and machine learning methods in order to provide systems capable to assist the scientist to obtain a better understanding of massive amounts of information. Such systems will be able to recognize and highlight the most informative regions of the dataset by reducing the amount of information displayed and guiding the observer attention. Head: C. Kervrann (Serpico), other EPIs: Aviz, Beagle, Hybrid, Morpheme, Mosaic, Parietal, and MaIage (INRA unit).

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners

- Anton Crombach collaborates with Dr. Alexander Fleischmann, who moved this year from CIRB, College de France (Paris), to Brown University, USA.
- Carole Knibbe collaborates with Kirsty Spalding and Peter Arner from Karolinska University Hospital in Stockholm, Sweden.
- Eric Tannier collaborates with Gergerly Szollosi, Eotvos University, Budapest, Cedric Chauve, SFU, Vancouver, Igor Sharakov, from Virginia Tech, Rob Waterhous, Univ Lausanne, Tom Williams, Univ Bristol, ...
- Eric Tannier has leaded the publication of a collaborative paper on a phylogenetic format co-signed by 27 researchers from 10 nationalities.

8.3.2. Participation in Other International Programs

Program: CNRS-Royal society

Project title: Modeling lateral gene transfer on a new bacterial tree

Duration: 2018

Coordinator: Bastien Boussau

Other partners: Eric Tannier (Beagle), LBBE (Lyon), Eotvos University (Budapest), University of Bristol (UK)

Abstract: Bacteria play a major role in human health and in the functioning of all ecosystems. Most of the genomes available in public databases come from Bacteria. However, we understand little of their evolution: both their phylogeny and their timeline of diversification are highly uncertain. This is due in great part to the difficulty of reconstructing events that happened billions of years ago, and also to the fact that individual genes are often transferred across species and therefore have a history that differs from that of the species that contain them. Recently we have proposed novel methods for dealing with the very problems that make reconstructing the bacterial phylogeny challenging. This proposal aims to support a joint project that will reconstruct and date the bacterial phylogeny by combining novel methods (Boussau, Lyon) with existing skills in microbial phylogenetics, genomics and evolution (Williams, Bristol).

8.4. International Research Visitors

8.4.1. Visits to International Teams

Audrey Denizot stayed 4 months in Okinawa, OIST University, Erik de Schutter's team, from june to october 2018

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- Guillaume Beslon co-organized a symposium on “in vivo, in vitro, in silico experimental evolution” during the World Evolution conference. Montpellier, August 2018.

9.1.1.2. Member of the Organizing Committees

- Hugues Berry was a member of the Local Organizing committee of MedInfo2019 (https://medinfo-lyon.org/)
- Eric Tannier was a member of the local Organizing committee of ICGT 2018 (https://projet.liris.cnrs.fr/~icgt2018/)

- Eric Tannier was a member of the local Organizing committee of RECOMB 2018 (http://recomb2018.fr/)
- Anton Crombach is a member of the organizing and scientific committee of the "Advanced Lecture Course on Computational Systems Biology", in Aussois, March 2019. Also,
- Anton Crombach is a member of the organizing committee of "Mathematical Models in Ecology and Evolution", that will take place in Lyon, July 2019.
- Guillaume Beslon served as a member of the Organizing Committee of ICSB (International Conference on Systems Biology), Lyon, October 2018.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Christophe Rigotti was a member of the program committee of the 34rd ACM Symposium On Applied Computing.
- Eric Tannier was a member of the program committee of RECOMB Comparative Genomics
- Eric Tannier was a member of the program committee of ISMB

9.1.2.2. Reviewer

- Eric Tannier reviewed 3 conference articles for Recomb-CG, 3 conference articles for ISMB

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- H. Berry: Associate Editor for PLoS Computational Biology
- Eric Tannier : Recommender for Peer Community in Evolutionary Biology
- Eric Tannier : guest editor for a special issue of Discrete Mathematics and Theoretical Computer Science
- Carole Knibbe and Guillaume Beslon are guest editors for a special issue of the Artificial Life journal, in preparation.

9.1.3.2. Reviewer - Reviewing Activities

- Anton Crombach has reviewed two articles for Nature Communications and a book proposal for Springer-Verlag.
- Eric Tannier has reviewed articles fo Bioinformatics, Bulletin of mathematical biology, PeerJ, Genome biology and evolution, a book chapter for a book on genome evolution,

9.1.4. Invited Talks

- H. Berry - May 2018: Stochastic simulations of calcium signaling in fine astrocytic processes: spatial properties, invited talk at the workshop "Modelisation Stochastique en Biologie", Tours
- H. Berry - Nov 2018: "Modelling the signaling pathways of neuronal plasticity", invited talk at the 10ème journée ITMO Technologies pour la Santé, Strasbourg.
- E Tannier, December 2018: "Gene transfers can date the tree of life" Lille
- Eric Tannier, August 2018: "Treerecs: a fast and easy to use phylogenetic reconciliation software", for the "Software in phylogenomics" workshop, satellite of Evolution 2018, in Montpellier.
- Eric Tannier, February 2018: "Comparative Genomics on artificial life", ALPHY 2018, Montpellier
- Guillaume Beslon gave an invited Keynote at the 10th Symposium on Seach-Based Software Engineering (Montpellier, September 2018)
- Guillaume Beslon gave an invited Keynote at the First International TRANSIT workshop on Cross-disciplinary Research (TWCR 2018). York (UK), April 2018.
- Guillaume Beslon gave an invited conference for the Systems Biology research group, Université Pierre et Marie Curie, Paris, May 2018.

9.1.5. Scientific Expertise

- Guillaume Beslon was a member of the ANR CES45 panel (Mathématique, informatique, automatique, traitement du signal pour répondre aux défis de la biologie et de la santé)

9.1.6. Research and Teaching Administration

- Hugues Berry is Inria's Deputy Scientific Director for the field "Digital Health, Biology and Earth" and has been:
  - Inria's representative on the board of Aviesan
  - Conseiller Scientifique, ITMO Technologies pour la sante
  - Member of the Working Group for the prefiguration of the French Health Data Hub
  - Member of the Comité d'Orientation Scientifique et Stratégique, Institut Français de Bioinformatique (IFB)
  - Memnber of the Comité Directeur of GIS IBISA
  - Member of the Comité des Tutelles for INBS France Life Imaging, FranceBioImaging
  - Member of the steering committee for the Action de Recherche Prospective "Biologie Predictive" of INRA
  - Member of the "Commission des thèses" of the Doctoral School "Info-Maths" (ED 512)
  - Member of the Steering Commitee of GdR IMaBIO (Imagerie et Microscopie pour la BIOlogie, http://imabio-cnrs.fr)
- Eric Tannier is
  - Member of the committee for Open Science, Ministry of Research
  - Member of the Administration Council of Inria
  - Member of the scientific Committee of the Ethics Platform of Université de Lyon
  - Member and founder of the scientific committee of the GTGC (National Working group on Comparative Genomics)
- Guillaume Beslon is a member of the IRD Commission Scientifique Sectorielle 5 (CSS5, Science des données et des modèles). He served as a president of the hiring committee for a position of Chargé de Recherche number 12.
- Christophe Rigotti is an elected member of Insa Scientific board (Conseil Scientifique)
- Carole Knibbe is
  - Member of the Doctoral Studies Committee of Inria Grenoble-Rhone Alpes
  - Head of the Bioinformatics and Modeling master program of INSA-Lyon Since Sept 2017

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master : Audrey Denizot, enzymology and cellular biology, 64h, INSA Lyon

Licence: Christophe Rigotti, Object-Oriented Programming and Graphical User Interfaces, 86h, L2, Department 1er cycle of INSA-Lyon.

Licence: Christophe Rigotti, Simulation of Chemical Reactions, 26h, L2, Department 1er cycle of INSA-Lyon.

Licence: Christophe Rigotti, Numerical Modelling for Engineering, 60h, L2, Department 1er cycle of INSA-Lyon.

Master: Christophe Rigotti, Data Mining, 25h, M1, Bioinformatics and Modeling Department of INSA-Lyon.

Master: Eric Tannier, Bioinformatics, 12h INSA Lyon

Master: Eric Tannier, Bioinformatics, 12h M2 UCBL Lyon

Master: Eric Tannier, Ethique de la recherche, 2h, Lille

Licence: C. Knibbe, Fundamentals of algorithmics and programming, 80 heqTD, L3, Bioinformatics and Modelling program of INSA-Lyon

Licence: C. Knibbe, Architecture of computer systems, 19 heqTD, L3, Bioinformatics and Modelling program of INSA-Lyon

Licence: C. Knibbe, Software development, 32 heqTD, L3, Bioinformatics and Modelling program of INSA-Lyon

Licence: C. Knibbe, HTML/CSS, 4 heqTD, L3, Bioinformatics and Modelling program of INSA-Lyon

Master: C. Knibbe, Careers in bioinformatics and modelling, 20 heqTD, M1, Bioinformatics and Modelling program of INSA-Lyon

License: Jonathan Rouzaud-Cornabas, Object-Oriented Programming, 64h, L3, Computer Science Department, INSA de Lyon

Master: Jonathan Rouzaud-Cornabas, Advanced Operating System Programming, 64h, M1, Computer Science Department, INSA de Lyon

Master: Jonathan Rouzaud-Cornabas, Graphical User Interfaces, 64h, M1, Computer Science Department, INSA de Lyon

Master: Jonathan Rouzaud-Cornabas, High Performance Computing, 84h, M2, Computer Science Department, INSA de Lyon

Master: Jonathan Rouzaud-Cornabas, High Performance Computing, 84h, M2, Bioinformatics and Modeling Department, INSA de Lyon

CNRS GDR Calcul: Jonathan Rouzaud-Cornabas, High Performance Computing, 6h, CNRS

E-learning

Eric Tannier was in the pedagogical team of a Mooc "Ethique de la Recherche", for Ph-D students in France. 4000 participants. He participated to 2 introductory videos (7min).

9.2.2. Supervision

PhD in progress : Vincent Liard, Towards a quantitative digital genetics platform, INSA-Lyon, started Oct 2016, co-supervision: G Beslon, J Rouzaud-Cornabas, C Ofria (Michigan State University, BEACON Center

PhD : Tuan Nguyen, Handling data quality in extraction and selection of evolutions from displacement field time series obtained by satellite imagery, at LISTIC laboratory of University Grenoble Alpes co-supervised by C. Rigotti, defended 10 October 2018.

PhD in progress : Audrey Denizot Simulating calcium signaling in fine astrocytic processes, september 2016, Hugues Berry

PhD in progress : Theo Tricou, Lateral Gene Transfer as a document from extinct and unknown species, 2018-2020 Eric Tannier (50%)

PhD in progress : Alexandre Laverré, cross-influcence between gene regulatory landscape and 1D-3D genome organization 2018-2020, Eric Tannier (40%)

M2 student Alexandre Laverré (Eric Tannier)

M1 student Adrien Stadler (Hugues Berry)

M2 student Elisa Denier (Eric Tannier)

M1 student Yvan Cluet (Anton Crombach)

M1 student Damien Agopian (Carole Knibbe)

9.2.3. Juries

- Hugues Berry- HDR: Samuel Bottani, Univ Denis Diderot, Paris, Nov 2018 (reviewer)
- Hugues Berry- PhD: Yi Cui, Univ Descartes, Paris, Dec 2018 (reviewer)
- Hugues Berry- PhD: Franziska Oschmann, TU Berlin, Oct 2018 (reviewer)
- Hugues Berry- search pannel: two assistant professor positions (Univ. Pierre et Marie Curie, INSA Lyon)
- Anton Crombach - tutor on the thesis committee of Elise Parey, IBENS, ENS Paris.
- Guillaume Beslon served as a member of the defence committee for the HDR of Alexandre Muzy. Nice, October 2018.
- Jonathan Rouzaud-Cornabas - Associate Professor Recruitment - Université de Toulouse

9.3. Popularization

9.3.1. Interventions

- Eric Tannier gave a series of four lectures at the Université Populaire de Lyon, on "science and democracy".
- Eric Tannier participated to a round table on scientific communication for the BMIC days, november 2018.
- Audrey Denizot has participated in the organization and animation of activities for children about research during the event "Dans la blouse d'un chercheur" in ENS de Lyon. More specifically, she worked on a project aiming at explaining to children what microbes are, where they are and if/when you should worry about them and what is immunodeficiency. She made a small movie about it.
- Audrey Denizot is a member of a public outreach association : DéMesures. She is in charge of collaborations with external entities, including laboratories, museums and artists collectives. Website: https://demesures.jimdo.com/ Facebook: https://www.facebook.com/DMesures/ Twitter: https://twitter.com/DMesures Radio podcasts: https://demesures.jimdo.com/radio-1/radio-brume/ Interviews are available on our YouTube channel: https://www.youtube.com/channel/UCIn3CucWk1mLgbzE0noO2Iw
- National events: Audrey Denizot organized several events including Fête de la Science 2017 and 2018 and A Nous de Voir 2018, as well as conducted projects such as radio podcasts on science/society topics, interviews of researchers, projects mixing Science and Art.

9.3.2. Creation of media or tools for science outreach

- In collaboration with Sylvain Charlat (LBBE, Lyon), we developed game to present the effect of random mutations on evolution to young children. The "GreenMice" game has been presented to the general public in the "Maison des Mathématiques et de l'Informatique" during the exhibition "Comme par hasard".

10. Bibliography

Major publications by the team in recent years

[1] P. Biller, L. Guéguen, C. Knibbe, E. Tannier.*Breaking Good: Accounting for Fragility of Genomic Regions in Rearrangement Distance Estimation*, in "Genome Biology and Evolution", 2016, vol. 8, n^{o} 5, p. 1427-1439 [*DOI :* 10.1093/gbe/evw083], https://hal.archives-ouvertes.fr/hal-01334923

[2] A.-S. COQUEL, J.-P. JACOB, M. PRIMET, A. DEMAREZ, M. DIMICCOLI, T. JULOU, L. MOISAN, A. B. LINDNER, H. BERRY.*Localization of protein aggregation in Escherichia coli is governed by diffusion and nucleoid macromolecular crowding effect*, in "PLoS Computational Biology", 2013, vol. 9, n^{o} 4 [*DOI :* 10.1371/JOURNAL.PCBI.100303], http://hal.inria.fr/hal-00798053

[3] Y. CUI, I. PROKIN, H. XU, B. DELORD, S. GENET, L. VENANCE, H. BERRY.*Endocannabinoid dynamics gate spike-timing dependent depression and potentiation*, in "eLife", February 2016, vol. 5, 32, e13185 [*DOI :* 10.7554/ELIFE.13185], https://hal.inria.fr/hal-01279901

[4] W. DUCHEMIN, V. DAUBIN, E. TANNIER.*Reconstruction of an ancestral Yersinia pestis genome and comparison with an ancient sequence*, in "BMC Genomics", 2015, vol. 16, n^{o} Suppl 10, S9 [*DOI :* 10.1007/978-3-319-19048-8_16], https://hal.inria.fr/hal-01179197

[5] S. FISCHER, S. BERNARD, G. BESLON, C. KNIBBE.*A model for genome size evolution*, in "Bulletin of Mathematical Biology", September 2014, vol. 76, n^{o} 9, p. 2249-2291 [*DOI :* 10.1007/S11538-014-9997-8], https://hal.archives-ouvertes.fr/hal-01090964

[6] T. HINDRÉ, C. KNIBBE, G. BESLON, D. SCHNEIDER.*New insights into bacterial adaptation through in vivo and in silico experimental evolution*, in "Nature Reviews Microbiology", 2012, vol. 10, p. 352-365, http://hal.inria.fr/hal-00696231

[7] P.-N. MOUGEL, C. RIGOTTI, M. PLANTEVIT, O. GANDRILLON.*Finding maximal homogeneous clique sets*, in "Knowledge and Information Systems", March 2013, vol. 35, n^{o} 1, p. 1-30 [*DOI :* 10.1007/S10115-013-0625-Y], http://hal.inria.fr/hal-00827164

[8] C. ROCABERT, C. KNIBBE, G. BESLON.*Towards a Integrated Evolutionary Model to Study Evolution of Evolution*, in "EvoEvo Workshop (Satellite workshop of ECAL 2015)", York, United Kingdom, July 2015, https://hal.inria.fr/hal-01252796

[9] H. SOULA, B. CARÉ, G. BESLON, H. BERRY.*Anomalous versus slowed-down Brownian diffusion in the ligand-binding equilibrium*, in "Biophysical Journal", 2013, vol. 105, n^{o} 9, p. 2064-2073 [*DOI :* 10.1016/J.BPJ.2013.07.023], http://hal.inria.fr/hal-00720515

[10] G. SZÖLLŐSI, B. BOUSSAU, S. S. ABBY, E. TANNIER, V. DAUBIN.*Phylogenetic modeling of lateral gene transfer reconstructs the pattern and relative timing of speciations*, in "Proceedings- National Academy of Sciences Usa", October 2012, vol. 109, n^{o} 43, p. 17513-17518 [*DOI :* 10.1073/PNAS.1202997109], http://hal.inria.fr/hal-00740292

[11] J. VIÑUELAS, G. KANEKO, A. COULON, E. VALLIN, V. MORIN, C. MEJIA-POUS, J.-J. KUPIEC, G. BESLON, O. GANDRILLON.*Quantifying the contribution of chromatin dynamics to stochastic gene expression reveals long, locus-dependent periods between transcriptional bursts*, in "BMC Biology", February 2013, vol. 11, n^{o} 1, 15 [*DOI :* 10.1186/1741-7007-11-15], http://hal.inria.fr/inserm-00817963

Publications of the year

Articles in International Peer-Reviewed Journal

[12] Y. ANSELMETTI, W. DUCHEMIN, E. TANNIER, C. CHAUVE, S. BÉRARD.*Phylogenetic signal from rearrangements in 18 Anopheles species by joint scaffolding extant and ancestral genomes*, in "BMC Genomics",

May 2018, vol. 19, n^o S2, p. 1-15 [*DOI :* 10.1186/S12864-018-4466-7], https://hal.archives-ouvertes.fr/hal-01791884

[13] Y. CHEN, M. ARNAL-LEVRON, F. HULLIN-MATSUDA, C. KNIBBE, P. MOULIN, C. LUQUAIN-COSTAZ, I. DELTON.*In vitro oxidized HDL and HDL from type 2 diabetes patients have reduced ability to efflux oxysterols from THP-1 macrophages*, in "Biochimie", October 2018, vol. 153, p. 232-237 [*DOI :* 10.1016/J.BIOCHI.2018.04.018], https://hal.inria.fr/hal-01952968

[14] Y. CUI, I. PROKIN, A. MENDES, H. BERRY, L. VENANCE.*Robustness of STDP to spike timing jitter*, in "Scientific Reports", May 2018, vol. 8, n^o 1, p. 8139:1-15 [*DOI :* 10.1038/S41598-018-26436-Y], https://hal.inria.fr/hal-01788826

[15] Y. CUI, Y. YANG, Z. NI, Y. DONG, G. CAI, A. FONCELLE, S. MA, K. SANG, S. TANG, Y. LI, Y. SHEN, H. BERRY, S. WU, H. HU.*Astroglial-Kir4.1 in Lateral Habenula Drives Neuronal Bursts to Mediate Depression*, in "Nature", February 2018, vol. 554, p. 323-327 [*DOI :* 10.1038/NATURE25752], https://hal.archives-ouvertes.fr/hal-01683191

[16] A. DAVÍN, E. TANNIER, T. WILLIAMS, B. BOUSSAU, V. DAUBIN, G. SZÖLLŐSI.*Gene transfers can date the tree of life*, in "Nature Ecology & Evolution", May 2018, vol. 2, n^o 5, p. 904 - 909 [*DOI :* 10.1038/S41559-018-0525-3], https://hal.archives-ouvertes.fr/hal-01774884

[17] W. DUCHEMIN, G. GENCE, A.-M. ARIGON CHIFOLLEAU, L. ARVESTAD, M. BANSAL, V. BERRY, B. BOUSSAU, F. CHEVENET, N. COMTE, A. DAVÍN, C. DESSIMOZ, D. DYLUS, D. HASIC, D. MALLO, R. PLANEL, D. POSADA, C. SCORNAVACCA, G. SZÖLLŐSI, L. ZHANG, E. TANNIER, V. DAUBIN.*RecPhyloXML - a format for reconciled gene trees*, in "Bioinformatics", May 2018, p. 1-8 [*DOI :* 10.1093/BIOINFORMATICS/BTY389], https://hal-lirmm.ccsd.cnrs.fr/lirmm-01800296

[18] A. FONCELLE, A. MENDES, J. JEDRZEJEWSKA-SZMEK, S. VALTCHEVA, H. BERRY, K. T. BLACKWELL, L. VENANCE.*Modulation of spike-timing dependent plasticity: towards the inclusion of a third factor in computational models*, in "Frontiers in Computational Neuroscience", July 2018, vol. 12, 49 [*DOI :* 10.3389/FNCOM.2018.00049], https://hal.inria.fr/hal-01809075

[19] C. KNIBBE.*Souvenirs from ECAL 2017: create, play, experiment, discover – revealing the experimental power of virtual worlds*, in "Adaptive Behavior", February 2018, vol. 26, n^o 1, p. 37-40, https://hal.inria.fr/hal-01952933

[20] N. MÉGER, C. RIGOTTI, C. POTHIER, T. NGUYEN, F. LODGE, L. GUEGUEN, R. ANDRÉOLI, M.-P. DOIN, M. DATCU.*Ranking evolution maps for Satellite Image Time Series exploration: application to crustal deformation and environmental monitoring*, in "Data Mining and Knowledge Discovery", October 2018, p. 1-37 [*DOI :* 10.1007/S10618-018-0591-9], https://hal.archives-ouvertes.fr/hal-01898015

[21] T. NGUYEN, N. MÉGER, C. RIGOTTI, C. POTHIER, E. TROUVÉ, N. GOURMELEN, J.-L. MUGNIER.*A pattern-based method for handling confidence measures while mining satellite displacement field time series. Application to Greenland ice sheet and Alpine glaciers*, in "IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing", October 2018, vol. 11, n^o 11, p. 1 - 13 [*DOI :* 10.1109/JSTARS.2018.2874499], https://hal.archives-ouvertes.fr/hal-01912708

[22] F. OSCHMANN, H. BERRY, K. OBERMAYER, K. LENK.*From in silico astrocyte cell models to neuron-astrocyte network models: A review*, in "Brain Research Bulletin", 2018, vol. 136, p. 76-84 [*DOI :* 10.1016/J.BRAINRESBULL.2017.01.027], https://hal.archives-ouvertes.fr/hal-01461928

[23] H. XU, S. PÉREZ, A. CORNIL, B. DETRAUX, I. PROKIN, Y. CUI, B. DEGOS, H. BERRY, A. DE KERCHOVE D'EXAERDE, L. VENANCE.*Dopamine-endocannabinoid interactions mediate spike-timing dependent potentiation in the striatum*, in "Nature Communications", October 2018, vol. 9, 4118 [*DOI :* 10.1038/S41467-018-06409-5], https://hal.inria.fr/hal-01865929

[24] C. YI, J. THEILLON, A. KOULAKOFF, H. BERRY, C. GIAUME.*Monitoring gap junctional communication in astrocytes from acute adult mouse brain slices using the gap-FRAP technique*, in "Journal of Neuroscience Methods", 2018, vol. 303, p. 103-113 [*DOI :* 10.1016/J.JNEUMETH.2018.03.005], https://hal.archives-ouvertes.fr/hal-01734652

International Conferences with Proceedings

[25] T. NGUYEN, N. MÉGER, C. RIGOTTI, C. POTHIER, E. TROUVÉ, J.-L. MUGNIER.*Finding Complementary and Reliable Patterns in Displacement Field Time Series of Alpine Glaciers*, in "IGARSS 2018 - IEEE International Geoscience and Remote Sensing Symposium", Valencia, Spain, Proceedings of the 38th IEEE International Geoscience and Remote Sensing Symposium, July 2018, p. 4213-4216, https://hal.archives-ouvertes.fr/hal-01868749

[26] *Best Paper*
S. PEIGNIER, C. RIGOTTI, A. ROSSI, G. BESLON.*Weight-based search to find clusters around medians in subspaces*, in "SAC 2018 - ACM Symposium On Applied Computing", Pau, France, Proceedings of the 33rd ACM Symposium On Applied Computing, April 2018, p. 1-10, https://hal.archives-ouvertes.fr/hal-01869974.

Conferences without Proceedings

[27] G. BESLON, S. F. ELENA, P. HOGEWEG, D. SCHNEIDER, S. STEPNEY.*Evolving Living Technologies-Insights from the EvoEvo project*, in "SSBSE 2018 - 10th Symposium on Search-Based Software Engineering", Montpellier, France, LNCS, Springer, September 2018, vol. 11036, p. 46-62 [*DOI :* 10.1007/978-3-319-99241-9_2], https://hal.archives-ouvertes.fr/hal-01882636

[28] *Best Paper*
V. F. LIARD, D. P. PARSONS, J. ROUZAUD-CORNABAS, G. BESLON.*The Complexity Ratchet: Stronger than selection, weaker than robustness*, in "ALIFE 2018 - the 2018 conference on artificial Life", Tokyo, Japan, July 2018, p. 1-8 [*DOI :* 10.1162/ISAL_A_00051], https://hal.archives-ouvertes.fr/hal-01882628.

Scientific Books (or Scientific Book chapters)

[29] Y. ANSELMETTI, N. LUHMANN, S. BÉRARD, E. TANNIER, C. CHAUVE.*Comparative Methods for Reconstructing Ancient GenomeOrganization*, in "Comparative Genomics: Methods and Protocols", 2018, https://hal.archives-ouvertes.fr/hal-01690366

Other Publications

[30] G. BESLON, V. F. LIARD, S. F. ELENA.*Jump then Climb: can rearrangements predict the occurrence of mutational bursts?*, August 2018, 1, Evolution 2018 - Congress on Evolutionary Biology, Poster, https://hal.archives-ouvertes.fr/hal-01938800

[31] G. BESLON, V. F. LIARD, D. P. PARSONS, J. ROUZAUD-CORNABAS.*In silico experimental evolution shows that complexity can rise even in simple environments*, October 2018, 1, ICSB 2018 - 19th International Conference on Systems Biology, Poster, https://hal.inria.fr/hal-01939365

[32] J. LEHMAN, J. CLUNE, D. MISEVIC, C. ADAMI, J. BEAULIEU, P. J. BENTLEY, S. BERNARD, G. BESLON, D. M. BRYSON, N. CHENEY, A. CULLY, S. DONCIUEX, F. C. DYER, K. O. ELLEFSEN, R. FELDT, S. FISCHER, S. FORREST, A. FRENOY, C. GAGNEÉ, L. LE GOFF, L. M. GRABOWSKI, B. HODJAT, L. KELLER, C. KNIBBE, P. KRCAH, R. E. LENSKI, H. LIPSON, R. MACCURDY, C. MAESTRE, R. MIIKKULAINEN, S. MITRI, D. E. MORIARTY, J.-B. MOURET, A. D. NGUYEN, C. OFRIA, M. PARIZEAU, D. PARSONS, R. T. PENNOCK, W. F. PUNCH, T. S. RAY, M. SCHOENAUER, E. SHULTE, K. SIMS, K. O. STANLEY, F. TADDEI, D. TARAPORE, S. THIBAULT, W. WEIMER, R. WATSON, J. YOSINKSI.*The Surprising Creativity of Digital Evolution: A Collection of Anecdotes from the Evolutionary Computation and Artificial Life Research Communities*, March 2018, https://arxiv.org/abs/1803.03453 - working paper or preprint, https://hal.inria.fr/hal-01735473

[33] V. F. LIARD, J. ROUZAUD-CORNABAS, D. P. PARSONS, G. BESLON.*The complexity ratchet: stronger than selection*, August 2018, 1, Beacon congress 2018, Poster, https://hal.archives-ouvertes.fr/hal-01938802

References in notes

[34] P. GALISON.*Image and Logic: A Material Culture of Microphysics*, University Of Chicago Press, 1997

[35] B. VERD, E. CLARK, K. WOTTON, H. JANSSENS, E. JIMÉNEZ-GURI, A. CROMBACH, J. JAEGER.*A damped oscillator imposes temporal order on posterior gap gene expression in Drosophila*, in "PLoS Biology", February 2018, vol. 16, n^{o} 2 [*DOI :* 10.1371/JOURNAL.PBIO.2003174], https://hal.inria.fr/hal-01934923

[36] C. VORS, G. PINEAU, L. GABERT, AL..*Modulating absorption and postprandial handling of dietary fatty acids by structuring fat in the meal: a randomized crossover clinical trial*, in "Am J Clin Nutr", 2013, vol. 97, p. 23–36

Team CASH

Compilation and Analysis, Software and Hardware

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Architecture, Languages and Compilation

Table of contents

Team CASH

Creation of the Team: 2018 April 01

Keywords:

Computer Science and Digital Science:

A1.1.1. - Multicore, Manycore
A1.1.2. - Hardware accelerators (GPGPU, FPGA, etc.)
A1.1.4. - High performance computing
A1.1.10. - Reconfigurable architectures
A1.1.12. - Non-conventional architectures
A2.1.1. - Semantics of programming languages
A2.1.6. - Concurrent programming
A2.1.7. - Distributed programming
A2.2.1. - Static analysis
A2.2.3. - Memory management
A2.2.4. - Parallel architectures
A2.2.6. - GPGPU, FPGA...
A2.2.8. - Code generation
A2.3.1. - Embedded systems

Other Research Topics and Application Domains:

B9.5.1. - Computer science

1. Team, Visitors, External Collaborators

Research Scientists

Christophe Alias [Inria, Researcher]
Ludovic Henrio [CNRS, Researcher, from Oct 2018, HDR]

Faculty Members

Matthieu Moy [Team leader, Univ Claude Bernard, Associate Professor, HDR]
Laure Gonnord [Univ Claude Bernard, Associate Professor, HDR]

PhD Students

Julien Braine [Ecole Normale Supérieure Lyon, from Sep 2018]
Paul Iannetta [Ecole Normale Supérieure Lyon, from Sep 2018]

Administrative Assistant

Laetitia Gauthe [Inria, until Nov 2018]

2. Overall Objectives

2.1. Introduction

Until 2006, the typical power-consumption of a chip remained constant for a given silicon area as the transistor size decreased (this evolution is referred to as Dennard scaling). In other words, energy efficiency was following an exponential law similar to Moore's law. This is no longer true, hence radical changes are needed to further improve power efficiency, which is the limiting factor for large-scale computing. Improving the performance under a limited energy budget must be done by rethinking computing systems at all levels: hardware, software, compilers, and runtimes.

On the hardware side, new architectures such as multi-core processors, Graphics Processing Units (GPUs), many-core and FPGA accelerators are introduced, resulting into complex heterogeneous platforms. In particular, FPGAs are now a credible solution for energy-efficient HPC. An FPGA chip can deliver the same computing power as a GPU for an energy budget 10 times smaller.

A consequence of this diversity and heterogeneity is that a given computation can be implemented in many different ways, with different performance characteristics. An obvious example is changing the degree of parallelism: this allows trading execution time for number of cores used. However, many choices are less obvious: for example, augmenting the degree of parallelism of a memory-bounded application will not improve performance. Most architectures involve a complex memory hierarchy, hence memory access patterns have a considerable impact on performance too. The design-space to be explored to find the best performance is much wider than it used to be with older architectures, and new tools are needed to help the programmer explore it. The problem is even stronger for FPGA accelerators, where programmers are expected to design a circuit for their application! Traditional synthesis tools take as input low-level languages like VHDL and Verilog. As opposed to this, high-level languages and hardware compilers (HLS, High-Level Synthesis, that takes as input a C or C-like language and produces a circuit description) are required.

One of the bottlenecks of performance and energy efficiency is data movement. The operational intensity (ratio computation/communication) must be optimized to avoid memory-bounded performance. Compiler analyses are strongly required to explore the trade-offs (operational intensity vs. local memory size, operational intensity vs. peak performance for reconfigurable circuits).

These issues are considered as one of the main challenges in the Hipeac roadmap [27] which, among others, cites the two major issues:

- Applications are moving towards global-scale services, accessible across the world and on all devices. Low power processors, systems, and communications are key to computing at this scale. (*Strategic Area 2, Data Center Computing*).
- Today data movement uses more power than computation. [...] To adapt to this change, we need to expose data movement in applications and optimize them at runtime and compile time and to investigate communication-optimized algorithms (*cross-cutting challenge 1, energy efficiency*).

2.2. Overall Objectives

The overall objective of the CASH team is to take advantage of the characteristics of the specific hardware (generic hardware, hardware accelerators, or reconfigurable chips) to *compile energy efficient software and hardware*. More precisely, we plan to work on:

1. Definition of dataflow representations of parallel programs that can capture the parallelism at all levels: fine-grain vs. coarse-grain, data & task parallelism, programming language, and intermediate representation (Section 3.1).
2. Scalable and expressive static program analyses. CASH will work on improving the scalability of analyses to allow a global analysis of large-scale programs, and on the expressiveness of analysis to find better program invariants. Analysis will be performed both on the representation defined above and on general programs (Section 3.2).
3. Transformations from and to the dataflow representation, combining traditional tools dedicated to dataflow and specific methods like the polyhedral model (Section 3.3).
4. A high-level synthesis (HLS) tool, built on the above item (instantiated with the particularities of FPGAs) and a code generation tool (Section 3.4). This HLS tool will focus on early stages of compilation and rely on an external tool for the back-end.
5. A parallel and scalable simulation of hardware systems, which, combined with the preceding activity, will result in an end-to-end workflow for circuit design (Section 3.5).

To ensure the coherency and the correctness of our approach these different tasks will rely on a *precise definition of the manipulated languages and their semantics*. The formalization of the different representations of the programs and of the analyses will allow us to show that these different tasks will be performed with the same understanding of the program semantics.

Note that these directions are strongly tied together. We use 5 research axis for the sake of the presentation, but their complementarity enables each member of the team to share common research goals while having their own research directions. Most of our results will contribute to several directions.

3. Research Program

3.1. Definition of dataflow representations of parallel programs

In the last decades, several frameworks have emerged to design efficient compiler algorithms. The efficiency of all the optimizations performed in compilers strongly relies on effective *static analyses* and *intermediate representations*. Dataflow models are a natural intermediate representation for hardware compilers (HLS) and more generally for parallelizing compilers. Indeed, dataflow models capture task-level parallelism and can be mapped naturally to parallel architectures. In a way, a dataflow model is a partition of the computation into processes and a partition of the flow dependences into channels. This partitioning prepares resource allocation (which processor/hardware to use) and medium-grain communications.

The main goal of the CASH team is to provide efficient analyses and the optimizing compilation frameworks for dataflow programming models. The results of the team will rely on programming languages and representation of programs in which parallelism and dataflow play a crucial role. This first research direction aims at defining these dataflow languages and intermediate representations, both from a practical perspective (syntax or structure), and from a theoretical point of view (semantics). This first research direction thus defines the models on which the other directions will rely. It is important to note that we do not restrict ourself to a strict definition of dataflow languages and, more generally, we are interested in the parallel languages in which dataflow synchronization plays a significant role.

Intermediate dataflow model. The intermediate dataflow model is a representation of the program that is adapted for optimization and scheduling. It will be obtained from the analysis of a (parallel or sequential) program and should at some point be used for compilation. The dataflow model must specify precisely its semantics and parallelism granularity. It must also be analyzable with polyhedral techniques, where powerful concepts exist to design compiler analysis, e.g., scheduling or resource allocation. Polyhedral Process Networks [55] extended with a module system could be a good starting point. But then, how to fit non-polyhedral parts of the program? A solution is to hide non-polyhedral parts into processes with a proper polyhedral abstraction. This organization between polyhedral and non-polyhedral processes will be a key aspect of our medium-grain dataflow model. The design of our intermediate dataflow model and the precise definition of its semantics will constitute a reliable basis to formally define and ensure the correctness of algorithms proposed by CASH: compilation, optimizations and analyses.

Dataflow programming languages. Dataflow paradigm has also been explored quite intensively in programming languages. Indeed, there exists a large panel of dataflow languages, whose characteristics differ notably, the major point of variability being the scheduling of agents and their communications. There is indeed a continuum from the synchronous dataflow languages like Lustre [37] or Streamit [51], where the scheduling is fully static, and general communicating networks like KPNs [39] or RVC-Cal [19] where a dedicated runtime is responsible for scheduling tasks dynamically, when they *can* be executed. These languages share some similarities with actor languages that go even further in the decoupling of processes by considering them as independent reactive entities. Another objective of the CASH team is to study dataflow programming *languages*, their semantics, their expressiveness, and their compilation. The specificity of the CASH team will be that these languages will be designed taking into consideration the compilation using polyhedral techniques. In particular, we will explore which dataflow constructs are better adapted for our static analysis, compilation, and scheduling techniques. In practice we want to propose high-level primitives to express data dependency, this way the programmer will express parallelism in a dataflow way instead of the classical communication-oriented dependencies. The higher-level more declarative point of view will make programming easier but also give more optimization opportunities. These primitives will be inspired by the existing works in the polyhedral model framework, as well as dataflow languages, but also in the actors and active object languages [26] that nowadays introduce more and more dataflow primitives to enable data-driven interactions between agents, particularly with *futures* [24], [31].

3.1.1. Expected Impact

Consequently, the impact of this research direction is both the usability of our representation for static analyses and optimizations performed in Sections 3.2 and 3.3, and the usability of its semantics to prove the correctness of these analyses.

3.1.2. Scientific Program

3.1.2.1. Short-term and ongoing activities.

We obtained preliminary experimental [16], [17], [32] and theoretical [38] results, exploring several aspects of dataflow models. The next step is to define accurately the intermediate dataflow model and to study existing programming and execution models:

- Define our medium-grain dataflow model. So far, a modular Polyhedral Process Networks appears as a natural candidate but it may need to be extended to be adapted to a wider range of applications. Precise semantics will have to be defined for this model to ensure the articulation with the activities discussed in Section 3.3.
- Study precisely existing dataflow languages, their semantics, their programmability, and their limitations.

3.1.2.2. Medium-term activities.

In a second step, we will extend the existing results to widen the expressiveness of our intermediate representation and design new parallelism constructs. We will also work on the semantics of dataflow languages:

- Propose new stream programming models and a clean semantics where all kinds of parallelisms are expressed explicitly, and where all activities from code design to compilation and scheduling can be clearly expressed.
- Identify a core language that is rich enough to be representative of the dataflow languages we are interested in, but abstract and small enough to enable formal reasoning and proofs of correctness for our analyses and optimizations.

3.1.2.3. Long-term activities.

In a longer-term vision, the work on semantics, while remaining driven by the applications, would lead to to more mature results, for instance:

- Design more expressive dataflow languages and intermediate representations which would at the same time be expressive enough to capture all the features we want for aggressive HPC optimizations, and sufficiently restrictive to be (at least partially) statically analyzable at a reasonable cost.
- Define a module system for our medium-grain dataflow language. A program will then be divided into modules that can follow different compilation schemes and execution models but still communicate together. This will allow us to encapsulate a program that does not fit the polyhedral model into a polyhedral one and vice versa. Also, this will allow a compositional analysis and compilation, as opposed to global analysis which is limited in scalability.

3.2. Expressivity and Scalability of Static Analyses

The design and implementation of efficient compilers becomes more difficult each day, as they need to bridge the gap between *complex languages* and *complex architectures*. Application developers use languages that bring them close to the problem that they need to solve which explains the importance of high-level programming languages. However, high-level programming languages tend to become more distant from the hardware which they are meant to command.

In this research direction, we propose to design expressive and scalable static analyses for compilers. This topic is closely linked to Sections 3.1 and 3.3 since the design of an efficient intermediate representation is made while regarding the analyses it enables. The intermediate representation should be expressive enough to embed maximal information; however if the representation is too complex the design of scalable analyses will be harder.

The analyses we plan to design in this activity will of course be mainly driven by the HPC dataflow optimizations we mentioned in the preceding sections; however we will also target other kinds of analyses applicable to more general purpose programs. We will thus consider two main directions:

- Extend the applicability of the polyhedral model, in order to deal with HPC applications that do not fit totally in this category. More specifically, we plan to work on more complex control and also on complex data structures, like sparse matrices, which are heavily used in HPC.
- Design of specialized static analyses for memory diagnostic and optimization inside general purpose compilers.

For both activities, we plan to cross fertilize ideas coming from the abstract interpretation community as well as language design, dataflow semantics, and WCET estimation techniques.

Correct by construction analyses. The design of well-defined semantics for the chosen programming language and intermediate representation will allow us to show the correctness of our analyses. The precise study of the semantics of Section 3.1 will allow us to adapt the analysis to the characteristics of the language, and prove that such an adaptation is well founded. This approach will be applicable both on the source language and on the intermediate representation.

Such wellfoundedness criteria relatively to the language semantics will first be used to design our analyses, and then to study which extensions of the languages can be envisioned and analyzed safely, and which extensions (if any) are difficult to analyze and should be avoided. Here the correct identification of a core language for our formal studies (see Section 3.1) will play a crucial role as the core language should feature all the characteristics that might make the analysis difficult or incorrect.

Scalable abstract domains. We already have experience in designing low-cost semi relational abstract domains for pointers [44], [42], as well as tailoring static analyses for specialized applications in compilation [30], [50], Synchronous Dataflow scheduling [49], and extending the polyhedral model to irregular applications [15]. We also have experience in the design of various static verification techniques adapted to different programming paradigms.

3.2.1. Expected impact

The impact of this work is the significantly widened applicability of various tools/compilers related to parallelization: allow optimizations for a larger class of programs, and allow low-cost analysis that scale to very large programs.

We target both analysis for optimization and analysis to detect, or prove the absence of bugs.

3.2.2. Scientific Program

3.2.2.1. Short-term and ongoing activities.

Together with Paul Iannetta and Lionel Morel (INSA/CEA LETI), we are currently working on the *semantic rephrasing* of the polyhedral model [34]. The objective is to clearly redefine the key notions of the polyhedral model on general flowchart programs operating on arrays, lists and trees. We reformulate the algorithms that are performed to compute dependencies in a more semantic fashion, i.e. considering the program semantics instead of syntactical criteria. The next step is to express classical scheduling and code generation activities in this framework, in order to overcome the classical syntactic restrictions of the polyhedral model.

3.2.2.2. Medium-term activities.

In medium term, we want to extend the polyhedral model for more general data-structures like lists and sparse matrices. For that purpose, we need to find polyhedral (or other shapes) abstractions for non-array data-structures; the main difficulty is to deal with non-linearity and/or partial information (namely, over-approximations of the data layout, or over-approximation of the program behavior). This activity will rely on a formalization of the optimization activities (dependency computation, scheduling, compilation) in a more general Abstract-Interpretation based framework in order to make the approximations explicit.

At the same time, we plan to continue to work on scaling static analyses for general purpose programs, in the spirit of Maroua Maalej's PhD [41], whose contribution is a sequence of memory analyses inside production compilers. We already began a collaboration with Sylvain Collange (PACAP team of IRISA Laboratory) on the design of static analyses to optimize copies from the global memory of a GPU to the block kernels (to increase locality). In particular, we have the objective to design specialized analyses but with an explicit notion of cost/precision compromise, in the spirit of the paper [36] that tries to formalize the cost/precision compromise of interprocedural analyses with respect to a "context sensitivity parameter".

3.2.2.3. Long-term activities.

In a longer-term vision, the work on scalable static analyses, whether or not directed from the dataflow activities, will be pursued in the direction of large general-purpose programs.

An ambitious challenge is to find a generic way of adapting existing (relational) abstract domains within the Single Static Information [20] framework so as to improve their scalability. With this framework, we would be able to design static analyses, in the spirit of the seminal paper [25] which gave a theoretical scheme for classical abstract interpretation analyses.

We also plan to work on the interface between the analyses and their optimization clients inside production compilers.

3.3. Compiling and Scheduling Dataflow Programs

In this part, we propose to design the compiler analyses and optimizations for the *medium-grain* dataflow model defined in section 3.1. We also propose to exploit these techniques to improve the compilation of dataflow languages based on actors. Hence our activity is split into the following parts:

- Translating a sequential program into a medium-grain dataflow model. The programmer cannot be expected to rewrite the legacy HPC code, which is usually relatively large. Hence, compiler techniques must be invented to do the translation.
- Transforming and scheduling our medium-grain dataflow model to meet some classic optimization criteria, such as throughput, local memory requirements, or I/O traffic.
- Combining agents and polyhedral kernels in dataflow languages. We propose to apply the techniques above to optimize the processes in actor-based dataflow languages and combine them with the parallelism existing in the languages.

We plan to rely extensively on the polyhedral model to define our compiler analysis. The polyhedral model was originally designed to analyze imperative programs. Analysis (such as scheduling or buffer allocation) must be redefined in light of dataflow semantics.

Translating a sequential program into a medium-grain dataflow model. The programs considered are compute-intensive parts from HPC applications, typically big HPC kernels of several hundreds of lines of C code. In particular, we expect to analyze the process code (actors) from the dataflow programs. On short ACL (Affine Control Loop) programs, direct solutions exist [53] and rely directly on array dataflow analysis [29]. On bigger ACL programs, this analysis no longer scales. We plan to address this issue by *modularizing* array dataflow analysis. Indeed, by splitting the program into processes, the complexity is mechanically reduced. This is a general observation, which was exploited in the past to compute schedules [28]. When the program is no longer ACL, a clear distinction must be made between polyhedral parts and non polyhedral parts. Hence, our medium-grain dataflow language must distinguish between polyhedral process networks, and non-polyhedral code fragments. This structure raises new challenges: How to abstract away non-polyhedral parts while keeping the polyhedrality of the dataflow program? Which trade-off(s) between precision and scalability are effective?

Medium-grain data transfers minimization. When the system consists of a single computing unit connected to a slow memory, the roofline model [56] defines the optimal ratio of computation per data transfer (*operational intensity*). The operational intensity is then translated to a partition of the computation (loop tiling) into *reuse units*: inside a reuse unit, data are transfered locally; between reuse units, data are transfered through the slow memory. On a *fine-grain* dataflow model, reuse units are exposed with loop tiling; this is the case for example in Data-aware Process Network (DPN) [17]. The following questions are however still open: How does that translate on *medium-grain* dataflow models? And fundamentally what does it mean to *tile* a dataflow model?

Combining agents and polyhedral kernels in dataflow languages. In addition to the approach developed above, we propose to explore the compilation of dataflow programming languages. In fact, among the applications targeted by the project, some of them are already thought or specified as dataflow actors (video compression, machine-learning algorithms,...).

So far, parallelization techniques for such applications have focused on taking advantage of the decomposition into agents, potentially duplicating some agents to have several instances that work on different data items in parallel [35]. In the presence of big agents, the programmer is left with the splitting (or merging) of these agents by-hand if she wants to further parallelize her program (or at least give this opportunity to the runtime, which in general only sees agents as non-malleable entities). In the presence of arrays and loop-nests, or, more generally, some kind of regularity in the agent's code, however, we believe that the programmer would benefit from automatic parallelization techniques such as those proposed in the previous paragraphs. To achieve the goal of a totally integrated approach where programmers write the applications they have in mind (application flow in agents where the agents' code express potential parallelism), and then it is up to the system (compiler, runtime) to propose adequate optimizations, we propose to build on solid formal definition of the language semantics (thus the formal specification of parallelism occurring at the agent level) to provide hierarchical solutions to the problem of compilation and scheduling of such applications.

3.3.1. Expected impact

In general, splitting a program into simpler processes simplifies the problem. This observation leads to the following points:

- By abstracting away irregular parts in processes, we expect to structure the long-term problem of handling irregular applications in the polyhedral model. The long-term impact is to widen the applicability of the polyhedral model to irregular kernels.
- Splitting a program into processes reduces the problem size. Hence, it becomes possible to scale traditionally expensive polyhedral analysis such as scheduling or tiling to quote a few.

As for the third research direction, the short term impact is the possibility to combine efficiently classical dataflow programming with compiler polyhedral-based optimizations. We will first propose ad-hoc solutions coming from our HPC application expertise, but supported by strong theoretical results that prove their correctness and their applicability in practice. In the longer term, our work will allow specifying, designing, analyzing, and compiling HPC dataflow applications in a unified way. We target semi-automatic approaches where pertinent feedback is given to the developer during the development process.

3.3.2. Scientific Program

3.3.2.1. Short-term and ongoing activities.

We are currently working on the RTM (Reverse-Time Migration) kernel for oil and gas applications (≈ 500 lines of C code). This kernel is long enough to be a good starting point, and small enough to be handled by a polyhedral splitting algorithm. We figured out the possible splittings so the polyhedral analysis can scale and irregular parts can be hidden. In a first step, we plan to define splitting metrics and algorithms to optimize the usual criteria: communication volume, latency and throughput.

Together with Lionel Morel (INSA/CEA LETI), we currently work on the evaluation of the practical advantage of combining the dataflow paradigm with the polyhedral optimization framework. We empirically build a proof-of-concept tooling approach, using existing tools on existing languages [33]. We combine dataflow programming with polyhedral compilation in order to enhance program parallelization by leveraging both inter-agent parallelism and intra-agent parallelism (i.e., regarding loop nests inside agents). We evaluate the approach practically, on benchmarks coming from image transformation or neural networks, and the first results demonstrate that there is indeed a room for further improvement.

3.3.2.2. Medium-term activities.

The results of the preceding paragraph are partial and have been obtained with a simple experimental approach only using off-the-shelf tools. We are thus encouraged to pursue research on combining expertise from dataflow programming languages and polyhedral compilation. Our long term objective is to go towards a formal framework to express, compile, and run dataflow applications with intrinsic instruction or pipeline parallelism.

We plan to investigate in the following directions:

- Investigate how polyhedral analysis extends on modular dataflow programs. For instance, how to modularize polyhedral scheduling analysis on our dataflow programs?
- Develop a proof of concept and validate it on linear algebra kernels (SVD, Gram-Schmidt, etc.).
- Explore various areas of applications from classical dataflow examples, like radio and video processing, to more recent applications in deep learning algorithmic. This will enable us to identify some potential (intra and extra) agent optimization patterns that could be leveraged into new language idioms.

3.3.2.3. Long-term activities.

Current work focus on purely polyhedral applications. Irregular parts are not handled. Also, a notion of tiling is required so the communications of the dataflow program with the outside world can be tuned with respect to the local memory size. Hence, we plan to investigate the following points:

- Assess simple polyhedral/non polyhedral partitioning: How non-polyhedral parts can be hidden in processes/channels? How to abstract the dataflow dependencies between processes? What would be the impact on analyses? We target programs with irregular control (e.g., while loop, early exits) and regular data (arrays with affine accesses).
- Design tiling schemes for modular dataflow programs: What does it mean to tile a dataflow program? Which compiler algorithms to use?
- Implement a mature compiler infrastructure from the front-end to code generation for a reasonable subset of the representation.

3.4. HLS-specific Dataflow Optimizations

The compiler analyses proposed in section 3.3 do not target a specific platform. In this part, we propose to leverage these analysis to develop source-level optimizations for high-level synthesis (HLS).

High-level synthesis consists in compiling a kernel written in a high-level language (typically in C) into a circuit. As for any compiler, an HLS tool consists in a *front-end* which translates the input kernel into an *intermediate representation*. This intermediate representation captures the control/flow dependences between computation units, generally in a hierarchical fashion. Then, the *back-end* maps this intermediate representation to a circuit (e.g. FPGA configuration). We believe that HLS tools must be thought as fine-grain automatic parallelizers. In classic HLS tools, the parallelism is expressed and exploited at the back-end level during the scheduling and the resource allocation of arithmetic operations. We believe that it would be far more profitable to derive the parallelism at the front-end level.

Hence, CASH will focus on the *front-end* pass and the *intermediate representation*. Low-level *back-end* techniques are not in the scope of CASH. Specifically, CASH will leverage the dataflow representation developed in Section 3.1 and the compilation techniques developed in Section 3.3 to develop a relevant intermediate representation for HLS and the corresponding front-end compilation algorithms.

Our results will be evaluated by using existing HLS tools (e.g., Intel HLS compiler, Xilinx Vivado HLS). We will implement our compiler as a source-to-source transformation in front of HLS tools. With this approach, HLS tools are considered as a "back-end black box". The CASH scheme is thus: (i) *front-end*: produce the CASH dataflow representation from the input C kernel. Then, (ii) turn this dataflow representation to a C program with pragmas for an HLS tool. This step must convey the characteristics of the dataflow representation found by step (i) (e.g. dataflow execution, fifo synchronisation, channel size). This source-to-source approach will allow us to get a full source-to-FPGA flow demonstrating the benefits of our tools while relying on existing tools for low-level optimizations. Step (i) will start from the DCC tool developed by Christophe Alias, which already produces a dataflow intermediate representation: the Data-aware Process Networks (DPN) [17]. Hence, the very first step is then to chose an HLS tool and to investigate which input should be fed to the HLS tool so it "respects" the parallelism and the resource allocation suggested by the DPN. From this basis, we plan to investigate the points described thereafter.

Roofline model and dataflow-level resource evaluation. Operational intensity must be tuned according to the roofline model. The roofline model [56] must be redefined in light of FPGA constraints. Indeed, the peak performance is no longer constant: it depends on the operational intensity itself. The more operational intensity we need, the more local memory we use, the less parallelization we get (since FPGA resources are limited), and finally the less performance we get! Hence, multiple iterations may be needed before reaching an efficient implementation. To accelerate the design process, we propose to iterate at the dataflow program level, which implies a fast resource evaluation at the dataflow level.

Reducing FPGA resources. Each parallel unit must use as little resources as possible to maximize parallel duplication, hence the final performance. This requires to factorize the control and the channels. Both can be achieved with source-to-source optimizations at dataflow level. The main issue with outputs from polyhedral optimization is large piecewise affine functions that require a wide silicon surface on the FPGA to be computed. Actually we do not need to compute a closed form (expression that can be evaluated in bounded time on the FPGA) *statically*. We believe that the circuit can be compacted if we allow control parts to be evaluated dynamically. Finally, though dataflow architectures are a natural candidate, adjustments are required to fit FPGA constraints (2D circuit, few memory blocks). Ideas from systolic arrays [48] can be borrowed to re-use the same piece of data multiple times, despite the limitation to regular kernels and the lack of I/O flexibility. A trade-off must be found between pure dataflow and systolic communications.

Improving circuit throughput. Since we target streaming applications, the throughput must be optimized. To achieve such an optimization, we need to address the following questions. How to derive an optimal upper bound on the throughput for polyhedral process network? Which dataflow transformations should be performed to reach it? The limiting factors are well known: I/O (decoding of burst data), communications through addressable channels, and latencies of the arithmetic operators. Finally, it is also necessary to find the right methodology to measure the throughput statically and/or dynamically.

3.4.1. Expected Impact

So far, the HLS front-end applies basic loop optimizations (unrolling, flattening, pipelining, etc.) and use a Hierarchical Control Flow Graph-like representation with data dependencies annotations (HCDFG). With this approach, we intend to demonstrate that polyhedral analysis combined with dataflow representations is an effective solution for HLS tools.

3.4.2. Scientific Program

3.4.2.1. Short-term and ongoing activities.

The HLS compiler designed in the CASH team currently extracts a fine-grain parallel intermediate representation (DPN [17], [16]) from a sequential program. We will not write a back-end that produces code for FPGA but we need to provide C programs that can be fed into existing C-to-FPGA compilers. However we obviously need an end-to-end compiler for our experiments. One of the first task of our HLS activity is to develop a DPN-to-C code generator suitable as input to an existing HLS tool like Vivado HLS. The generated code should exhibit the parallelism extracted by our compiler, and allow generating a final circuit more efficient than the one that would be generated by our target HLS tool if ran directly on the input program. Source-to-source approaches have already been experimented successfully, e.g. in Alexandru Plesco's PhD [45].

3.4.2.2. Medium-term activities.

Our DPN-to-C code generation will need to be improved in many directions. The first point is the elimination of redundancies induced by the DPN model itself: buffers are duplicated to allow parallel reads, processes are produced from statements in the same loop, hence with the same control automaton. Also, multiplexing uses affine constraints which can be factorized [18]. We plan to study how these constructs can be factorized at C-level and to design the appropriate DPN-to-C translation algorithms.

Also, we plan to explore how on-the-fly evaluation can reduce the complexity of the control. A good starting point is the control required for the load process (which fetch data from the distant memory). If we want to avoid multiple load of the same data, the FSM (Finite State Machine) that describes it is usually very complex. We believe that dynamic construction of the load set (set of data to load from the main memory) will use less silicon than an FSM with large piecewise affine functions computed statically.

3.4.2.3. Long-term activities.

The DPN-to-C compiler will open new research perspectives. We will explore the roofline model accuracy for different applications by playing on DPN parameters (tile size). Unlike the classical roofline model, the peak performance is no longer assumed to be constant, but decreasing with operational intensity [58]. Hence, we expect a *unique* optimal set of parameters. Thus, we will build a DPN-level cost model to derive an interval containing the optimal parameters.

Also, we want to develop DPN-level analysis and transformation to quantify the optimal reachable throughput and to reach it. We expect the parallelism to increase the throughput, but in turn it may require an operational intensity beyond the optimal point discussed in the first paragraph. We will assess the trade-offs, build the cost-models, and the relevant dataflow transformations.

3.5. Simulation of Hardware

Complex systems such as systems-on-a-chip or HPC computer with FPGA accelerator comprise both hardware and software parts, tightly coupled together. In particular, the software cannot be executed without the hardware, or at least a simulator of the hardware.

Because of the increasing complexity of both software and hardware, traditional simulation techniques (Register Transfer Level, RTL) are too slow to allow full system simulation in reasonable time. New techniques such as Transaction Level Modeling (TLM) [14] in SystemC [13] have been introduced and widely adopted in the industry. Internally, SystemC uses discrete-event simulation, with efficient context-switch using cooperative scheduling. TLM abstracts away communication details, and allows modules to communicate using function calls. We are particularly interested in the loosely timed coding style where the timing of the platform is not modeled precisely, and which allows the fastest simulations. This allowed gaining several orders of magnitude of simulation speed. However, SystemC/TLM is also reaching its limits in terms of performance, in particular due to its lack of parallelism.

Work on SystemC/TLM parallel execution is both an application of other work on parallelism in the team and a tool complementary to HLS presented in Sections 3.1 (dataflow models and programs) and 3.4 (application to FPGA). Indeed, some of the parallelization techniques we develop in CASH could apply to SystemC/TLM programs. Conversely, a complete design-flow based on HLS needs fast system-level simulation: the full-system usually contains both hardware parts designed using HLS, handwritten hardware components, and software.

We will also work on simulation of the DPN intermediate representation. Simulation is a very important tool to help validate and debug a complete compiler chain. Without simulation, validating the front-end of the compiler requires running the full back-end and checking the generated circuit. Simulation can avoid the execution time of the backend and provide better debugging tools.

Automatic parallelization has shown to be hard, if at all possible, on loosely timed models [23]. We focus on semi-automatic approaches where the programmer only needs to make minor modifications of programs to get significant speedups.

3.5.1. Expected Impact

The short term impact is the possibility to improve simulation speed with a reasonable additional programming effort. The amount of additional programming effort will thus be evaluated in the short term.

In the longer term, our work will allow scaling up simulations both in terms of models and execution platforms. Models are needed not only for individual Systems on a Chip, but also for sets of systems communicating together (e.g., the full model for a car which comprises several systems communicating together), and/or heterogeneous models. In terms of execution platform, we are studying both parallel and distributed simulations.

3.5.2. Scientific Program

3.5.2.1. Short-term and ongoing activities.

We started the joint PhD (with Tanguy Sassolas) of Gabriel Busnot with CEA-LIST. The research targets parallelizing SystemC heterogeneous simulations. CEA-LIST already developed SCale [54], which is very efficient to simulate parallel homogeneous platforms such as multi-core chips. However, SCale cannot currently load-balance properly the computations when the platform contains different components modeled at various levels of abstraction. Also, SCale requires manual annotations to identify accesses to shared variables. These annotations are given as address ranges in the case of a shared memory. This annotation scheme does not work when the software does non-trivial memory management (virtual memory using a memory management unit, dynamic allocation), since the address ranges cannot be known statically. We started working on the "heterogeneous" aspect of simulations with an approach allowing changing the level of details in a simulation at runtime, and started tackling the virtual and dynamic memory management problem by porting Linux on our simulation platform.

We also started working on an improved support for simulation and debugging of the DPN internal representation of our parallelizing compiler (see Section 3.3). A previous quick experiment with simulation was to generate C code that simulates parallelism with POSIX-threads. While this simulator greatly helped debug the compiler, this is limited in several ways: simulations are not deterministic, and the simulator does not scale up since it would create a very large number of threads for a non-trivial design.

We are working in two directions. The first is to provide user-friendly tools to allow graphical inspection of traces. For example, we will work on the visualization of the sequence of steps leading to a deadlock when the situation occurs, and give hints on how to fix the problem in the compiler. The second is to use an efficient simulator to speed up the simulation. We plan to generate SystemC/TLM code from the DPN representation to benefit from the ability of SystemC to simulate a large number of processes.

3.5.2.2. Medium-term activities.

Several research teams have proposed different approaches to deal with parallelism and heterogeneity. Each approach targets a specific abstraction level and coding style. While we do not hope for a universal solution, we believe that a better coordination of different actors of the domain could lead to a better integration of solutions. We could imagine, for example, a platform with one subsystem accelerated with SCale [54] from CEA-LIST, some compute-intensive parts delegated to sc-during [43] from Matthieu Moy, and a co-simulation with external physical solvers using SystemC-MDVP [21] from LIP6. We plan to work on the convergence of approaches, ideally both through point-to-point collaborations and with a collaborative project.

A common issue with heterogeneous simulation is the level of abstraction. Physical models only simulate one scenario and require concrete input values, while TLM models are usually abstract and not aware of precise physical values. One option we would like to investigate is a way to deal with loose information, e.g. manipulate intervals of possible values instead of individual, concrete values. This would allow a simulation to be symbolic with respect to the physical values.

Obviously, works on parallel execution of simulations would benefit to simulation of data-aware process networks (DPN). Since DPN are generated, we can even tweak the generator to guarantee some properties on the generated code, which will give us more freedom on the parallelization and partitioning techniques.

3.5.2.3. Long-term activities.

In the long term, our vision is a simulation framework that will allow combining several simulators (not necessarily all SystemC-based), and allow running them in a parallel way. The Functional Mockup Interface (FMI) standard is a good basis to build upon, but the standard does not allow expressing timing and functional constraints needed for a full co-simulation to run properly.

4. Application Domains

4.1. Compute-intensive Loop Kernels

The CASH team targets HPC programs, at different levels. Small computation kernels (tens of lines of code) that can be analyzed and optimized aggressively, medium-size kernels (hundreds of lines of code) that require modular analysis, and assembly of compute kernels (either as classical imperative programs or written directly in a dataflow language).

The work on various application domains and categories of programs is driven by the same idea: exploring various topics is a way to converge on unifying representations and algorithms even for specific applications. All these applications share the same research challenge: find a way to integrate computations, data, mapping, and scheduling in a common analysis and compilation framework.

Typical HPC kernels include linear solvers, stencils, matrix factorizations, BLAS kernels, etc. Many kernels can be found in the Polybench/C benchmark suite [46]. The irregular versions can be found in [47]. Numerical kernels used in quantitative finance [57] are also good candidates, e.g., finite difference and Monte-Carlo simulation.

4.2. Medium-size HPC applications

The medium-size applications we target are streaming algorithms [19], scientific workflows [52], and also the now very rich domain of deep learning applications [40]. We explore the possibilities of writing (see Section 3.1) and compiling (see Section 3.3) applications using a dataflow language. As a first step, we will target dataflow programs written in SigmaC [22] for which the fine grain parallelism is not taken into account. In parallel, we will also study the problem of deriving relevant (with respect to safety or optimization) properties on dataflow programs with array iterators.

Obviously, large applications are not limited to assembly of compute kernels. Our languages and formalism definitions (3.1) and analyses (3.2) must also be able to deal with general programs. Our targets also include generalist programs with complex behaviors such as recursive programs operating on arrays, lists and trees; worklist algorithms (lists are not handled within the polyhedral domain). Analysis on these programs should be able to detect non licit memory accesses, memory consumption, hotspots, ..., and to prove functional properties.

The simulation activities (3.5) are both applied internally in CASH, to simulate intermediate representations, and for embedded systems. We are interested in Transaction-Level Models (TLM) of Systems-on-a-Chip (SoCs) including processors and hardware accelerators. TLM provides an abstract but executable model of the chip, with enough details to run the embedded software. We are particularly interested in models written in a loosely timed coding style. We plan to extend these to heterogeneous simulations including a SystemC/TLM part to model the numerical part of the chip, and other simulators to model physical parts of the system.

5. Highlights of the Year

5.1. Highlights of the Year

- CASH has been validated as a *équipe projet commune* (EPC) by the *comité des projets*.
- We designed a dataflow transformation which always recovers all the FIFO in our dataflow model (DPN) after a loop tiling [1], [9], a program transformation widely used in automatic parallelization. This is an important enabling transformation which reinforces DPN as an intermediate representation in the CASH HLS project.
- We obtained new results on the comparison between different forms of synchronisation on futures, bringing a better understanding on the impact dataflow synchronisation and future typing on program synchronisation.

6. New Software and Platforms

6.1. DCC

DPN C Compiler
KEYWORDS: Polyhedral compilation - Automatic parallelization - High-level synthesis
FUNCTIONAL DESCRIPTION: Dcc (Data-aware process network C compiler) analyzes a sequential regular program written in C and generates an equivalent architecture of parallel computer as a communicating process network (Data-aware Process Network, DPN). Internal communications (channels) and external communications (external memory) are automatically handled while fitting optimally the characteristics of the global memory (latency and throughput). The parallelism can be tuned. Dcc has been registered at the APP ("Agence de protection des programmes") and transferred to the XtremLogic start-up under an Inria license.

- Participants: Alexandru Plesco and Christophe Alias
- Contact: Christophe Alias

6.2. PoCo

Polyhedral Compilation Library
KEYWORDS: Polyhedral compilation - Automatic parallelization
FUNCTIONAL DESCRIPTION: PoCo (Polyhedral Compilation Library) is a compilation framework allowing to develop parallelizing compilers for regular programs. PoCo features many state-of-the-art polyhedral program analysis and a symbolic calculator on execution traces (represented as convex polyhedra). PoCo has been registered at the APP (“agence de protection des programmes”) and transferred to the XtremLogic start-up under an Inria licence.

- Participant: Christophe Alias
- Contact: Christophe Alias

6.3. MPPcodegen

Source-to-source loop tiling based on MPP
KEYWORDS: Source-to-source compiler - Polyhedral compilation

FUNCTIONAL DESCRIPTION: MPPcodegen applies a monoparametric tiling to a C program enriched with pragmas specifying the tiling and the scheduling function. The tiling can be generated by any convex polyhedron and translation functions, it is not necessarily a partition. The result is a C program depending on a scaling factor (the parameter). MPPcodegen relies on the MPP mathematical library to tile the iteration sets.

- Partner: Colorado State University
- Contact: Christophe Alias
- URL: http://foobar.ens-lyon.fr/mppcodegen/

7. New Results

7.1. Monoparametric Tiling of Polyhedral Programs

Participant: Christophe Alias.

Tiling is a crucial program transformation, adjusting the ops-to-bytes balance of codes to improve locality. Like parallelism, it can be applied at multiple levels. Allowing tile sizes to be symbolic parameters at compile time has many benefits, including efficient autotuning, and run-time adaptability to system variations. For polyhedral programs, parametric tiling in its full generality is known to be non-linear, breaking the mathematical closure properties of the polyhedral model. Most compilation tools therefore either perform fixed size tiling, or apply parametric tiling in only the final, code generation step.

We introduce monoparametric tiling, a restricted parametric tiling transformation. We show that, despite being parametric, it retains the closure properties of the polyhedral model. We first prove that applying monoparametric partitioning (i) to a polyhedron yields a union of polyhedra, and (ii) to an affine function produces a piecewise-affine function. We then use these properties to show how to tile an entire polyhedral program. Our monoparametric tiling is general enough to handle tiles with arbitrary tile shapes that can tesselate the iteration space (e.g., hexagonal, trapezoidal, etc). This enables a wide range of polyhedral analyses and transformations to be applied.

This is a joint work with Guillaume Iooss (Inria Parkas) and Sanjay Rajopadhye (Colorado State University).

This work is under submission [12].

7.2. Improving Communication Patterns in Polyhedral Process Networks

Participant: Christophe Alias.

Process networks are a natural intermediate representation for HLS and more generally automatic parallelization. Compiler optimizations for parallelism and data locality restructure deeply the execution order of the processes, hence the read/write patterns in communication channels. This breaks most FIFO channels, which have to be implemented with addressable buffers. Expensive hardware is required to enforce synchronizations, which often results in dramatic performance loss. In this paper, we present an algorithm to partition the communications so that most FIFO channels can be recovered after a loop tiling, a key optimization for parallelism and data locality. Experimental results show a drastic improvement of FIFO detection for regular kernels at the cost of a few additional storage. As a bonus, the storage can even be reduced in some cases.

This work has been published in the HIP3ES workshop [1]

7.3. FIFO Recovery by Depth-Partitioning is Complete on Data-aware Process Networks

Participant: Christophe Alias.

In this paper, we build on our algorithm for FIFO recovery based on depth partitioning. We describe a class of process networks where the algorithm can recover all the FIFO channels. We point out the limitations of the algorithm outside of that class. Experimental results confirm the completeness of the algorithm on the class and reveal good performance outside of the class.

This work is under submission [9]

7.4. Parallel code generation of synchronous programs for a many-core architecture

Participant: Matthieu Moy.

Embedded systems tend to require more and more computational power. Many-core architectures are good candidates since they offer power and are considered more time predictable than classical multi-cores. Data-flow Synchronous languages such as Lustre or Scade are widely used for avionic critical software. Programs are described by networks of computational nodes. Implementation of such programs on a many-core architecture must ensure a bounded response time and preserve the functional behavior by taking interference into account. We consider the top-level node of a Lustre application as a software architecture description where each sub-node corresponds to a potential parallel task. Given a mapping (tasks to cores), we automatically generate code suitable for the targeted many-core architecture. This minimizes memory interferences and allows usage of a framework to compute the Worst-Case Response Time.

This is a joint work with Amaury Graillat, Pascal Raymond (IMAG) and Benoît Dupont de Dinechin (Kalray).

This work has been published at the DATE conference [6].

7.5. Estimation of the Impact of Architectural and Software Design Choices on Dynamic Allocation of Heterogeneous Memories

Participant: Matthieu Moy.

Reducing energy consumption is a key challenge to the realization of the Internet of Things. While emerging memory technologies may offer power reduction, they come with major drawbacks such as high latency or limited endurance. As a result, system designers tend to juxtapose several memory technologies on the same chip. This paper studies the interactions between dynamic memory allocation and architectural choices regarding this heterogeneity. We provide cycle accurate simulations of embedded platforms with various memory technologies and we show that different dynamic allocation strategies have a major impact on performance. We demonstrate that interesting performance gains can be achieved even for a low fraction of heap objects in fast memory, but only with a clever data placement strategy between memory banks.

This is a joint work with Tristan Delizy, Kevin Marquet, Tanguy Risset, Guillaume Salagnac (Inria Socrate) and Stéphane Gros (eVaderis).

This work has been published at the French Compas workshop [8] and the RSP Symposium [2].

7.6. Dataflow-explicit futures

Participant: Ludovic Henrio.

A future is a place-holder for a value being computed, and we generally say that a future is resolved when the associated value is computed. In existing languages futures are either implicit, if there is no syntactic or typing distinction between futures and non-future values, or explicit when futures are typed by a parametric type and dedicated functions exist for manipulating futures. We defined a new form of future, named data-flow explicit futures [38], with specific typing rules that do not use classical parametric types. The new futures allow at the same time code reuse and the possibility for recursive functions to return futures like with implicit futures, and let the programmer declare which values are futures and where synchronisation occurs, like with explicit futures. We prove that the obtained programming model is as expressive as implicit futures but exhibits a different behaviour compared to explicit futures. The current status of this work is the following:

- A paper showing formally the difference between implicit and explicit futures is under submission
- We are working with collaborators from University of Uppsala and University of Oslo on the design of programming constructs mixing implicit and dataflow-explicit futures
- Amaury Maillé will do his internship in the Cash team (advised by Matthieu Moy and Ludovic Henrio), working on an implementation of dataflow-explicit futures and further experiments with the model.

7.7. Locally abstract globally concrete semantics

Participant: Ludovic Henrio.

This research direction aims at designing a new way to write semantics for concurrent languages. The objective is to design semantics in a compositional way, where each primitive has a local behavior, and to adopt a style much closer to verification frameworks so that the design of an automatic verifier for the language is easier. The local semantics is expressed in a symbolic and abstract way, a global semantics gathers the abstract local traces and concretizes them. We have a reliable basis for the semantics of a simple language (a concurrent while language) and for a complex one (ABS), but the exact semantics and the methodology for writing it is still under development, we expect to submit a journal article during 2019 on the subject.

This is a joint with Reiner Hähnle (TU Darmstadt), Einar Broch Johnsen, Crystal Chang Din, Lizeth Tapia Tarifa (Univ Oslo), Ka I Pun (Univ Oslo and Univ of applied science).

7.8. Memory consistency for heterogeneous systems

Participant: Ludovic Henrio.

Together with Christoph Kessler (Linköping University), we worked on the formalization of the cache coherency mechanism used in the VectorPU library developed at Linköping University. Running a program on disjoint memory spaces requires to address memory consistency issues and to perform transfers so that the program always accesses the right data. Several approaches exist to ensure the consistency of the memory accessed, we are interested here in the verification of a declarative approach where each component of a computation is annotated with an access mode declaring which part of the memory is read or written by the component. The programming framework uses the component annotations to guarantee the validity of the memory accesses. This is the mechanism used in VectorPU, a C++ library for programming CPU-GPU heterogeneous systems and this article proves the correctness of the software cache-coherence mechanism used in the library. Beyond the scope of VectorPU, this article can be considered as a simple and effective formalisation of memory consistency mechanisms based on the explicit declaration of the effect of each component on each memory space. This year, we have the following new results:

- provided a formalization showing the correctness of VectorPU approach (published in 4PAD 2018, a symposium affiliated to HPCS).
- extended the work to support the manipulation of overlapping array (submitted as an extended version of the 4PAD paper)

We now plan to extend the work with support for concurrency.

7.9. PNets: Parametrized networks of automata

Participant: Ludovic Henrio.

pNets (parameterised networks of synchronised automata) are semantic objects for defining the semantics of composition operators and parallel systems. We have used pNets for the behavioral specification and verification of distributed components, and proved that open pNets (i.e. pNets with holes) were a good formalism to reason on operators and parameterized systems. This year, we have the following new results:

- A weak bisimulation theory for open pNets is under development (a strong isimulation had already been defined in the past) and its properties are being proven, especially in terms of compositionality. This work is realized with Eric Madelaine (Inria Sophia-Antipolis) and Rabéa Ameur Boulifa (Telecom ParisTech).
- A translation from BIP model to open pNets is being formalized and encoded, this work is done in collaboration with Simon Bliudze (Inria Lille).

These works are under progress and should be continued in 2019.

7.10. Decidability results on the verification of phaser programs

Participant: Ludovic Henrio.

Together with Ahmed Rezine and Zeinab Ganjei (Linköping University) we investigated the possibility to analyze programs with phasers (a construct for synchronizing processes that generalizes locks, barrier, and publish-subscribe patterns). They work with signal and wait messages from the processes (comparing the number of wait and signal received to synchronize the processes). We proved that in many conditions, if the number of phasers or processes cannot be bounded, or if the difference between the number of signal and the number of wait signal is unbounded, then many reachability problems are undecidable. We also proposed fragments where these problems become decidable, and proposed an analysis algorithm in these cases. The results are currently under review in a conference.

7.11. Practicing Domain-Specific Languages: From Code to Models

Participant: Laure Gonnord.

Together with Sebastien Mosser, we proposed a new Domain- Specific Language course at the graduate level whose objectives is to reconciliate concepts coming from Language Design as well as Modeling domains. We illustrate the course using the reactive systems application domain, which prevents us to fall back in a toy example pitfall. This paper describes the nine stages used to guide students through a journey starting at low-level C code to end with the usage of a language design workbench. This course was given as a graduate course available at Université Côte d'Azur (8 weeks, engineering-oriented) and École Normale Supérieure de Lyon (13 weeks, research-oriented).

The results have been published in a national software engineering conference [4] and the Models Educator Symposium [5].

7.12. Polyhedral Dataflow Programming: a Case Study

Participant: Laure Gonnord.

With Lionel Morel and Romain Fontaine (Insa Lyon), we have studied the benefits of jointly using polyhedral compilation with dataflow languages. We have proposed to expend the parallelization of dataflow programs by taking into account the parallelism exposed by loop nests describing the internal behavior of the program's agents. This approach is validated through the development of a prototype toolchain based on an extended version of the SigmaC language. We demonstrated the benefit of this approach and the potentiality of further improvements on several case studies.

The results have been published in the Sbac-PAD conference on High Performance computing [3].

7.13. Semantic Array Dataflow Analysis

Participants: Laure Gonnord, Paul Iannetta.

Together with Lionel Morel (Insa/CEA) and Tomofumi Yuki (Inria, Rennes), we revisited the polyhedral model's key analysis, dependency analysis. The semantic formulation we propose allows a new definition of the notion of dependency and the computation of the dependency set. As a side effect, we propose a general algorithm to compute *an over-approximation of* the dependency set of general imperative programs.

We argue that this new formalization will later allow for a new vision of the polyhedral model in terms of semantics, which will help us fully characterize its expressivity and applicability. We also believe that abstract semantics will be the key for designing an approximate abstract model in order to enhance the applicability of the polyhedral model.

The results is published in a research report [11].

7.14. Static Analysis Of Binary Code With Memory Indirections Using Polyhedra

Participant: Laure Gonnord.

Together with Clement Ballabriga, Julien Forget, Giuseppe Lipari, and Jordy Ruiz (University of Lille), we proposed a new abstract domain for static analysis of binary code. Our motivation stems from the need to improve the precision of the estimation of the Worst-Case Execution Time (WCET) of safety-critical real-time code. WCET estimation requires computing information such as upper bounds on the number of loop iterations, unfeasible execution paths, etc. These estimations are usually performed on binary code, mainly to avoid making assumptions on how the compiler works. Our abstract domain, based on polyhedra and on two mapping functions that associate polyhedra variables with registers and memory, targets the precise computation of such information. We prove the correctness of the method, and demonstrate its effectiveness on benchmarks and examples from typical embedded code.

The results have been accepted to VMCAI'19 on Model Checking and Abstract Interpretation [7].

7.15. Polyhedral Value Analysis as Fast Abstract Interpretation

Participant: Laure Gonnord.

Together with Tobias Grosser, (ETH Zurich, Switzerland), Siddhart Bhat, (IIIT Hydrabad, India), Marcin Copik (ETH Zurich, Switzerland), Sven Verdoolaege (Polly Labs, Belgium) and Torsten Hoefler (ETH Zurich, Switzerland), we tried to bridge the gap between the well founded classical abstract interpretation techniques and their usage in production compilers.

We formulate the polyhedral value analysis (a classical algorithm in production compilers like LLVM, scalar evolution based on Presburger set as abstract interpretation), present a set of fast join operators, and show that aggressively falling back to top (rather than continuing with approximations) results in a scalable analysis. By formally describing the required analysis, we provide the necessary theoretical foundations for analysing large program systems with hundred thousands of loops and complex control flow structures at a precision high enough to cater for high-precision users such as polyhedral optimization frameworks, at a compile-time cost comparable with just compiling the application.

The paper is under redaction process.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

- Matthieu Moy submitted an ANR project as scientific coordinator entitled "Distributed Efficient Architecture for the Rapid (Co)simulation of Multiphysics Objects" (48 months, partners Verimag, TIMA and LIP6).
- Laure Gonnord's "Jeune Chercheur" ANR, CODAS, has started in January 2018 (42 months).

8.1.2. Scientific Advising

- Christophe Alias is scientific advisor (concours scientifique, 20%) for the XTREMLOGIC start-up.

8.2. International Initiatives

8.2.1. Informal International Partners

- Christophe Alias has regular collaborations with Sanjay Rajopadhye from Colorado State University, USA (3 publications, one publication submitted 7.1).

- Ludovic Henrio has regular collaborations with university of Linköping (one publication last year, 2 submissions); University of Oslo, University of Uppsala, and TU Darmstadt on active objects (2 publications being written); Chalmers university and Univ of Twente (one publication submitted).
- Laure Gonnord has regular collaborations with Fernando Pereira from UFMG, Brasil (5 publications in total, last in 2017). In 2018 she has began a collaboration with Tobias Grösser, from ETH Zurich.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Member of the Organizing Committees

- Laure Gonnord animates the french compilation community since 2010 (http://compilfr.ens-lyon.fr
- Laure Gonnord has coorganised (with other members of the LIP lab) the doctoral school EJCP (École jeunes chercheur.se.s en programmation), in June 2018.

9.1.2. Chair of Conference Program Committees

- Ludovic Henrio was one of the chairs of the **ICE** 2018 workshop (Interaction and Concurrency Experiences), he will also be chairing **ICE** 2019.

9.1.3. Member of the Conference Program Committees

- Christophe Alias is a PC member of **HiP3ES** 2018 – 6th High Performance Energy Efficient Embedded Systems Workshop. He will be a PC member for **COMPAS** 2019 – Conférence d'informatique en Parallélisme, Architecture et Système.
- Matthieu Moy is a PC member of **DUHDe** 2018 — 5th Workshop on Design Automation for Understanding Hardware Designs.
- Laure Gonnord is a PC member of **TAPAS** 2018 - Ninth Workshop on Tools for Automatic Program Analysis, and **VMCAI** 2018 - 19th International Conference on Verification, Model Checking, and Abstract Interpretation. She will be a PC member of **CAV** 2019 Conference on Computer-Aided Verification.
- Ludovic Henrio is a PC member of **4PAD** 2018 (Formal Approaches to Parallel and Distributed Systems), **FASE** 2019 (Fundamental Approaches to Software Engineering). He will be a PC member of **Coordination** 2019 (Conference on Coordination Models and Languages), **ACSD** 2019 (Application of Concurrency to System Design).

9.1.4. Reviewer

- Christophe Alias was reviewer for **HCW** 2018 – 27th International Workshop on Heterogeneity in Computing.
- Matthieu Moy was reviewer for **DATE** 2018 - Design, Automation and Test in Europe, **MCSoC** 2019 - International Symposium on Embedded Multicore/Many-core Systems-on-Chip
- Laure Gonnord was reviewer for **STACS** 2018 - Symposium of Theorical Aspects of Computer Science.
- Ludovic Henrio was reviewer for **FM'2018**.

9.1.5. Journal

9.1.5.1. Reviewer - Reviewing Activities

- Christophe Alias was reviewer for ACM Transactions on Architectures and Code Optimization (**TACO**, ACM) and **Proceedings of the IEEE**.

- Matthieu Moy was reviewer for the Microprocessors and Microsystems Journal (**MICPRO**, Elsevier).
- Ludovic Henrio was reviewer for **JLAMP** (Journal of Logical and Algebraic Methods in Programming), Elsevier.

9.1.6. Invited Talks

- Matthieu Moy presented a talk “Response Time Analysis of Dataflow Applications on a Many-Core Processor with Shared-Memory and Network-on-Chip” for the 30 years of the LIP laboratory, November 2018.
- Laure Gonnord gave two invited talks at the workshop **FAC Days** (Toulouse) and at the International Conference **LOPSTR 2018** “Experiences in designing scalable static analyses”.
- Ludovic Henrio gave an invited talk at **HLPP 2018** “An Overview of (some) Active-object Languages” and at **PASS** 2018 workshop “SafePlace: Trustable Virtual Machine Scheduling”.

9.1.7. Research Administration

- Laure Gonnord is member of the doctoral commitee of the Inria Grenoble Rhône Alpes center.
- Laure Gonnord is elected member of the LIP council and the Fédération d’Informatique de Lyon council.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence:

- Christophe Alias, Compilation, CM+TD, 27h, 3A, INSA Centre Val de Loire.
- Laure Gonnord, Algorithmic, C++ Programming, TP: L2, UCBL
- Laure Gonnord, Operating Systems, CM+TD+TP, 24h, L2, UCBL
- Laure Gonnord, Formal Languages, TD+TP, 15h, L3, UCBL
- Laure Gonnord, Logics, TD+TP, 18h, L3, UCBL.
- Matthieu Moy, Concurrent Programming, CM+TD+TP, 57h, L3, UCBL.
- Matthieu Moy, Recursive Programming, TD+TP, 48h, L1, UCBL.
- Matthieu Moy, Software Engineering, CM+TD+TP, 25h, M1, UCBL.
- Matthieu Moy, Git, CM+TP: 12h, L3, UCBL.
- Paul Iannetta, ACM, TD, 32h, L3, ENS de Lyon.
- Julien Braine, Théorie de la programmation, TD/TP, 32h, L3, ENS de Lyon.
- Julien Braine, Projet 1, TP, 32h, L3, ENS de Lyon.

Master:

- Christophe Alias, Compiler optimizations for embedded applications, CM+TD, 27h, 4A, INSA Centre Val de Loire.
- Christophe Alias and Matthieu Moy, Hardware Compilation and Simulation, CM+TD, 36h, M2 Informatique Fondamentale, ENS de Lyon.
- Laure Gonnord, Compilation and Program Analysis, CM, 28h, M1, ENS de Lyon.
- Laure Gonnord, Compilation and program transformations, CM+TD+TP, 35h, M1, UCBL.
- Laure Gonnord, Real Time Systems, CM+TD+TP, 30h, M1, UCBL.
- Laure Gonnord (25%), with Sebastien Mosser, Software Engineering and Compilation, CM+TP, 36h, M2 Informatique Fondamentale, ENS de Lyon.

- Laure Gonnord, Graphs, Complexity, Algorithmics, M1 MEEF (CAPES prepa), CM+TD+TP+oral training, 30h, UCBL.
- Matthieu Moy, Compilation and Program Analysis, TP, 16h, M1, ENS de Lyon.
- Matthieu Moy, Compilation and program transformations, TD+TP, 25h, M1, UCBL.
- Paul Iannetta, Projet intégré, 28h, M1, ENS de Lyon.
- Ludovic Henrio, Distributed Systems: an algorithmic approach, CM+TD, 7h, M2 Specialite IFI (Ingénierie et Fondements de l'Informatique), parcours CSSR, and UBINET, Université de Nice Sophia-Antipolis.

9.2.2. Supervision

- PhD: Amaury Graillat, "Parallel Code Generation of Synchronous Programs for a Many-core Architecture", Univ. Grenoble Alpes, defended on November 16th 2018, supervised by Matthieu Moy (LIP), Pascal Raymond (Verimag) and Benoît Dinechin (Kalray).
- PhD in progress: Gabriel Busnot, "Accélération SystemC pour la co-simulation multi-physique et la simulation de modèles hétérogènes en complexité", Univ. Lyon 1, started in october 2017, supervised by Matthieu Moy (LIP) and Tanguy Sassolas (CEA-LIST).
- PhD in progress: Tristan Delizy, "Dynamic Memory Management For Embedded Non-Volatile Memory", INSA Lyon, started in October 2016, supervised by Guillaume Salagnac (CITI), Tanguy Risset (CITI), Kevin Marquet (CITI) and Matthieu Moy (LIP).
- PhD in progress (from Sept. 2018): Paul Iannetta "Complex data structures scheduling for optimizing compilers", supervised by Lionel Morel (CITI/CEA) and Laure Gonnord (LIP).
- PhD in progress (from Sept. 2018): Julien Braine "Horn Clauses as an Efficient Intermediate Representation for Data Structure Verification", supervised by David Monniaux (CNRS/Verimag) and Laure Gonnord (LIP).
- PhD in progress: Pierre Leca, "Distributed BSP: Active Objects for BSPlib programs", CIFRE Huawei/UNS, started in August 2017, supervised by Gaëtan Hains (Huawei), Wijnand Suijlen (Huawei), Françoise Baude (UNS./I3S), Ludovic Henrio (LIP).

9.2.3. Juries

- Christophe Alias was an expert for the midterm PhD evaluation of Hang Yu from Université Grenoble-Alpes. Hang Yu is supervised by Michaël Perrin and David Monniaux.
- Laure Gonnord was an expert for the midterm PhD evaluation of Sébastien Bonnieux from University Nice Côte d'Azur. Sébastien Bonnieux was supervised by Sébastien Mosser and Mireille Blay-Fornarino.
- Laure Gonnord was reviewer for the PhD of Vincent Botbol from Sorbonne Université entitled "Analyse Statique de programmes concurrents avec variables numériques" and supervised by Emmanuel Chailloux and Tristan Le Gall.
- Laure Gonnord was external jury member for the PhD of Hoby Rakotoarivelo from Université Paris-Saclay entitled "Approche de co-design de noyaux irréguliers sur accélérateurs manycore. Application au cas du remaillage adaptatif pour le calcul intensif" and supervised by Franck Ledoux et Franck Pommereau.
- Laure Gonnord was local jury member for the PhD of Mohammed Amer from Université de Lyon entitled "Centralized Optimization of the Association in IEEE 802.11 Networks" and supervised by Anthony Busson and Isabelle Guérin-Lassous.
- Matthieu Moy was an expert for the midterm PhD evaluation of Joumana Lagha from Ecole centrale de Nantes. Joumana Lagha is supervised by Prof. Olivier H. Roux, Sebastien Faucou and Jean-luc Bechennec.

- Matthieu Moy was reviewer for the Ph.D of Benjamin Rouxel from Université de Rennes 1 entitled “Minimising communication costs impact when scheduling real-time applications on multi-core architectures” and supervised by Isabelle Puaut and Steven Derrien.

9.2.4. Internships

- Bilel Aouadhi, a last year engineer student from Faculté des sciences de Tunis, worked from April 2018 to July 2018 on the implementation of a visualization tool for Data-aware process networks. His was supervised by Christophe Alias and Matthieu Moy, his internship was funded by Université Lyon 1.
- Ligia Novacean, a L3 student from University of Cluj-Napoca (Romania), worked from July 2018 to September 2018 on a DPN-to-C translator for the HLS tool of Xilinx, VivadoHLS. She was supervised by Christophe Alias and Matthieu Moy, her internship was funded by Inria.
- Alexandra Dobre, a L3 student from University of Cluj-Napoca (Romania), worked from July 2018 to September 2018 on the generation of a SystemC simulator from a Data-aware process network. She was supervised by Matthieu Moy and Christophe Alias, her internship was funded by Inria.
- Arthur Gontier, a L3 student from University of Nantes, worked from April 2018 to July 2018 on the formal functional verification of Lustre code with Horn Clauses. He was supervised by Lionel Morel and Laure Gonnord. His internship was founded by the Codas ANR.
- Paul Iannetta, a M2 student from ENS de Lyon, worked from March 2018 to July 2018 on a semantic formalisation of the polyhedral model. He was supervised by Lionel Morel and Laure Gonnord. His internship was founded by the Codas ANR.

9.3. Popularization

9.3.1. Education

- Laure Gonnord is part of the local organisation of the Computer Science preparation for the Agregation examination for future maths teachers (MEEF).

9.3.2. Interventions

- Talk at “Campus du libre” by Matthieu Moy, Doua Lyon, “Pourquoi et comment se lancer dans le libre quand on est étudiant (ou pas) ?”, November 2018.

9.3.3. Internal action

- Café développeur by Matthieu Moy, “Utilisation avancée de Git” at LIRIS (2 sessions), October 2018.

9.3.4. Creation of media or tools for science outreach

- Video “Mon équipe en 180 secondes” by Matthieu Moy for the CASH team.

10. Bibliography

Publications of the year

International Conferences with Proceedings

[1] C. ALIAS.*Improving Communication Patterns in Polyhedral Process Networks*, in "HIP3ES 2018 - Sixth International Workshop on High Performance Energy Efficient Embedded Systems", Manchester, United Kingdom, January 2018, p. 1-6, https://hal.inria.fr/hal-01725143

[2] T. DELIZY, S. GROS, K. MARQUET, M. MOY, T. RISSET, G. SALAGNAC.*Estimating the Impact of Architectural and Software Design Choices on Dynamic Allocation of Heterogeneous Memories*, in "RSP 2018 - 29th International Symposium on Rapid System Prototyping", Turin, Italy, October 2018, p. 1-7, https://hal.archives-ouvertes.fr/hal-01891599

[3] R. FONTAINE, L. GONNORD, L. MOREL.*Polyhedral Dataflow Programming: a Case Study*, in "SBAC-PAD 2018 - 30th International Symposium on Computer Architecture and High-Performance Computing", Lyon, France, IEEE, September 2018, p. 1-9, https://hal-cea.archives-ouvertes.fr/cea-01855997

[4] L. GONNORD, S. MOSSER.*Du code aux modèles, des modèles au code: enseigner les langages dédiés (DSL)*, in "CIEL 2018 : 7ème Conférence en IngénieriE du Logiciel", Grenoble, France, June 2018, p. 1-4, https://hal.archives-ouvertes.fr/hal-01816239

[5] L. GONNORD, S. MOSSER.*Practicing Domain-Specific Languages: From Code to Models*, in "14th Educators Symposium at MODELS 2018", Copenaghen, Denmark, October 2018, p. 1-8 [*DOI :* 10.1145/3270112.3270116], https://hal.archives-ouvertes.fr/hal-01865448

[6] A. GRAILLAT, M. MOY, P. RAYMOND, B. DUPONT DE DINECHIN.*Parallel Code Generation of Synchronous Programs for a Many-core Architecture*, in "DATE 2018 - Design, Automation and Test in Europe", Dresden, Germany, IEEE, March 2018, p. 1139-1142 [*DOI :* 10.23919/DATE.2018.8342182], https://hal.inria.fr/hal-01667594

Conferences without Proceedings

[7] C. BALLABRIGA, J. FORGET, L. GONNORD, G. LIPARI, J. RUIZ.*Static Analysis Of Binary Code With Memory Indirections Using Polyhedra*, in "International Conference on Verification, Model Checking, and Abstract Interpretation (VMCAI'19)", Lisbon, Portugal, January 2019, https://hal.archives-ouvertes.fr/hal-01939659

[8] T. DELIZY, S. GROS, K. MARQUET, M. MOY, T. RISSET, G. SALAGNAC.*Quels objets en NVRAM ? Placement en mémoires de travail hétérogènes*, in "Compas 2018 - Conférence d'informatique en Parallélisme, Architecture et Système", Toulouse, France, July 2018, p. 1-8, https://hal.archives-ouvertes.fr/hal-01891398

Research Reports

[9] C. ALIAS.*FIFO Recovery by Depth-Partitioning is Complete on Data-aware Process Networks*, Inria Grenoble - Rhone-Alpes, June 2018, n° RR-9187, https://hal.inria.fr/hal-01818585

[10] L. GONNORD, P. IANNETTA, L. MOREL.*Semantic Polyhedral Model for Arrays and Lists*, Inria Grenoble Rhône-Alpes ; UCBL ; LIP - ENS Lyon ; CEA List, June 2018, n° RR-9183, p. 1-28, https://hal.archives-ouvertes.fr/hal-01815759

[11] P. IANNETTA, L. GONNORD, L. MOREL, T. YUKI.*Semantic Array Dataflow Analysis*, Inria Grenoble Rhône-Alpes, December 2018, n° RR-9232, https://hal.archives-ouvertes.fr/hal-01954396

[12] G. IOOSS, C. ALIAS, S. RAJOPADHYE.*Monoparametric Tiling of Polyhedral Programs*, Inria Grenoble - Rhone-Alpes, December 2018, https://hal.inria.fr/hal-01952593

References in notes

[13] *IEEE 1666 Standard: SystemC Language Reference Manual*, Open SystemC Initiative, 2011, http://www.accellera.org/

[14] *OSCI TLM-2.0 Language Reference Manual*, Open SystemC Initiative (OSCI), June 2008, http://www.accellera.org/downloads/standards

[15] C. ALIAS, A. DARTE, P. FEAUTRIER, L. GONNORD.*Multi-dimensional Rankings, Program Termination, and Complexity Bounds of Flowchart Programs*, in "International Static Analysis Symposium (SAS'10)", 2010

[16] C. ALIAS, A. PLESCO.*Method of Automatic Synthesis of Circuits, Device and Computer Program associated therewith*, April 2014, Patent FR1453308

[17] C. ALIAS, A. PLESCO.*Data-aware Process Networks*, Inria - Research Centre Grenoble – Rhône-Alpes, June 2015, n^o RR-8735, 32, https://hal.inria.fr/hal-01158726

[18] C. ALIAS, A. PLESCO.*Optimizing Affine Control with Semantic Factorizations*, in "ACM Transactions on Architecture and Code Optimization (TACO) ", December 2017, vol. 14, n^o 4, 27

[19] I. AMER, C. LUCARZ, G. ROQUIER, M. MATTAVELLI, M. RAULET, J.-F. NEZAN, O. DEFORGES.*Reconfigurable video coding on multicore*, in "Signal Processing Magazine, IEEE", 2009, vol. 26, n^o 6, p. 113–123, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5230810

[20] S. ANANIAN.*The Static Single Information Form*, MIT, September 1999

[21] C. B. AOUN, L. ANDRADE, T. MAEHNE, F. PÊCHEUX, M.-M. LOUËRAT, A. VACHOUXY.*Pre-simulation elaboration of heterogeneous systems: The SystemC multi-disciplinary virtual prototyping approach*, in "Embedded Computer Systems: Architectures, Modeling, and Simulation (SAMOS), 2015 International Conference on", IEEE, 2015, p. 278–285

[22] P. AUBRY, P.-E. BEAUCAMPS, F. BLANC, B. BODIN, S. CARPOV, L. CUDENNEC, V. DAVID, P. DORÉ, P. DUBRULLE, B. DUPONT DE DINECHIN, F. GALEA, T. GOUBIER, M. HARRAND, S. JONES, J.-D. LESAGE, S. LOUISE, N. MOREY CHAISEMARTIN, T. H. NGUYEN, X. RAYNAUD, R. SIRDEY.*Extended Cyclostatic Dataflow Program Compilation and Execution for an Integrated Manycore Processor*, in "Alchemy 2013 - Architecture, Languages, Compilation and Hardware support for Emerging ManYcore systems", Barcelona, Spain, Proceedings of the International Conference on Computational Science, ICCS 2013, June 2013, vol. 18, p. 1624-1633 [*DOI :* 10.1016/J.PROCS.2013.05.330], https://hal.inria.fr/hal-00832504

[23] D. BECKER, M. MOY, J. CORNET.*Parallel Simulation of Loosely Timed SystemC/TLM Programs: Challenges Raised by an Industrial Case Study*, in "MDPI Electronics", 2016, vol. 5, n^o 2, 22 [*DOI :* 10.3390/ELECTRONICS5020022], https://hal.archives-ouvertes.fr/hal-01321055

[24] D. CAROMEL, L. HENRIO.*A Theory of Distributed Objects*, Springer-Verlag, 2004

[25] P. COUSOT, R. COUSOT.*Abstract interpretation: A unified lattice model for static analysis of programs by construction or approximation of fixpoints*, in "4th ACM Symposium on Principles of Programming Languages (POPL'77)", Los Angeles, January 1977, p. 238-252

[26] F. DE BOER, V. SERBANESCU, R. HÄHNLE, L. HENRIO, J. ROCHAS, C. C. DIN, E. BROCH JOHNSEN, M. SIRJANI, E. KHAMESPANAH, K. FERNANDEZ-REYES, A. M. YANG.*A Survey of Active Object Languages*, in "ACM Comput. Surv.", October 2017, vol. 50, n^{o} 5, p. 76:1–76:39, http://doi.acm.org/10.1145/3122848

[27] M. DURANTON, D. BLACK-SCHAFFER, K. D. BOSSCHERE, J. MAEBE.*The HIPEAC VISION FOR ADVANCED COMPUTING IN HORIZON 2020, https://www.hipeac.net/v13*, 2013, https://www.hipeac.net/v13

[28] P. FEAUTRIER.*Scalable and Structured Scheduling*, in "International Journal of Parallel Programming", October 2006, vol. 34, n^{o} 5, p. 459–487

[29] P. FEAUTRIER.*Dataflow analysis of array and scalar references*, in "International Journal of Parallel Programming", 1991, vol. 20, n^{o} 1, p. 23–53

[30] P. FEAUTRIER, A. GAMATIÉ, L. GONNORD.*Enhancing the Compilation of Synchronous Dataflow Programs with a Combined Numerical-Boolean Abstraction*, in "CSI Journal of Computing", 2012, vol. 1, n^{o} 4, p. 8:86–8:99, http://hal.inria.fr/hal-00860785

[31] K. FERNANDEZ-REYES, D. CLARKE, E. CASTEGREN, H.-P. VO.*Forward to a Promising Future*, in "Conference proceedings COORDINATION 2018", 2018

[32] R. FONTAINE, L. GONNORD, L. MOREL.*Polyhedral Dataflow Programming: a Case Study*, in "SBAC-PAD 2018 - 30th International Symposium on Computer Architecture and High-Performance Computing", Lyon, France, IEEE, September 2018, p. 1-9, https://hal-cea.archives-ouvertes.fr/cea-01855997

[33] R. FONTAINE, L. GONNORD, L. MOREL.*Polyhedral Dataflow Programming: a Case Study*, in "International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD)", IEEE, September 2018

[34] L. GONNORD, P. IANNETTA, L. MOREL.*Semantic Polyhedral Model for Arrays and Lists*, Inria Grenoble Rhône-Alpes ; UCBL ; LIP - ENS Lyon ; CEA List, June 2018, n^{o} RR-9183, https://hal.archives-ouvertes.fr/hal-01815759

[35] M. I. GORDON.*Compiler techniques for scalable performance of stream programs on multicore architectures*, Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science, 2010

[36] O. HAKJOO, L. WONCHAN, H. KIHONG, Y. HONGSEOK, Y. KWANGKEUN.*Selective context-sensitivity guided by impact pre-analysis*, in "ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI '14, Edinburgh, United Kingdom - June 09 - 11, 2014", ACM, 2014, 49

[37] N. HALBWACHS, P. CASPI, P. RAYMOND, D. PILAUD.*The synchronous data flow programming language LUSTRE*, in "Proceedings of the IEEE", Sep 1991, vol. 79, n^{o} 9, p. 1305-1320

[38] L. HENRIO.*Data-flow Explicit Futures*, I3S, Université Côte d'Azur, April 2018, https://hal.archives-ouvertes.fr/hal-01758734

[39] G. KAHN.*The semantics of a simple language for parallel programming*, in "Information processing", North-Holland, 1974

[40] A. KRIZHEVSKY, I. SUTSKEVER, G. E. HINTON.*Imagenet classification with deep convolutional neural networks*, in "Advances in neural information processing systems", 2012, p. 1097–1105

[41] M. MAALEJ KAMMOUN.*Low-cost memory analyses for efficient compilers*, Université Lyon 1, 2017, Thèse de doctorat, Université Lyon1, http://www.theses.fr/2017LYSE1167

[42] M. MAALEJ KAMMOUN, V. PAISANTE, P. RAMOS, L. GONNORD, F. PEREIRA.*Pointer Disambiguation via Strict Inequalities*, in "Code Generation and Optimisation", Austin, United States, February 2017, https://hal.archives-ouvertes.fr/hal-01387031

[43] M. MOY.*Parallel Programming with SystemC for Loosely Timed Models: A Non-Intrusive Approach*, in "DATE", Grenoble, France, March 2013, 9, https://hal.archives-ouvertes.fr/hal-00761047

[44] V. PAISANTE, M. MAALEJ KAMMOUN, L. BARBOSA, L. GONNORD, F. M. Q. PEREIRA.*Symbolic Range Analysis of Pointers*, in "International Symposium of Code Generation and Optmization", Barcelon, Spain, March 2016, p. 791-809, https://hal.inria.fr/hal-01228928

[45] A. PLESCO.*Program Transformations and Memory Architecture Optimizations for High-Level Synthesis of Hardware Accelerators*, Ecole normale supérieure de lyon - ENS LYON, September 2010, https://tel.archives-ouvertes.fr/tel-00544349

[46] L.-N. POUCHET.*Polybench: The polyhedral benchmark suite*, 2012, http://www.cs.ucla.edu/~pouchet/software/polybench/

[47] W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERLING, B. P. FLANNERY.*Numerical recipes in C++*, in "The art of scientific computing", 2015

[48] P. QUINTON.*Automatic synthesis of systolic arrays from uniform recurrent equations*, in "ACM SIGARCH Computer Architecture News", 1984, vol. 12, n^{o} 3, p. 208–214

[49] H. RIHANI, M. MOY, C. MAIZA, R. I. DAVIS, S. ALTMEYER.*Response Time Analysis of Synchronous Data Flow Programs on a Many-Core Processor*, in "Proceedings of the 24th International Conference on Real-Time Networks and Systems", New York, NY, USA, RTNS '16, ACM, 2016, p. 67–76, http://doi.acm.org/10.1145/2997465.2997472

[50] H. N. W. SANTOS, I. MAFFRA, L. OLIVEIRA, F. PEREIRA, L. GONNORD.*Validation of Memory Accesses Through Symbolic Analyses*, in "Proceedings of the 2014 ACM International Conference on Object Oriented Programming Systems Languages And Applications (OOPSLA'14)", Portland, Oregon, United States, October 2014, http://hal.inria.fr/hal-01006209

[51] W. THIES.*Language and compiler support for stream programs*, Massachusetts Institute of Technology, 2009

[52] J. TRAVIS, J. KRING.*LabVIEW for everyone: graphical programming made easy and fun*, Prentice-Hall, 2007

[53] A. TURJAN.*Compiling Nested Loop Programs to Process Networks*, Universiteit Leiden, 2007

[54] N. VENTROUX, T. SASSOLAS.*A new parallel SystemC kernel leveraging manycore architectures*, in "Design, Automation & Test in Europe Conference & Exhibition (DATE), 2016", IEEE, 2016, p. 487–492

[55] S. VERDOOLAEGE.*Polyhedral Process Networks*, Handbook of Signal Processing Systems, Springer, 2010, p. 931–965

[56] S. WILLIAMS, A. WATERMAN, D. PATTERSON.*Roofline: an insightful visual performance model for multicore architectures*, in "Communications of the ACM", 2009, vol. 52, n^{o} 4, p. 65–76

[57] P. WILMOTT.*Quantitative Finance*, Wiley, 2006

[58] B. DA SILVA, A. BRAEKEN, E. H. D'HOLLANDER, A. TOUHAFI.*Performance modeling for FPGAs: extending the roofline model with high-level synthesis tools*, in "International Journal of Reconfigurable Computing", 2013, vol. 2013, 7

Project-Team Chroma

Cooperative and Human-aware Robot Navigation in Dynamic Environments

IN COLLABORATION WITH: Centre of Innovation in Telecommunications and Integration of services

IN PARTNERSHIP WITH:
Institut national des sciences appliquées de Lyon

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Robotics and Smart environments

Table of contents

Project-Team Chroma

Creation of the Team: 2015 March 01, updated into Project-Team: 2017 December 01

Keywords:

Computer Science and Digital Science:

A1.5.2. - Communicating systems
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.3. - Reinforcement learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.8. - Deep learning
A5.1. - Human-Computer Interaction
A5.4.2. - Activity recognition
A5.4.4. - 3D and spatio-temporal reconstruction
A5.4.5. - Object tracking and motion analysis
A5.4.6. - Object localization
A5.4.7. - Visual servoing
A5.10.2. - Perception
A5.10.3. - Planning
A5.10.4. - Robot control
A5.10.5. - Robot interaction (with the environment, humans, other robots)
A5.10.6. - Swarm robotics
A5.10.7. - Learning
A5.11.1. - Human activity analysis and recognition
A6.1.2. - Stochastic Modeling
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A6.2.3. - Probabilistic methods
A6.2.6. - Optimization
A6.4.3. - Observability and Controlability
A6.4.6. - Optimal control
A8.2. - Optimization
A8.2.1. - Operations research
A8.2.2. - Evolutionary algorithms
A8.11. - Game Theory
A8.12. - Optimal transport
A9.2. - Machine learning
A9.5. - Robotics
A9.6. - Decision support
A9.7. - AI algorithmics
A9.9. - Distributed AI, Multi-agent
A9.10. - Hybrid approaches for AI

Other Research Topics and Application Domains:

B5.2.1. - Road vehicles
B5.6. - Robotic systems
B7.1.2. - Road traffic
B8.4. - Security and personal assistance

1. Team, Visitors, External Collaborators

Research Scientists

Christian Laugier [Inria, Emeritus, HDR]
Agostino Martinelli [Inria, Researcher]

Faculty Members

Olivier Simonin [Team leader, INSA Lyon, Professor, HDR]
Jilles Dibangoye [INSA Lyon, Associate Professor]
Laetitia Matignon [Univ de Claude Bernard, Associate Professor]
Anne Spalanzani [Univ Grenoble Alpes, Associate Professor, HDR]
Christian Wolf [INSA Lyon, Associate Professor, HDR]

External Collaborators

Fabrice Jumel [Ecole supérieure de chimie physique électronique de Lyon]
Jacques Saraydaryan [Ecole supérieure de chimie physique électronique de Lyon]

Technical Staff

Alexander Oliva [Inria, until Sep 2018]
Rabbia Asghar [Inria, from May 2018]
Stephane d'Alu [INSA Lyon]
Jean-Alix David [Inria]
Thomas Genevois [Inria]
Vincent Le Doze [Inria]
Jerome Lussereau [Inria]
Anshul Paigwar [Inria, from Oct 2018]
Lukas Rummelhard [Inria, from Sep 2018]
Blanche Baudouin [Inria, until May 2018]

PhD Students

Mathieu Barbier [Renault/Inria]
Edward Beeching [Inria, from Oct 2018]
Guillaume Bono [INSA Lyon]
Remi Cambuzat [Région, Univ. Grenoble Alpes]
Niranjan Deshpande [Inria, from Mar 2018]
Remy Grunblatt [Inria/DGA, Lyon 1]
Maria Kabtoul [Inria, from Sep 2018]
Mihai-Ioan Popescu [INSA Lyon]
Manon Predhumeau [Univ Grenoble Alpes, from Sep 2018]
Benoit Renault [INSA Lyon, from Sep 2018]
Luiz Serafim Guardini [Renault, from Oct 2018]
David Sierra Gonzalez [Inria]
Pavan Vasishta [Inria]

Post-Doctoral Fellows

Ozgur Erkent [Inria]
Mario Andrei Garzon Oviedo [Inria, from Feb 2018]
Andres Gomez Hernandez [Inria, from Nov 2018]

Mohamad Hobballah [INSA Lyon, from Feb 2018]
Alessandro Renzaglia [Inria]

Administrative Assistant

Sophie Karsai [Inria, until Nov 2018]

2. Overall Objectives

2.1. Origin of the project

Chroma is a bi-localized project-team at Inria Grenoble Rhone-Alpes in Grenoble and Lyon cities. The project was launched at the beginning of the year 2015 (March) before it became an Inria project-team on December 1st, 2017. It brings together experts in perception and decision-making for mobile robotics, all of them sharing common approaches that mainly relate to the field of Artificial Intelligence. It was originally founded by members of Inria project-team eMotion led by Christian Laugier (2002-2014) and teacher-researchers from INSA Lyon [0][0] working in the robotic group led by Prof. Olivier Simonin in CITI Lab. [0] (since 2013). Earlier members include Olivier Simonin (Prof. INSA Lyon), Christian Laugier (Inria researcher DR1), Jilles Dibangoye (Asso. Prof. INSA Lyon), Agostino Martinelli (Inria researcher CR1). On December 1st, 2015, Anne Spalanzani (Asso. Prof. Univ. Grenoble, habilite) joined the team (she was previously member of eMotion and Prima Inria project-teams).

2.2. Overall Objectives

The overall objective of Chroma is to address fundamental and open issues that lie at the intersection of the emerging research fields called “Human Centered Robotics” [0] and “Multi-Robot Systems (MRS) [0]".

More precisely, our goal is to design algorithms and develop models allowing mobile robots to navigate and cooperate in dynamic and human-populated environments. Chroma is involved in all decision aspects pertaining to single and multi robot navigation tasks, including perception and motion-planning.

The general objective is to build robotic behaviors that allow one or several robots to operate safely among humans in partially known environments, where time, dynamics and interactions play a significant role. Recent advances in embedded computational power, sensor and communication technologies, and miniaturized mechatronic systems, make the required technological breakthroughs possible (including from the scalability point of view).

Chroma is clearly positioned in the "Artificial Intelligence and Autonomous systems" research theme of the **Inria 2018-2022 Strategic Plan**. More specifically we refer to the "Augmented Intelligence" challenge (connected autonomous vehicles) and to the "Human centred digital world" challenge (interactive adaptation).

2.3. Research themes

To address the mentionned challenges, we take advantage of recent advances in all: probabilistic methods, planning techniques, multi-agent decision making, and machine learning. We also draw inspiration from other disciplines such as Sociology to take into account human models.

[0] National Institute of Applied Sciences
[0] INSA Lyon is part of the University de Lyon
[0] Centre of Innovation in Telecommunications and Integration of Service, see http://www.citi-lab.fr/
[0] Montreuil, V.; Clodic, A.; Ransan, M.; Alami, R., "Planning human centered robot activities," in Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on , vol., no., pp.2618-2623, 7-10 Oct. 2007
[0] IEEE RAS Multi-Robot Systems http://multirobotsystems.org/

Two main research themes of mobile robotics are addressed : i) Perception and Situation Awareness ii) Navigation and Cooperation in Dynamic Environments. Next, we elaborate more about these themes.

- **Perception and Situation Awareness.** This theme aims at understanding complex dynamic scenes, involving mobile objects and human beings, by exploiting prior knowledge and streams of perceptual data coming from various sensors. To this end, we investigate three complementary research problems:
  - **Bayesian Perception**: How to take into account prior knowledge and uncertain sensory data in a dynamic context?
  - **Situation awareness** : How to interpret the perceived scene and to predict their likely future motion (including near future collision risk) ?
  - **Robust state estimation**: acquire a deep understanding on several sensor fusion problems and investigate their observability properties in the case of unknown inputs.
- **Navigation and Cooperation in Dynamic Environments.** This theme aims at designing models and algorithms allowing robots to move and to coordinate efficiently in dynamic environments. We focus on two problems: navigation in human-populated environment (social navigation) and cooperation in large distributed fleet of robots (scalability and robustness issues).
  - **Motion-planning in human-populated environment.** How to plan trajectories that take into account the uncertainty of human-populated environments and respect the social rules of human beings? Such a challenge requires models of human behavior to be learnt or designed as well as dedicated learning or planning algorithms.
  - **Multi-robot decision making in complex environments.** How to design models and algorithms that can achieve both scalability and performance guarantees in real-world robotic systems? Our methodology builds upon advantages of two complementary approaches, Multi-Agent Sequential Decision Making (MA-SDM) and Swarm Intelligence (SI).

Chroma is also concerned with applications and transfer of the scientific results. Our main applications include autonomous and connected vehicles as well as service robotics. They are presented in Sections 4.2 and 4.3, respectively. Chroma is currently involved on several projects in collaboration with automobile companies (Renault, Toyota and Volvo) and some startups.

3. Research Program

3.1. Introduction

The Chroma team aims to deal with different issues of autonomous mobile robotics : perception, decision-making and cooperation. Figure 1 schemes the different themes and sub-themes investigated by Chroma.

We present here after our approaches to address these different themes of research, and how they combine altogether to contribute to the general problem of robot navigation. Chroma pays particular attention to the problem of autonomous navigation in highly dynamic environments populated by humans and cooperation in multi-robot systems. We share this goal with other major robotic laboratories/teams in the world, such as Autonomous Systems Lab at ETH Zurich, Robotic Embedded Systems Laboratory at USC, KIT [0] (Prof Christoph Stiller lab and Prof Ruediger Dillmann lab), UC Berkeley, Vislab Parma (Prof. Alberto Broggi), and iCeiRA [0] laboratory in Taipei, to cite a few. Chroma collaborates at various levels (visits, postdocs, research projects, common publications, etc.) with most of these laboratories, see Sections 9.3 and 9.4.

[0] Karlsruhe Institut fur Technologie

[0] International Center of Excellence in Intelligent Robotics and Automation Research.

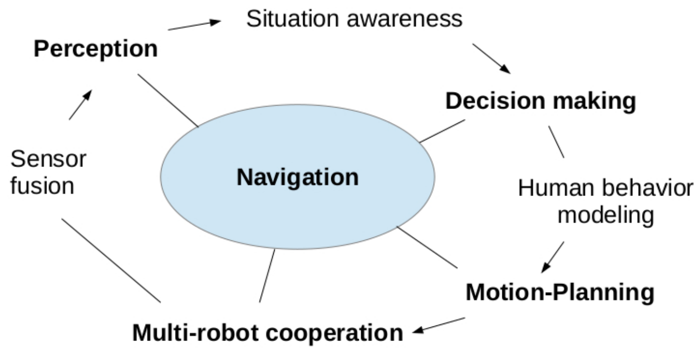


Figure 1. Research themes of the team and their relation

3.2. Perception and Situation Awareness

Participants: Christian Laugier, Agostino Martinelli, Jilles S. Dibangoye, Anne Spalanzani, Olivier Simonin, Christian Wolf, Ozgur Erkent, Alessandro Renzaglia, Rabbia Asghar, Jean-Alix David, Thomas Genevois, Jerome Lussereau, Anshul Paigwar, Lukas Rummelhard.

Robust perception in open and dynamic environments populated by human beings is an open and challenging scientific problem. Traditional perception techniques do not provide an adequate solution for these problems, mainly because such environments are uncontrolled [0] and exhibit strong constraints to be satisfied (in particular high dynamicity and uncertainty). This means that **the proposed solutions have to simultaneously take into account characteristics such as real time processing, temporary occultations, dynamic changes or motion predictions**.

3.2.1. Bayesian perception

Context. Perception is known to be one of the main bottlenecks for robot motion autonomy, in particular when navigating in open and dynamic environments is subject to strong real-time and uncertainty constraints. In order to overcome this difficulty, we have proposed in the scope of the former e-Motion team, a new paradigm in robotics called "Bayesian Perception". The foundation of this approach relies on the concept of "Bayesian Occupancy Filter (BOF)" initially proposed in the Ph.D. thesis of Christophe Coue [55] and further developed in the team [0]. The basic idea is to combine a Bayesian filter with a probabilistic grid representation of both the space and the motions. It allows the filtering and the fusion of heterogeneous and uncertain sensors data, by taking into account the history of the sensors measurements, a probabilistic model of the sensors and of the uncertainty, and a dynamic model of the observed objects motions.

In the scope of the Chroma team and of several academic and industrial projects (in particular the IRT Security for autonomous vehicle and Toyota projects), we went on with the development and the extension under strong embedded implementation constraints, of our Bayesian Perception concept. This work has already led to the development of more powerful models and more efficient implementations, e.g. the *CMCDOT* (Conditional Monte Carlo Dense Occupancy Tracker) framework [83] which is still under development.

This work is currently mainly performed in the scope of the "Security for Autonomous Vehicle (SAV)" project (IRT Nanoelec), and more recently in cooperation with some Industrial Companies (see section New Results for more details on the non confidential industrial cooperation projects).

[0] partially unknown and open

[0] The Bayesian programming formalism developed in e-Motion, pioneered (together with the contemporary work of Thrun, Burgard and Fox [91]) a systematic effort to formalize robotics problems under Probability theory–an approach that is now pervasive in Robotics.

Objectives. We aim at defining a complete framework extending the Bayesian Perception paradigm to the object level. The main objective is to be simultaneously more robust, more efficient for embedded implementations, and more informative for the subsequent scene interpretation step (Figure 2.a illustrates). Another objective is to improve the efficiency of the approach (by exploiting the highly parallel characteristic of our approach), while drastically reducing important factors such as the required memory size, the size of the hardware component, its price and the required energy consumption. This work is absolutely necessary for studying embedded solutions for the future generation of mobile robots and autonomous vehicles. We also aim at developing strong partnerships with non-academic partners in order to adapt and move the technology closer to the market.

Figure 2. a. Illustration of the Bayesian Perception Paradigm: Filtered occupancy grids, enhanced with motion estimations (vectors) and object detection (colored boxes) b. Autonomous Zoe car of Inria/Chroma.

3.2.2. System validation

Context. Testing and validating Cyber Physical Systems which are designed for operating in various real world conditions, is both an open scientific question and a necessity for a future deployment of such systems. In particular, this is the case for Embedded Perception and Decision-making Systems which are designed for future ADAS [0] and Autonomous Vehicles. Indeed, it is unrealistic to try to be exhaustive by making a huge number of experiments in various real situations. Moreover, such experiments might be dangerous, highly time consuming, and expensive. This is why we have decided to develop appropriate *realistic simulation and statistical analysis tools* for being able to perform a huge number of tests based on some previously recorded real data and on random changes of some selected parameters (the "co-simulation" concept). Such an approach might also be used in a training step of a machine learning process. This is why simulation-based validation is getting more and more popular in automotive industry and research.

This work is performed in the scope of both the SAV [0] project (IRT Nanoelec) and of the EU Enable-S3 project; it is also performed in cooperation with the Inria team Tamis in Rennes, with the objective to integrate the Tamis "Statistical Model Checking" (SMC) approach into our validation process. We are also starting to work on this topic with the Inria team Convecs, with the objective to also integrate formal methods into our validation process.

Objectives. We started to work on this new research topic in 2017. The first objective is to build a "simulated navigation framework" for: (1) constructing realistic testing environments (including the possibility of using real experiments records), (2) developing for each vehicle a simulation model including various physical and dynamic characteristics (e.g. physics, sensors and motion control), and (3) evaluating the performances of a simulation run using appropriate statistical software tools.

[0] Advance Driving Assistance System
[0] Security for Autonomous Vehicles

The second objective is to develop models and tools for automating the Simulation & Validation process, by using a selection of relevant randomized parameters for generating large database of tests and statistical results. Then, a metric based on the use of some carefully selected "Key Performance Indicator" (KPI) has to be defined for performing a statistical evaluation of the results (e.g. by using the above-mentioned SMC approach).

3.2.3. Situation Awareness and Prediction

Context. Predicting the evolution of the perceived moving agents in a dynamic and uncertain environment is mandatory for being able to safely navigate in such an environment. We have recently shown that an interesting property of the Bayesian Perception approach is to generate short-term conservative [0] predictions on the likely future evolution of the observed scene, even if the sensing information is temporary incomplete or not available [79]. But in human populated environments, estimating more abstract properties (e.g. object classes, affordances, agent's intentions) is also crucial to understand the future evolution of the scene. This work is carried out in the scope of the Security of Autonomous Vehicle (SAV) project (IRT Nanoelec) and of several cooperative and PhD projects with Toyota and with Renault.

Objectives. The first objective is to develop an integrated approach for "Situation Awareness & Risk Assessment" in complex dynamic scenes involving multiples moving agents (e.g vehicles, cyclists, pedestrians ...), whose behaviors are most of the time unknown but predictable. Our approach relies on combining machine learning to build a model of the agent behaviors and generic motion prediction techniques (e.g. Kalman-based, GHMM , or Gaussian Processes). In the perspective of a long-term prediction we will consider the semantic level [0] combined with planning techniques.

The second objective is to build a general framework for perception and decision-making in multi-robot/vehicle environments. The navigation will be performed under both dynamic and uncertainty constraints, with contextual information and a continuous analysis of the evolution of the probabilistic collision risk. Interesting published and patented results [67] have already been obtained in cooperation with Renault and UC Berkeley, by using the "Intention / Expectation" paradigm and Dynamic Bayesian Networks. We are currently working on the generalization of this approach, in order to take into account the dynamics of the vehicles and multiple traffic participants. The objective is to design a new framework, allowing us to overcome the shortcomings of rules-based reasoning approaches which often show good results in low complexity situations, but which lead to a lack of scalability and of long terms predictions capabilities.

3.2.4. Robust state estimation (Sensor fusion)

Context. In order to safely and autonomously navigate in an unknown environment, a mobile robot is required to estimate in real time several physical quantities (e.g., position, orientation, speed). These physical quantities are often included in a common state vector and their simultaneous estimation is usually achieved by fusing the information coming from several sensors (e.g., camera, laser range finder, inertial sensors). The problem of fusing the information coming from different sensors is known as the *Sensor Fusion* problem and it is a fundamental problem which plays a major role in robotics.

Objective. A fundamental issue to be investigated in any sensor fusion problem is to understand whether the state is observable or not. Roughly speaking, we need to understand if the information contained in the measurements provided by all the sensors allows us to carry out the estimation of the state. If the state is not observable, we need to detect a new observable state. This is a fundamental step in order to properly define the state to be estimated. To achieve this goal, we apply standard analytic tools developed in control theory together with some new theoretical concepts we introduced in [71] (concept of continuous symmetry). Additionally, we want to account the presence of disturbances in the observability analysis.

Our approach is to introduce general analytic tools able to derive the observability properties in the nonlinear case when some of the system inputs are unknown (and act as disturbances). We recently obtained a simple analytic tool able to account the presence of unknown inputs [73], which extends a heuristic solution derived by the team of Prof. Antonio Bicchi [51] with whom we collaborate (Centro Piaggio at the University of Pisa).

[0] i.e. when motion parameters are supposed to be stable during a small amount of time
[0] knowledge about agent's activities and tasks

Fusing visual and inertial data. A special attention is devoted to the fusion of inertial and monocular vision sensors (which have strong application for instance in UAV navigation). The problem of fusing visual and inertial data has been extensively investigated in the past. However, most of the proposed methods require a state initialization. Because of the system nonlinearities, lack of precise initialization can irreparably damage the entire estimation process. In literature, this initialization is often guessed or assumed to be known [49], [69], [61]. Recently, this sensor fusion problem has been successfully addressed by enforcing observability constraints [63], [64] and by using optimization-based approaches [68], [60], [70], [65], [78]. These optimization methods outperform filter-based algorithms in terms of accuracy due to their capability of relinearizing past states. On the other hand, the optimization process can be affected by the presence of local minima. We are therefore interested in a deterministic solution that analytically expresses the state in terms of the measurements provided by the sensors during a short time-interval.

For some years we explore deterministic solutions as presented in [72] and [74]. Our objective is to improve the approach by taking into account the biases that affect low-cost inertial sensors (both gyroscopes and accelerometers) and to exploit the power of this solution for real applications. This work is currently supported by the ANR project VIMAD [0] and experimented with a quadrotor UAV. We have a collaboration with Prof. Stergios Roumeliotis (the leader of the MARS lab at the University of Minnesota) and with Prof. Anastasios Mourikis from the University of California Riverside. Regarding the usage of our solution for real applications we have a collaboration with Prof. Davide Scaramuzza (the leader of the Robotics and Perception group at the University of Zurich) and with Prof. Roland Siegwart from the ETHZ.

3.3. Navigation and cooperation in dynamic environments

Participants: Olivier Simonin, Anne Spalanzani, Jilles S. Dibangoye, Christian Wolf, Laetitia Matignon, Fabrice Jumel, Jacques Saraydaryan, Christian Laugier, Alessandro Renzaglia, Mohamad Hobballah, Vincent Le Doze.

In his reference book *Planning algorithms*[0] S. LaValle discusses the different dimensions that made the motion-planning problem complex, which are the number of robots, the obstacle region, the uncertainty of perception and action, and the allowable velocities. In particular, it is emphasized that complete algorithms require at least exponential time to deal with multiple robot planning in complex environments, preventing them to be scalable in practice. Moreover, dynamic and uncertain environments, as human-populated ones, expand this complexity.

In this context, we aim at **scale up decision-making in human-populated environments and in multi-robot systems, while dealing with the intrinsic limits of the robots (computation capacity, limited communication)**.

3.3.1. Motion-planning in human-populated environment

Context. Motion planning in dynamic and human-populated environments is a current challenge of robotics. Many research teams work on this topic. We can cite the Institut of robotic in Barcelone [59], the MIT [48], the Autonomous Intelligent Systems lab in Freiburg [52], or the LAAS [80]. In Chroma, we explore different issues : **integrating the risk (uncertainty) in planning processes, modeling and taking into account human behaviors and flows.**

Objective We aim to give the robot some socially compliant behaviors by anticipating the near future (trajectories of mobile obstacle in the robot's surroundings) and by integrating knowledge from psychology, sociology and urban planning. In this context, we will focus on the following 3 topics.

[0] Navigation autonome des drones aériens avec la fusion des données visuelles et inertielles, lead by A. Martinelli, Chroma.
[0] Steven M. LaValle, Planning Algorithms, Cambridge University Press, 2006.

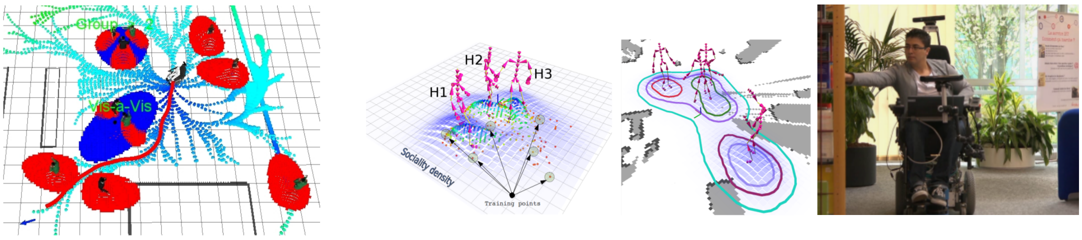


Figure 3. Illustrations of a. the Risk-RRT planning b. the human interaction space model c. experiment with the wheelchair.

Risk-based planning. Unlike static or controlled environments [0] where global path planning approaches are suitable, dealing with highly dynamic and uncertain environments requires to integrate the notion of risk (risk of collision, risk of disturbance). Then, we examine how motion planning approaches can integrate this risk in the generation and selection of the paths. An algorithm called RiskRRT was proposed in the previous eMotion team. This algorithm plans goal oriented trajectories that minimize the risk estimated at each instant. It fits environments that are highly dynamic and adapts to a representation of uncertainty [90] (see Figure 3.a for illustration). Now, we extend this principle to be adapted to various risk evaluation methods (proposed in 3.2) and various situation (highways, urban environments, even in dense traffic).

We also investigate the problem of learning recurring human displacements - or flows of humans - from robots embedded sensors. It has been shown that such recurring behaviors can be mapped from spatial-temporal observations, as in [92]. In this context, we explore counting-based mapping models [66] to learn motion probabilities in cells of a grid representing the environment. Then we can revisit cost-function of path-planning algorithms (eg. A*) by integrating the risk to encountering humans in opposite direction. We also aim at demonstrating the efficiency of the approach with real robots evolving in dense human-populated environments.

Recently we investigated the automatic learning of robot navigation in complex environments based on specific tasks and from visual input. We address this problem by combining computer vision, machine learning (deep-learning), and robotics path planning (see 7.4.2).

Sharing the physical space with humans. Robots are expected to share their physical space with humans. Hence, robots need to take into account the presence of humans and to behave in a socially acceptable way. Their trajectories must be safe but also predictable, that is why they must follow social conventions, respecting proximity constraints, avoiding people interacting or joining a group engaged in conversation without disturbing. For this purpose, we proposed earlier to integrate some knowledge from the psychology domain (i.e. proxemics theory), see figure 3.b. We aim now to integrate semantic knowledge [0] and psycho-social theories of human behavior [00] in the navigation framework we have developed for a few years (i.e. the Risk-based navigation algorithms [62], [90], [96]). These concepts were tested on our automated wheelchair (see figure 3.c) but they have and will be adapted to autonomous cars, telepresence robots and companion robots. This work is currently supported by the ANR Valet and the ANR Hianic.

[0] known environment without uncertainty

[0] B. Kuipers, The Spatial Semantic Hierarchy, Artificial Intelligence, Volume 119, Issues 1–2, May 2000, Pages 191-233

[0] Gibson, J. (1977). The theory of affordances, in Perceiving, Acting, and Knowing. Towards an Ecological Psychology. Number eds Shaw R., Bransford J. Hoboken,NJ: John Wiley & Sons Inc.

[0] Hall, E. (1966). The hidden dimension. Doubleday Anchor Books.

3.3.2. Decision Making in Multi-robot systems

Context. A central challenge in Chroma is to define **decision-making algorithms that scale up to large multi-robot systems**. This work takes place in the general framework of Multi-Agent Systems (MAS). The objective is to compute/define agent behaviors that provide cooperation and adaptation abilities. Solutions must also take into account the agent/robot computational limits.

We can abstract the challenge in three objectives :
i) mastering the complexity of large fleet of robots/vehicles (scalability),
ii) dealing with limited computational/memory capacity,
iii) building adaptive solutions (robustness).

Combining Decision-theoretic models and Swarm intelligence.

Over the past few years, our attempts to address multi-robot decision-making are mainly due to Multi-Agent Sequential Decision Making (MA-SDM) and Swarm Intelligence (SI). MA-SDM builds upon well-known decision-theoretic models (e.g., Markov decision processes and games) and related algorithms, that come with strong theoretical guarantees. In contrast, the expressiveness of MA-SDM models has limited scalability in face of realistic multi-robot systems [0], resulting in computational overload. On their side, SI methods, which rely on local rules – generally bio-inspired – and relating to Self-Organized Systems [0], can scale up to multiple robots and provide robustness to disturbances, but with poor theoretical guarantees [0]. Swarm models can also answer to the need of designing tractable solutions [89], but they remain not geared to express complex realistic tasks or to handle (point-to-point) communication between robots. This motivates our work to go beyond these two approaches and to combine them.

First, we plan to investigate **incremental expansion mechanisms in anytime decision-theoretic planning**, starting from local rules (from SI) to complex strategies with performance guarantees (from MA-SDM) [57]. This methodology is grounded into our research on anytime algorithms, that are guaranteed to stop at anytime while still providing a reliable solution to the original problem. It further relies on decision theoretical models and tools including: Decentralized and Partially Observable Markov Decision Processes and Games, Dynamic Programming, Distributed Reinforcement Learning and Statistical Machine Learning.

Second, we plan to extend the SI approach by considering **the integration of optimization techniques at the local level**. The purpose is to force the system to explore solutions around the current stabilized state – potentially a local optimum – of the system. We aim at keeping scalability and self-organization properties by not compromising the decentralized nature of such systems. Introducing optimization in this way requires to measure locally the performances, which is generally possible from local perception of robots (or using learning techniques). The main optimization methods we will consider are Local Search (Gradient Descent), Distributed Stochastic Algorithm and Reinforcement Learning. We have shown in [93] the interest of such an approach for driverless vehicle traffic optimization.

[0]Martin L. Puterman, Markov Decision Processes; Stuart Russell and Peter Norvig, Artificial Intelligence - A Modern Approach

[0]D. Floreano and C. Mattiussi, Bio-Inspired Artificial Intelligence - Theories, Methods, and Technologies, MIT Press, 2008.

[0]S. A. Brueckner, G. Di Marzo Serugendo, A. Karageorgos, R. Nagpal (2005). Engineering Self-Organising Systems, Methodologies and Applications. LNAI 3464 State-of-the-Art Survey, Springer book.

Both approaches must lead to **master the complexity** inherent to large and open multi-robot systems. Such systems are prone to combinatorial problems, in term of state space and communication, when the number of robots grows. To cope with this complexity we explore several approaches :

- Combining MA-SDM, machine learning and RO [0] techniques to deal with global-local optimization in multi-agent/robot systems. In 2016, we started a collaboration with the VOLVO Group, in Lyon, to deal with VRP problems and optimization of goods distribution using a fleet of autonomous vehicles. We also explore such a methodology in the framework of the collaboration with the team of Prof. G. Czibula (Cluj Univiersity, Romania).
- Defining heuristics by decentralizing global exact solutions. We explore this methodology to deal with dynamic problems such as the patrolling of moving persons (see [86]). We also deal with dynamic-MRR (Multi-Robot Routing) problems in the context of the PhD of M. Popescu, see Section 7.5.1.2.
- Online incremental refining of the environment representation. This allows us to revisit mapping/coverage techniques and problems, see section 7.5.2.1 [77].

Beyond this methodological work, we aim to evaluate our models on benchmarks from the literature, by using simulation tools as a complement of robotic experiments. This will lead us to develop simulators, allowing to deploy thousands of humans and robots in constrained environments.

Towards adaptive connected robots.

Mobile robots and autonomous vehicles are becoming more connected to one another and to other devices in the environment (concept of cloud of robots [0] and V2V/V2I connectivity in transportation systems). Such robotic systems are open systems as the number of connected entities is varying dynamically. Network of robots brought with them new problems, as the need of (online) adaption to changes in the system and to the variability of the communication.

In Chroma, we address the problem of adaptation by considering machine learning techniques and local mechanisms as discussed above (SI models). More specifically we investigate the problem of maintaining the connectivity between robots which perform dynamic version of tasks such as patrolling, exploration or transportation, i.e. where the setting of the problem is continuously changing and growing (see [81]).

In Lyon, the CITI Laboratory conducts research in many aspects of telecommunication, from signal theory to distributed computation. In this context, Chroma develops cooperations with the Inria team Agora [81] (wireless communication protocols) and with Dynamid team [54] (middlleware and cloud aspects), that we wish to reinforce in the next years.

4. Application Domains

4.1. Introduction

Applications in Chroma are organized in two main domains : **i) Future cars and transportation systems and ii) Services robotics.** These domains correspond to the experimental fields initiated in Grenoble (eMotion team) and in Lyon (CITI lab). However, the scientific objectives described in the previous sections are intended to apply equally to both applicative domains. Even our work on Bayesian Perception is today applied to the intelligent vehicle domain, we aim to generalize to any mobile robots. The same remark applies to the work on multi-agent decision making. We aim to apply algorithms to any fleet of mobile robots (service robots, connected vehicles, UAVs). This is the philosophy of the team since its creation.

[0] Operations Research

[0] see for instance the first International Workshop on Cloud and Robotics, 2016.

Figure 4. Most of the Chroma platforms: the Pepper robot, a fleet of (22) Turtlebot 2, one of the 4 Bebop drones and the equipped Toyota Lexus.

4.2. Future cars and transportation systems

Thanks to the introduction of new sensor and ICT technologies in cars and in mass transportation systems, and also to the pressure of economical and security requirements of our modern society, this application domain is quickly changing. Various technologies are currently developed by both research and industrial laboratories. These technologies are progressively arriving at maturity, as it is witnessed by the results of large scale experiments and challenges such as the Google's car project and several future products announcements made by the car industry. Moreover, the legal issue starts to be addressed in USA (see for instance the recent laws in Nevada and in California authorizing autonomous vehicles on roads) and in several other countries (including France).

In this context, we are interested in the development of ADAS [0] systems aimed at improving comfort and safety of the cars users (e.g., ACC, emergency braking, danger warnings), and of Fully Autonomous Driving functions for controlling the displacements of private or public vehicles in some particular driving situations and/or in some equipped areas (e.g., automated car parks or captive fleets in downtown centers or private sites).

Since about 8 years, we are collaborating with Toyota and with Renault-Nissan on these applications (bilateral contracts, PhD Theses, shared patents), but also recently with Volvo group (PhD thesis started in 2016). We are also strongly involved (since 2012) in the innovation project Perfect then now Security for autonomous vehicle of the IRT [0] Nanoelec (transportation domain). In 2016, we have been awarded a European H2020 ECSEL project [0] involving major European automotive constructors and car suppliers. In this project, Chroma is focusing on the embedded perception component (models and algorithms, including the certification issue), in collaboration with Renault, Valeo and also with the Inria team TAMIS (Rennes). Chroma is also involved in the ANR project "Valet" (2015-2018) coordinated by the Inria team RITS (Rocquencourt), dealing with automatic redistribution of car-sharing vehicles and parking valet; Chroma is involved in the pedestrian-vehicle interaction for a safe navigation.

In this context, Chroma has two experimental vehicles equipped with various sensors (a Toyota Lexus and a Renault Zoe, see. Fig. 4 and Fig. 2.b), which are maintained by Inria-SED [0] and that allow the team to perform experiments in realistic traffic conditions (Urban, road and highway environments). The Zoe car has been automated in December 2016, through our collaboration with the team of P. Martinet (IRCCyN Lab, Nantes), that allow new experiments in the team.

4.3. Services robotics

[0] Advanced Driver Assistance Systems
[0] Institut de Recherche Technologique
[0] ENABLE-S3: European Initiative to Enable Validation for Highly Automated Safe and Secure Systems.
[0] Service Expérimentation et Développement

Service robotics is an application domain quickly emerging, and more and more industrial companies (e.g., IS-Robotics, Samsung, LG) are now commercializing service and intervention robotics products such as vacuum cleaner robots, drones for civil or military applications, entertainment robots ... One of the main challenges is to propose robots which are sufficiently robust and autonomous, easily usable by non-specialists, and marked at a reasonable cost. We are involved in developing observation and surveillance systems, by using ground robots and aerial ones, see Fig. 4. Since 2016, we develop solutions for 3D observation/exploration of complex scenes or environments with a fleet of UAVs (Inria ADT CORDES [0]) or mobile robots (COMODYS FIL project [32]).

A more recent challenge for the coming decade is to develop robotized systems for assisting elderly and/or disabled people. In the continuity of our work in the IPL PAL [0], we aim to propose smart technologies to assist electric wheelchair users in their displacements and also to conttrol autonomous cars in human crowds. This concerns our recent "Hianic" ANR project. Another emerging application is humanoid robots helping humans at their home or work. In this context, we address the problem of NAMO (Navigation Among Movable Obstacles) in human populated environments (eg. PhD of B. Renault started on 2018). More generally we address navigation and reconnaissance tasks with Pepper humanoids in the context of the RoboCup-Social League.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Best student paper, 15th International Conference on Control, Automation, Robotics and Vision, Nov 2018, Singapore, Singapore (ICARCV 2018), Pavan Vasishta, Dominique Vaufreydaz, Anne Spalanzani

5.1.2. Other highlights of 2018

- Success for several project applications in the field of Autonomous Vehicles : ANR"Hianic", PIA Ademe "CAMPUS", FUI "STAR" and "TORNADO".
- In 2018, Chroma published several papers in Artificial Intelligence A+ ranked conferences: CVPR [21], NIPS [27], ICML [26], AAMAS [32].
- Strong involvement of Chroma in the IEEE/RSJ IROS 2018 Conference (Madrid, October 2018, more than 4000 people): C. Laugier was Program co-chair and co-organized three interconnected events on Autonomous Vehicles: a one day Workshop having attracted more than 360 people [0], an Industrial Forum involving international companies (e.g. Renault, Nvidia, Baidu, EasyMile, Ambarella, etc) and having attracted about 80 people, and an Autonomous Vehicles Demonstration involving 5 international teams (including Chroma with our Autonomous Renault Zoe car) [0] [46].
- First participation to the international RoboCup competition (Montréal, Juin, 2018) : we created the 'LyonTech' team to compete in the robocup@Home Pepper league. We ranked 5th over 21 participants. LyonTech is composed of members from Chroma (F. Jumel, L. Matignon, J. Saraydaryan, O. Simonin, C. Wolf) and two engineers from CPE Lyon (R. Leber) and LIRIS lab/CNRS (E. Lombardi). In October 2018, we qualified for the next RoboCup final, to be organized in Sydney, on July 2019.
- Participation in several International Award Committees (C. Laugier): Several IEEE/RSJ IROS 2018 Award Committees (Best Paper Award, Fellow Award, Harashima Award, Distinguished Service Award, Young Professional Award), IEEE ICARCV 2018 Best Paper Award Committee, IEEE Chapter Award Committee 2018.
- French Robotics GDR : co-animation of the new GT « Apprentissage et Robotique » by Christian Wolf (with David Filiat), started in November 2018 ; O. Simonin will chair, with F. Charpillet (Inria Larsen), the next National Conference on Robotics Reseacrh (JNRR), on October 2019.

[0] Coordination d'une Flotte de Drones Connectés pour la Cartographie 3D d'édifices, led by O. Simonin.
[0] Personnaly assisted Living
[0] See website: http://project.inria.fr/ppniv18
[0] see website: http://avdemo.car.upm-csic.es/

BEST PAPERS AWARDS :
[36]
P. VASISHTA, D. VAUFREYDAZ, A. SPALANZANI.*Building Prior Knowledge: A Markov Based Pedestrian Prediction Model Using Urban Environmental Data*, in "ICARCV 2018 - 15th International Conference on Control, Automation, Robotics and Vision", Singapore, Singapore, November 2018, p. 1-12, https://arxiv.org/abs/1809.06045 , https://hal.inria.fr/hal-01875147

6. New Software and Platforms

6.1. kinetics

FUNCTIONAL DESCRIPTION: Software computing decision support strategies and decision-making

- Contact: Jilles Dibangoye

6.2. VI-SFM

FUNCTIONAL DESCRIPTION: Experimentary the closed Form Solution for usual-initial data fusion agains real and simulated fusion

- Authors: Agostino Martinelli and Jacques Kaiser
- Contact: Agostino Martinelli

6.3. Ground Elevation and Occupancy Grid Estimator (GEOG - Estimator)

KEYWORDS: Robotics - Environment perception
FUNCTIONAL DESCRIPTION: GEOG-Estimator is a system of joint estimation of the shape of the ground, in the form of a Bayesian network of constrained elevation nodes, and the ground-obstacle classification of a pointcloud. Starting from an unclassified 3D pointcloud, it consists of a set of expectation-maximization methods computed in parallel on the network of elevation nodes, integrating the constraints of spatial continuity as well as the influence of 3D points, classified as ground-based or obstacles. Once the ground model is generated, the system can then construct a occupation grid, taking into account the classification of 3D points, and the actual height of these impacts. Mainly used with lidars (Velodyne64, Quanergy M8, IBEO Lux), the approach can be generalized to any type of sensor providing 3D pointclouds. On the other hand, in the case of lidars, free space information between the source and the 3D point can be integrated into the construction of the grid, as well as the height at which the laser passes through the area (taking into account the height of the laser in the sensor model). The areas of application of the system spread across all areas of mobile robotics, it is particularly suitable for unknown environments. GEOG-Estimator was originally developed to allow optimal integration of 3D sensors in systems using 2D occupancy grids, taking into account the orientation of sensors, and indefinite forms of grounds. The ground model generated can be used directly, whether for mapping or as a pre-calculation step for methods of obstacle recognition or classification. Designed to be effective (real-time) in the context of embedded applications, the entire system is implemented on Nvidia graphics card (in Cuda), and optimized for Tegra X2 embedded boards. To ease interconnections with the sensor outputs and other perception modules, the system is implemented using ROS (Robot Operating System), a set of opensource tools for robotics.

- Authors: Amaury Nègre, Lukas Rummelhard, Lukas Rummelhard, Jean-Alix David and Christian Laugier
- Contact: Christian Laugier

6.4. CMCDOT

KEYWORDS: Robotics - Environment perception

FUNCTIONAL DESCRIPTION: CMCDOT is a Bayesian filtering system for dynamic occupation grids, allowing parallel estimation of occupation probabilities for each cell of a grid, inference of velocities, prediction of the risk of collision and association of cells belonging to the same dynamic object. Last generation of a suite of Bayesian filtering methods developed in the Inria eMotion team, then in the Inria Chroma team (BOF, HSBOF, ...), it integrates the managment of hybrid sampling methods (classical occupancy grids for static parts, particle sets for parts dynamics) into a Bayesian unified programming formalism , while incorporating elements resembling the Dempster-Shafer theory (state "unknown", allowing a focus of computing resources). It also offers a projection system of the estimated scene in the near future, to reference potential collisions with the ego-vehicle or any other element of the environment, as well as very low cost pre-segmentation of coherent dynamic spaces (taking into account speeds). It takes as input instantaneous occupation grids generated by sensor models for different sources, the system is composed of a ROS package, to manage the connectivity of I / O, which encapsulates the core of the embedded and optimized application on GPU Nvidia (Cuda), allowing real-time analysis of the direct environment on embedded boards (Tegra X1, X2). ROS (Robot Operating System) is a set of open source tools to develop software for robotics. Developed in an automotive setting, these techniques can be exploited in all areas of mobile robotics, and are particularly suited to highly dynamic and uncertain environment management (eg urban scenario, with pedestrians, cyclists, cars, buses, etc.).

- Authors: Amaury Nègre, Lukas Rummelhard, Jean-Alix David and Christian Laugier
- Partners: CEA - CNRS
- Contact: Christian Laugier

6.5. cuda_grid_fusion

KEYWORDS: Robotics - Environment perception
FUNCTIONAL DESCRIPTION: This module, directly implemented in ROS / Cuda, performs the merge of occupancy grids, defined in the format proposed in CMCDOT (probabilities integrating the "visibility" information of the cell, via the coefficients "unknown") thanks to an original method, allowing not only consistency with the rest of the system, but also a nuanced consideration of confidence criteria towards the various sources of information.

- Authors: Lukas Rummelhard and Jean-Alix David
- Contact: Lukas Rummelhard

6.6. cuda_laser_grid

KEYWORDS: Robotics - Environment perception
FUNCTIONAL DESCRIPTION: This module generates occupation grids from "almost" planar lidar. The sensor model, as well as the outputs, have been modified, in order to be fully consistent with the CMCDOT and grid fusion module formats.

- Authors: Amaury Nègre, Lukas Rummelhard and Jean-Alix David
- Contact: Lukas Rummelhard

6.7. CMCDOT-Tools

KEYWORD: Robotics
FUNCTIONAL DESCRIPTION: Tools for CMCDOT Software

- Authors: Amaury Nègre, Lukas Rummelhard, Jean-Alix David, Mathias Perrollaz, Procopio Silveira-Stein, Jérôme Lussereau and Nicolas Vignard
- Contact: Jean-Alix David

6.8. DWA Planner on occupancy grid

Dynamic Window Approach Planner based on occupancy grid

KEYWORD: Navigation
FUNCTIONAL DESCRIPTION: This program considers : - a given target - an occupancy grid which represents the environment - the odometry of the vehicle With these data, it computes the commands for a safe navigation towards the target.

- Authors: Christian Laugier and Thomas Genevois
- Partner: CEA
- Contact: Christian Laugier

6.9. Zoe Simulation

Simulation of Inria's Renault Zoe in Gazebo environment
KEYWORD: Simulation
FUNCTIONAL DESCRIPTION: This simulation represents the Renault Zoe vehicle considering the realistic physical phenomena (friction, sliding, inertia, ...). The simulated vehicle embeds sensors similar to the ones of the actual vehicle. They provide measurement data under the same format. Moreover the software input/output are identical to the vehicle's. Therefoe any program executed on the vehicle can be used with the simulation and reciprocally.

- Authors: Christian Laugier, Nicolas Turro and Thomas Genevois
- Contact: Christian Laugier

6.10. PedSim-ROS

FUNCTIONAL DESCRIPTION: Simulation of moving people and mobile robots that can detect agents around them. Integration of ROS mobile robots with the PedSim simulator.

- Contact: Jacques Saraydaryan

6.11. EKF Odom

EKF based localisation for vehicles
KEYWORDS: Localization - Autonomous Cars
FUNCTIONAL DESCRIPTION: This software fuses IMU data with wheel rotation or speed measurement inside an Extended Kalman Filter. It estimates the state position, orientation, speed, angular speed, acceleration.

- Authors: Thomas Genevois and Christian Laugier
- Contact: Christian Laugier
- URL: https://team.inria.fr/chroma/en/

6.12. Light Vehicle Simulation

Simulation of a light vehicle in Gazebo environment
KEYWORD: Simulation
FUNCTIONAL DESCRIPTION: This simulation represents a light vehicle considering the realistic physical phenomena (friction, sliding, inertia, ...). The simulated vehicle embeds sensors similar to the ones of the actual vehicle. They provide measurement data under the same format. Moreover the software input/output are identical to the vehicle's. Therefore any program executed on the vehicle can be used with the simulation and reciprocally.

- Authors: Thomas Genevois and Christian Laugier
- Contact: Christian Laugier
- URL: https://team.inria.fr/chroma/en/

6.13. CarHybridSim

Hybrid simulation for autonomous cars with high traffic
KEYWORDS: Simulation - Autonomous Cars
FUNCTIONAL DESCRIPTION: Open source tool for simulating autonomous vehicles in complex, high traffic, scenarios. The hybrid simulation fully integrates and synchronizes a microscopic, multi-modal traffic simulator and a complex 3D simulator.

- Contact: Mario Garzon Oviedo
- URL: https://github.com/marioney/hybrid_simulation

6.14. SimuDronesGR

Simultion of UAV fleets with Gazebo/ROS
KEYWORDS: Robotics - Simulation
FUNCTIONAL DESCRIPTION: The simulator includes the following functionality : 1) Simulation of the mechanical behavior of an Unmanned Aerial Vehicle : * Modeling of the body's aerodynamics with lift, drag and moment * Modeling of rotors' aerodynamics using the forces and moments' expressions from Philppe Martin's and Erwan Salaün's 2010 IEEE Conference on Robotics and Automation paper "The True Role of Accelerometer Feedback in Quadrotor Control". 2) Gives groundtruth informations : * Positions in East-North-Up reference frame * Linear velocity in East-North-Up and Front-Left-Up reference frames * Linear acceleration in East-North-Up and Front-Left-Up reference frames * Orientation from East-North-Up reference frame to Front-Left-Up reference frame (Quaternions) * Angular velocity of Front-Left-Up reference frame expressed in Front-Left-Up reference frame. 3) Simulation of the following sensors : * Inertial Measurement Unit with 9DoF (Accelerometer + Gyroscope + Orientation) * Barometer using an ISA model for the troposphere (valid up to 11km above Mean Sea Level) * Magnetometer with the earth magnetic field declination * GPS Antenna with a geodesic map projection.

RELEASE FUNCTIONAL DESCRIPTION: Initial version

- Author: Vincent Le Doze
- Partner: Insa de Lyon
- Contact: Vincent Le Doze

7. New Results

7.1. Bayesian Perception

Participants: Christian Laugier, Lukas Rummelhard, Jean-Alix David, Thomas Genevois, Jerome Lussereau, Nicolas Turro [SED], Jean-François Cuniberto [SED].

7.1.1. Conditional Monte Carlo Dense Occupancy Tracker (CMCDOT) Framework

Participants: Lukas Rummelhard, Jerome Lussereau, Jean-Alix David, Thomas Genevois, Christian Laugier, Nicolas Turro [SED].

Recognized as one of the core technologies developed within the team over the years (see related sections in previous activity report of Chroma, and previously e-Motion reports), the CMCDOT framework is a generic Bayesian Perception framework, designed to estimate a dense representation of dynamic environments [83] and the associated risks of collision [85], by fusing and filtering multi-sensor data. This whole perception system has been developed, implemented and tested on embedded devices, incorporating over time new key modules [84]. In 2018, this framework, and the corresponding software, has continued to be the core of many important industrial partnerships and academic contributions [17] [18] [16] [15] [45] [47], and to be the subject of important developments, both in terms of research and engineering. Some of those recent evolutions are detailed below.

- CMCDOT evolutions : important developments in the CMCDOT, in terms of calculation methods and fundamental equations, were introduced and tested this year. These developments could lead, in the coming months, to the proposal of a new patent, then to academic publications. These changes introduced, among other evolutions, a much higher update frequency, greater flexibility in the management of transitions between states (and therefore a better system reactivity), as well as the management of a high variability in sensor frequencies (for each sensor over time, and in the set of sensors). The technical documents describing those developments are currently being redacted, and will be described in the next annual report.
- Multi-sensor integration in the Ground Estimator : the module of dynamic estimation of the shape of the ground and data segmentation, based solely on the sensor point clouds (no prior map information), the first step of data interpretation in CMCDOT framework, has been developed since 2016, patented and published in 2017. The corresponding software, until this year, could not take into account more than one sensor. In case of multiple sensors, several different modules were to be launched, their respective occupancy grids then fused, not only increasing the global computation use, but also preventing each sensor from benefiting from the ground models generated by the others. This point was corrected this year, by introducing the management of multiple input sensors, unifying the ground estimation in a single model, thus leading to improved performance, both in terms of calculation and results.
- Velocity display : in the CMCDOT framework, velocity of every element of the scene is inferred at a cell level, without object segmentation. This low-level velocity estimation is one of the most original and important aspects of the method, and should be displayed accordingly. A velocity display module, displaying for each occupied cell of the grid the average of the estimated velocity, generating colors depending on the intensity and the orientation, has been developed, see Fig. 5.

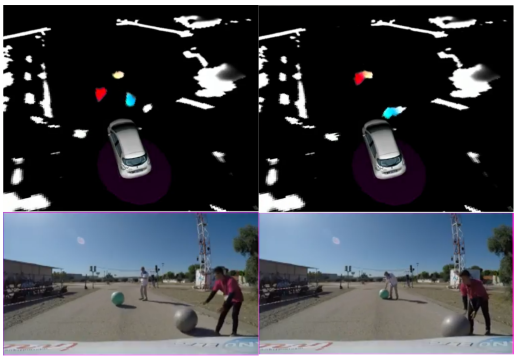


Figure 5. Image from the Velocity Display module : in every occupied cell of the grid, the average velocity is represented by a color code, the hue being based on the orientation, the saturation on its norm. A static cell is white, a cell moving in the same direction as the vehicle is red, in the opposite direction in blue. In the grid can be seen the moving balloons, the pedestrians being static.

- Software optimization : the whole CMCDOT framework has been developed on GPUs (implementations in C++/Cuda), an important focus of the engineering has always been, and continued to be in 2018, on the optimization of the software and methods to be embedded on low energy consumption embedded boards (now Nvidia Jetson TX2).
- IROS 2018 Autonomous Driving event : https://hal.inria.fr/medihal-01963296v1 As already mentioned in the highlights of the year, the experimental Zoe platform, funded by IRT Nanoelec, has participated at IROS2018 in the Autonomous Vehicle Demonstrations, a full day of demonstration of autonomous vehicle capacities from various research centers. During this successful event, it has

been presented and demonstrated on live conditions the effectiveness of the embedded CMCDOT framework, in connection with the newly developed control and decision making systems.

7.1.2. Simulation based validation

Participants: Thomas Genevois, Lukas Rummelhard, Nicolas Turro [SED], Christian Laugier, Anshul Paigwar, Alessandro Renzaglia.

Since 2017, we are working to address the concept of *simulation based validation* in the scope of the EU Enable-S3 project, with the objective of searching for novel approaches, methods, tools and experimental methodology for validating BOF-based algorithms. For that purpose, we have collaborated with the Inria Tamis team (Rennes) and with Renault for developing the simulation platform that is used in the test platform. The simulation of both the sensors and the driving environment are based on the Gazebo simulator. A simulation of the prototype car and its sensors has also been realized, meaning that the same implementation of *CMCDOT* can handle both real data and simulated data. The test management component that generates random simulated scenarios has also been developed. Output of *CMCDOT* computed from the simulated scenarios are recorded by *ROS* and analyzed through the Statistical Model Checker *(SMC)* developed by the Inria Tamis team. In [41], we presented the first results of this work, where a decision-making approach for intersection crossing (see Section 7.2.3) has been analyzed. In particular new KPIs expressed as Bounded Linear Temporal Logic (BLTL) formula have been defined. Temporal formulas allow a finer formulation of KPIs by taking into account the evolution of the metrics during time. A further work in this direction will be done in the next months to provide new results on the validation of the perception algorithm, namely for the velocity estimation and collision risk assessment. For this part, we are also exploring the advantages and potentiality of a new open-source vehicle simulator (Carla), which would allow considering more realistic scenarios with respect to Gazebo. This work on simulation-based validation will be continued in 2019.

Previously, in 2017, CHROMA has developed a model of the Renault Zoe demonstrator within the simulation framework Gazebo. In 2018, we have improved it to keep it up-to-date after several evolutions of the actual demonstrator. Namely, the drivers of the simulated lidars and the control law have been updated. Thus the model now provides the outputs corresponding to a simulated Inertial Measurement Unit.

7.1.3. Control and navigation

Participants: Thomas Genevois, Lukas Rummelhard, Jerome Lussereau, Jean-Alix David, Christian Laugier, Nicolas Turro [SED], Rabbia Asghar.

In 2018, we have updated the Renault Zoe demonstrator in collaboration with the LS2N (Laboratoire des Sciences Numérique de Nantes). The control codes have been transferred to the micro-controllers of the car for a faster and more precise control. An electric signal has been added to identify when the driver acts on the manual controls of the car. Finally the control law of the vehicle has been modified in order to consider a command in acceleration. These modifications allowed us to improve the software we use to control the vehicle. We have improved our implementation of DWA (Dynamic Window Approach) local planner in order to handle acceleration commands. This local planner has also been modified to take in account maxima of lateral acceleration and to integrate a path following module in its cost function. Thanks to this, the new version of this program provides a smooth command for a combination of path following and obstacle avoidance with the demonstrator Renault Zoe. This has been showed at the Autonomous Vehicle Demonstration event at IROS2018, Madrid, Figure 6 [46].

We have also experimented a driving assistant for autonomous obstacle avoidance. We showed that it is possible on the Renault Zoe demonstrator to let a driver drive manually the car and then, when a collision risk is identified, to take over the control with the autonomous drive and perform an avoidance maneuver. A simple ADAS [0] system has been developed for this purpose. In addition, we have developed on the Renault Zoe demonstrator, a localization system which merges the data of wheel speed, accelerometer, gyrometer, magnetometer and GPS into a position estimation. This relies on an Extended Kalman Filter. This will probably be extended later to consider the localization with respect to roads identified on a map.

[0] Advanced Driving Assistance System

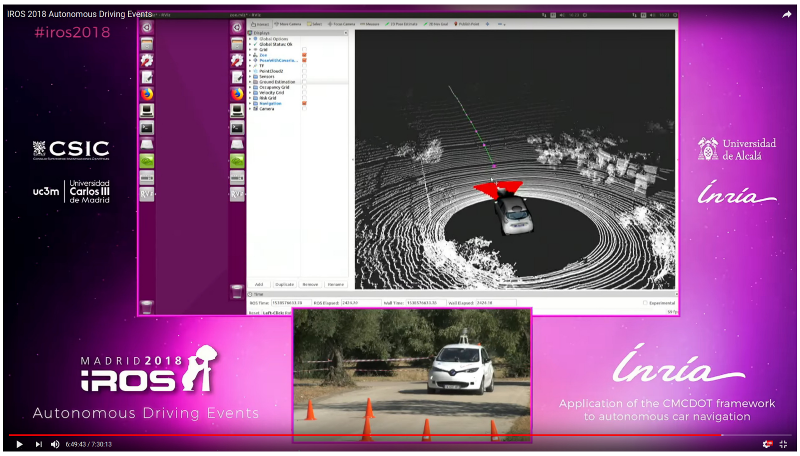


Figure 6. Image taken from the live diffusion of the Autonomous Vehicles event at IROS2018. The demonstrator Renault Zoe is about to go through an obstacle course

Finally a Dijkstra Algorithm have been tested in simulation to define a global navigation path allowing management of waypoints to give to the DWA planner for local navigation.

7.2. Situation Awareness & Decision-making

Participants: Christian Laugier, Olivier Simonin, Jilles Dibangoye, David Sierra-Gonzalez, Mathieu Barbier, Victor Romero-Cano [Universidad Autónoma de Occidente, Cali, Colombia], Ozgur Erkent, Christian Wolf.

7.2.1. Dense & Robust outdoor perception for autonomous vehicles

Participants: Christian Laugier, Victor Romero-Cano, Özgür Erkent, Christian Wolf.

Robust perception plays a crucial role in the development of autonomous vehicles. While perception in normal and constant environmental conditions has reached a plateau, robustly perceiving changing and challenging environments has become an active research topic, particularly due to the safety concerns raised by the introduction of autonomous vehicles to public streets. Solving the robustness issue in road and urban perception applications is the first challenge. Then, it is also mandatory to develop an appropriate framework for extracting relevant semantic information. Our approach is to reason about vision-based data and the output of our grid-based multi-sensors perception approach (see previous section).

The work presented in this section has partly been done in 2017 and completed in 2018, in the scope of our collaboration with Toyota Motor Europe (TME). The main objective was to develop a framework for integrate the outcomes of the deep learning methods with a well-established area, occupancy grids obtained with a Bayesian filtering method in the grid space.

In this work, we are interested in 2D egocentric representations. We propose a method, which estimates an occupancy grid containing detailed semantic information. The semantic characteristics include classes like *road, car, pedestrian, sidewalk, building, vegetation, etc.*. To this end, we leverage and fuse information from multiple sensors including Lidar, odometry and monocular RGB video. To benefit from the respective advantages of the two different methodologies, we propose a hybrid approach leveraging i) the high-capacity of deep neural networks as well as ii) Bayesian filtering, which is able to model uncertainty in a unique way.

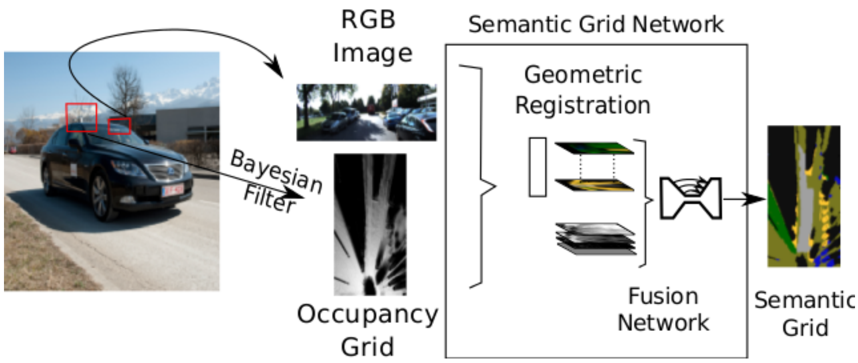


Figure 7. The Semantic Grid framework.

In the system depicted by Figure 7, Bayesian particle filtering processes the Lidar data as well as odometry information from the vehicle's motion in order to robustly estimate an egocentric bird's eye view in the form of an occupancy grid. This grid contains a 360° view of the environment around the car and integrates information from the observation history through temporal filtering; however, it does not include fine-grained semantic classes.

Deep Learning is used for two different tasks in our work. Firstly, a deep network performs semantic segmentation of monocular RGB images. This network has been pre-trained on large scale datasets for image classification and fine-tuned on the vehicle datasets. Secondly, a deep network fuses the occupancy grid with the segmented image of the projective view in order to estimate the semantic grid. Since the occupancy grid is dense, the semantic grid is also expected to be dense. We pay particular attention to correctly model the transformation from the egocentric projective view of the RGB image to the bird's eye view of the occupancy grid as input to the neural network. This work was filed for a patent [98] and published in [28], [14].

Novel approach: Semantic Grid Estimation with a Hybrid Bayesian and Deep Neural Network Approach.

Current and future work in the scope of our collaboration with TME, aims at constructing *Semantic Occupancy Grids*. We propose a hybrid approach, which combines the advantages of Bayesian filtering and deep neural networks. Bayesian filtering provides robust temporal/geometrical filtering and integration and allows for modelling of uncertainty. RGB information and deep neural networks provide knowledge about the semantic class labels like *sideway* vs *road*. The fusion process is fully learned and due to dense structure of occupancy grid, we can construct a dense semantic grid even if we have a sparse point cloud.

7.2.2. Towards Human-Like Motion Prediction and Decision-Making in Highway Scenarios

Participants: David Sierra González, Victor Romero-Cano, Özgür Erkent, Jilles Dibangoye, Christian Laugier.

The objective is to develop human-like motion prediction and decision-making algorithms to enable automated driving in highways. This research work is done in the scope of the Inria-Toyota long-term cooperation on Autonomous Driving and of the PhD thesis work of David Sierra González.

Previous work from our team has shown the predictive potential of driver behavioral models learned from demonstrations using Inverse Reinforcement Learning (IRL) [87] [88]. Unfortunately, these models are hard to learn from real-world driving data due to the inability of traditional IRL algorithms to handle continuous state spaces and dynamic environments. To facilitate this task, we have proposed in 2018 an approximated IRL algorithm for driver behavior modeling that successfully scales to continuous spaces with moving obstacles, by leveraging a spatio-temporal trajectory planner [35]. The proposed algorithm was validated using real-world data gathered with an instrumented vehicle. As an example, Figure 8 shows the similarity between the trajectory obtained using a driver model learned with the proposed method and that of a real human driver in a highway overtake scenario. Current efforts are directed towards integrating the learned behavioral models and the predictive models developed in the scope of this project into a decision-making framework for highways. David Sierra González will defend his PhD thesis in March 2019.

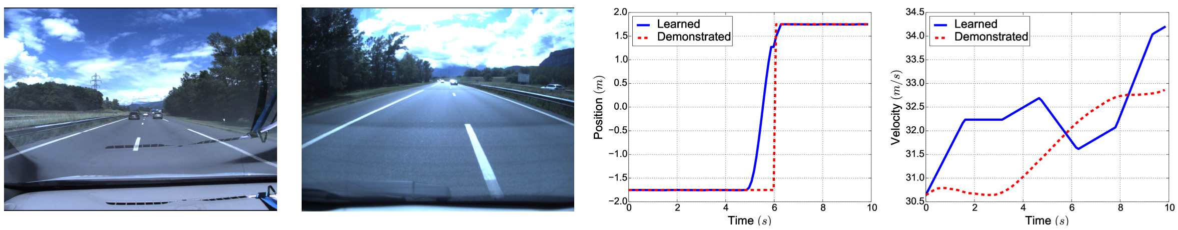


Figure 8.

Comparison of the trajectory obtained with a driver model learned from demonstrated driving data using the method proposed in [35] and that of a human driver for a typical highway overtake scenario a. Front view at $t = 5.0$ b. Back view at $t = 5.0$ c. Position prediction d. Velocity prediction

7.2.3. Decision-making for safe road intersection crossing

Participants: Mathieu Barbier, Christian Laugier, Olivier Simonin.

Road intersections are probably the most complex segment in a road network. Most major accidents occur at intersections, mainly caused by human errors due to failures in fully understanding the encountered situations. Indeed, as drivers approach a road intersection, they must assess the situation and quickly adapt their behaviour accordingly. When this task is performed by a computer, the available information is partial and uncertain. Any decision requires the system to use this information as well as taking into account the behaviour of other drivers to avoid collisions. However, metrics such as collision rate can remain low in an interactive environment because of other driver's actions. Consequently, evaluation metrics must depend on other driving aspects.

In this framework, we developed a decision-making mechanism and designed metrics to evaluate such a system at road intersection crossing [22]. For the former, a Partially Observable Markov Decision Process (POMDP) is used to model the system with respect to uncertainties in the behaviour of other drivers. For the latter, different key performance indicators are defined to evaluate the resulting behaviour of the system in different configurations and scenarios. The approach has been demonstrated within an automotive grade simulator.

Current work aims at increasing the complexity of the scenario, to include pedestrians and more vehicles, and improving the model used for the dynamics of the vehicle and the observation of the physical state to get closer to real world scenarios.

This work has been carried out in the framework of the PhD thesis of Mathieu Barbier, which will be defended in the first trimester of 2019.

7.3. Robust state estimation (Sensor fusion)

This research is the follow up of Agostino Martinelli's investigations carried out during the last five years, which are in the framework of the visual and inertial sensor fusion problem and the unknown input observability problem.

7.3.1. Visual-inertial structure from motion

Participants: Agostino Martinelli, Alexander Oliva, Alessandro Renzaglia.

During this year, we have obtained the full analytic solution of the cooperative visual inertial sensor fusion problem in the case of two agents, starting from the closed-form solution obtained in the last years (this latter solution will be published on the journal of Autonomous Robots [76]). Additionally, we also validated this solution with real experiments and in particular we showed that the analytic solution significantly outperforms our previous closed-form solution in [76]. The new analytic solution has been accepted for publication by the IEEE Robotics and Automation Letters [13].

Specifically, we obtained the analytic solution of the problem by first proving that, this sensor fusion problem, is equivalent to a simple polynomial equations system that consists of several linear equations and three polynomial equations of second degree. The analytic solution of this polynomial equations system was easily obtained by using an algebraic method (developed by Bernard Mourrain, the leader of AROMATH at Inria Sophia Antipolis). The power of the analytic solution is twofold. From one side, it allows us to determine the relative state between the agents (i.e., relative position, speed and orientation) without the need of an initialization. From another side, it provides fundamental insights into all the theoretical aspects of the problem. During this year, we focused on the first issue. Our next objective is to exploit the analytic solution to obtain basic structural properties of the problem.

7.3.2. Unknown Input Observability

Participant: Agostino Martinelli.

The Unknown Input Observability problem (UIO) in the nonlinear case was an open problem since the sixties years, when it was solved only in the linear case. In the last five years, I have obtained its general analytic solution. The mathematics apparatus necessary to obtain this solution is very sophisticated and is based on Ricci calculus, borrowed from theoretical physics. On the other hand, this mathematics can be avoided in the case of driftless systems and characterized by a single unknown input.

All the results (i.e., in the general case that also accounts for a drift and more than one unknown input) are fully described in a book available on ArXiv (arXiv:1704.03252).

During this year, my effort was devoted to make the analytic derivation of the solution palatable for a large audience (in particular, without knowledge of Ricci calculus). Hence, I focused on the simple case of a single unknown input and without drift. This solution has been published on a full paper on the IEEE Transaction on Automatic Control [75].

Regarding the general case available on ArXiv (arXiv:1704.03252), I was invited by the SIAM to write a book, palatable for a large audience. The scope of writing this book, is to present to the control theory and information theory communities a very powerful mathematics framework borrowed from theoretical physics. This could provide the possibility of revisiting many aspects of the control and information theory and bring new fundamental results, open new research domains etc. In this sense the book could be the kick-off of a new season of research in control and information theory. This will be the objective of the next years.

7.4. Motion-planning in human-populated environment

We study new motion planning algorithms to allow robots/vehicles to navigate in human populated environment, and to predict human motions. Since 2016, we investigate several directions exploiting vision sensors : prediction of pedestrian behaviors in urban environments (extended GHMM), mapping of human flows (statistical learning), and learning task-based motion planning (RL+Deep-Learning) . These works are presented here after.

7.4.1. Urban Behavioral Modeling

Participants: Pavan Vasishta, Anne Spalanzani, Dominique Vaufreydaz.

The objective of modeling urban behavior is to predict the trajectories of pedestrians in towns and around car or platoons (PhD work of P. Vasishta). In 2017 we proposed to model pedestrian behaviour in urban scenes by combining the principles of urban planning and the sociological concept of Natural Vision. This model assumes that the environment perceived by pedestrians is composed of multiple potential fields that influence their behaviour. These fields are derived from static scene elements like side-walks, cross-walks, buildings, shops entrances and dynamic obstacles like cars and buses for instance. This work was published in [95], [94]. In 2018, an extension to the Growing Hidden Markov Model (GHMM) method has been proposed to model behavior of pedestrian without observed data or with very few of them. This is achieved by building on existing work using potential cost maps and the principle of Natural Vision. As a consequence, the proposed model is able to predict pedestrian positions more precisely over a longer horizon compared to the state of the art. The method is tested over legal and illegal behavior of pedestrians, having trained the model with sparse observations and partial trajectories. The method, with no training data (see. Fig. 9.a), is compared against a trained state of the art model. It is observed that the proposed method is robust even in new, previously unseen areas. This work was published in [36] and won the **best student paper** of the conference.

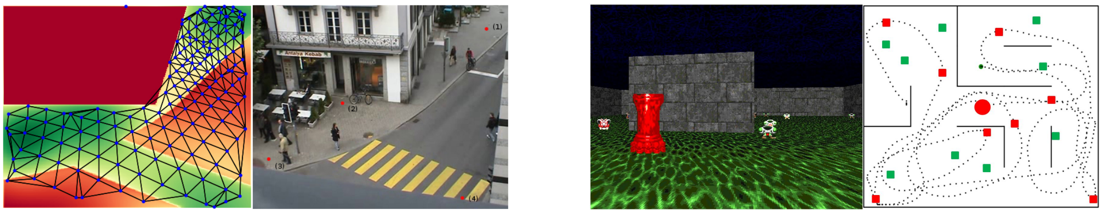


Figure 9.

a. Prior Topological Map of the dataset from the Traffic Anomaly Dataset : first figure shows the generated potential cost map and second figure the "Prior Topology" of the image from scene.
b. Illustration of learning task-based motion planning.

7.4.2. Learning task-based motion planning

Participants: Christian Wolf, Jilles Dibangoye, Laetitia Matignon, Olivier Simonin, Edward Beeching.

Our goal is the automatic learning of robot navigation in human populated environments based on specific tasks and from visual input. The robot automatically navigates in the environment in order to solve a specific problem, which can be posed explicitly and be encoded in the algorithm (e.g. recognize the current activities of all the actors in this environment) or which can be given in an encoded form as additional input. Addressing these problems requires competences in computer vision, machine learning, and robotics (navigation and paths planning).

We started this work in the end of 2017, following the arrival of C. Wolf, through combinations of reinforcement learning and deep learning. The underlying scientific challenge here is to automatic learn representations which allow the agent to solve multiple sub problems require for the task. In particular, the robot needs to learn a metric representation (a map) of its environment based from a sequence of ego-centric observations. Secondly, to solve the problem, it needs to create a representation which encodes the history of ego-centric observations which are relevant to the recognition problem. Both representations need to be connected, in order for the robot to learn to navigate to solve the problem. Learning these representations from limited information is a challenging goal. This is the subject of the PhD thesis of Edward Beeching who started on October 2018, see illustration Fig. 9.b.

7.4.3. Human-flows modeling and social robots

Participants: Jacques Saraydaryan, Fabrice Jumel, Olivier Simonin, Benoit Renault, Laetitia Matignon, Christian Wolf.

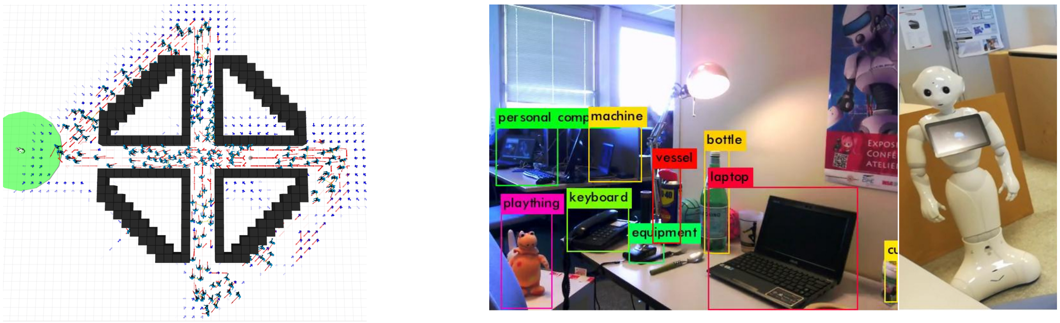


Figure 10.

(a) Flow-grid mapping in a cross-corridor where 200 moving pedestrians turns
(b) Object detection with Pepper based on vision/deep learning techniques.

In order to deal with robot/humanoid navigation in complex and populated environments such as homes, we investigate since 2 years several research avenues :

- Mapping humans flows. We defined a statistical learning approach (ie. a counting-based grid model) exploiting only data from robots embedded sensors. See illustration in Fig. 10.a and publication [66].
- Path-planning in human flows. We revisited the A* path-planning cost function under the hypothesis of the knowledge of a flow grid. See publication [66].
- In 2018 we started to study NAMO problems (Navigation Among Movable Obstacles) by considering populated environments and muti-robot cooperation. After his Master thesis on this subject, Benoit Renault started a PhD in Chroma focusing on the extension of NAMO algorithms to such dynamic environments.
- RoboCup competition. In the context of the **RoboCup** international competition, we created the 'LyonTech' team, joining members from Chroma (INSA/CPE/UCBL). We investigated several humanoid tasks in home environments with our Pepper robot : social aware architecture, decision making and navigation, deep-learning based human and object detection (see Fig. 10.b), human-robot interaction. In July 2018, we participated for the first time to the RoboCup and reaching the 5th rank of the SSL league (Pepper@home). We also published our social-aware architecture to the RoboCup Conference [31]. In October 2018, we qualified for the next final phase of RoboCup SSL (Pepper) to be organized on July 2019, in Sydney.

7.5. Decision Making in Multi-Robot Systems

7.5.1. Multi-robot planning in dynamic environments

7.5.1.1. Global-local optimization in autonomous multi-vehicles systems

Participants: Guillaume Bono, Jilles Dibangoye, Laetitia Matignon, Olivier Simonin, Florian Peyreron [VOLVO Group, Lyon].

This work is part of the PhD. thesis in progress of Guillaume Bono, with the VOLVO Group, in the context of the INSA-VOLVO Chair. The goal of this project is to plan and learn at both global and local levels how to act when facing a vehicle routing problem (VRP). We started with a state-of-the-art paper on vehicle routing problems as it currently stands in the literature [53]. We were surprise to notice that few attention has been devoted to deep reinforcement learning approaches to solving VRP instances. Hence, we investigated our own deep reinforcement learning approach that can help one vehicle to learn how to generalize strategies from solved instances of travelling salesman problems (an instance of VRPs) to unsolved ones. The difficulty of this problem lies in the fact that its Markov decision process' formulation is intractable, i.e., the number of states grows doubly exponentially with the number of cities to be visited by the salesman. To gain in scalability, we build inspiration on a recent work by DeepMind, which suggests using pointer-net, i.e., a novel deep neural network architecture, to address learning problems in which entries are sequences (here cities to be visited) and output are also sequences (here order in which cities should be visited). Preliminary results are encouraging and we are extending this work to the multi-agent setting.

7.5.1.2. *Multi-Robot Routing (MRR) for evolving missions*

Participants: Mihai Popescu, Olivier Simonin, Anne Spalanzani, Fabrice Valois [INSA/Inria, Agora team].

After considering Multi-Robot Patrolling of known targets in 2016 [81], we generalized to MRR (multi-robot routing) and to DMRR (Dynamic MRR) in the work of the PhD of M. Popescu. Target allocation problems have been frequently treated in contexts such as multi-robot rescue operations, exploration, or patrolling, being often formalized as multi-robot routing problems. There are few works addressing dynamic target allocation, such as allocation of previously unknown targets. We recently developed different solutions to variants of this problem :

- MRR : Multi-robot routing has been the main testbed in the domain of multi-robot task allocation, where decentralized solutions consist in auction-based methods. Our work addresses the MRR problem and proposes MRR with saturation constraints (MRR-Sat), where the cost of each robot treating its allocated targets cannot exceed a bound (called saturation). We provided a NP-Complete proof for the problem of MRR-Sat. Then, we proposed a new auction-based algorithm for MRR-Sat and MRR, which combines ideas of parallel allocations with target-oriented heuristics. An empirical analysis of the experimental results shows that the proposed algorithm outperforms state-of-the art methods, obtaining not only better team costs, but also a much lower running time. Results are submitted to RSS'2019 conference.
- DMRR : we defined the Dynamic-MRR problem as the continuous adaptation of the ongoing robot missions to new targets. We proposed a framework for dynamically adapting the existent robot missions to new discovered targets. Dynamic saturation-based auctioning (DSAT) is proposed for adapting the execution of robots to the new targets. Comparison was made with algorithms ranging from greedy to auction-based methods with provable sub-optimality. The results for DSAT shows it outperforms state-of-the-art methods, like standard SSI or SSI with regret clearing, especially in optimizing the target allocation w.r.t. the target coverage in time and the robot resource usage (e.g. minimizing the worst mission cost). First results have been published in [34].
- Synchronization : When patrolling targets along bounded cycles, robots have to meet periodically to exchange information, data (e.g. results of their tasks). Data will finally reach a delivery point (e.g. the base station). Hence, patrolling cycles sometimes have common points (rendezvous points), where the information needs to be exchanged between different cycles (robots). We investigated this problem by defining the following first solutions : random-wait, speed adaptation (first-multiple), primality of periods, greedy interval overlapping. We developed a simulator, allowing experiments that show the approaches have different performances and robustness. This work will be submitted to IROS'2019 conference.
- PHC DRONEM [0] : We started a collaboration in 2017 with the team of Prof. Gabriela Czibula from Babes-Bolyai University in Cluj-Napoca, Romania. The DRONEM project focuses on optimization and online adaptation of the multi-cycle patrolling with machine learning (RL) techniques in order to deal with the arrival of new targets in the environment.

[0] Hubert Curien Partnership

7.5.1.3. Middleware for open multi-robot systems

Participants: Stefan Chitic, Julien Ponge [INSA/CITI, Dynamid], Olivier Simonin.

Multi-robots systems (MRS) require dedicated software tools and models to face the complexity of their design and deployment. In the context of the PhD work of Stefan Chitic, we addressed service self-discovery and property proofs in an ad-hoc network formed by a fleet of robots. This led us to propose a robotic middleware, SDfR, that is able to provide service discovery, see [54]. In 2017, we defined a tool-chain based on timed automata, called ROSMDB, that offers a framework to formalize and implement multi-robot behaviors and to check some (temporal) properties (both offline and online). Stefan Chtic defended his Phd thesis on March 2018 [11].

7.5.2. Multi-robot Coverage and Mapping

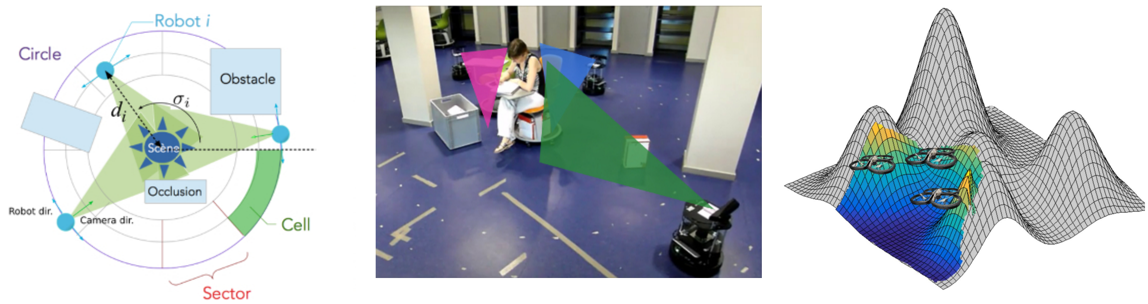


Figure 11. (a) Concentric navigation model and (b) its experimental setup. (c) Illustration of the local search method for multi-UAV coverage.

7.5.2.1. Human scenes observation

Participants: Laetitia Matignon, Olivier Simonin, Stephane d'Alu, Christian Wolf.

Solving complex tasks with a fleet of robots requires to develop generic strategies that can decide in real time (or time-bounded) efficient and cooperative actions. This is particularly challenging in complex real environments. To this end, we explore anytime algorihms and adaptive/learning techniques.

The "CROME" and "COMODYS" [0] projects [0] are motivated by the exploration of the joint-observation of complex (dynamic) scenes by a fleet of mobile robots. In our current work, the considered scenes are defined as a sequence of activities, performed by a person in a same place. Then, mobile robots have to cooperate to find a spatial configuration around the scene that maximizes the joint observation of the human pose skeleton. It is assumed that the robots can communicate but have no map of the environment and no external localisation.

To attack the problem, we proposed an original concentric navigation model allowing to keep easily each robot camera towards the scene (see fig. 11.a). This model is combined with an incremental mapping of the environment and exploration guided by meta-heuristics in order to limit the complexity of the exploration state space. Results have been published in AAMAS'2018 [32]. An extended version has been submitted to the Journal JAAMAS.

[0] COoperative Multi-robot Observation of DYnamic human poSes

[0] Funded by a LIRIS transversal project in 2016-2017 and a FIL project in 2017-2019 (led by L. Matignon)

For experiment with multi-robot systems, we defined an hybrid metric-topological mapping. Robots individually build a map that is updated cooperatively by exchanging only high-level data, thereby reducing the communication payload. We combined the on-line distributed multi-robot decision with this hybrid mapping. These modules has been evaluated on our platform composed of several Turtlebots2, see fig. 11.b. This robotic architecture has been presented in [77] (ECMR). A Demo has been done in AAMAS'2018 international conference [33].

7.5.2.2. Multi-UAV Visual Coverage of Partially Known 3D Surfaces

Participants: Alessandro Renzaglia, Olivier Simonin, Jilles Dibangoye, Vincent Le Doze.

It has been largely proved that the use of Unmanned Aerial Vehicles (UAVs) is an efficient and safe way to deploy visual sensor networks in complex environments. In this context, a widely studied problem is the cooperative coverage of a given environment. In a typical scenario, a team of UAVs is called to achieve the mission without a perfect knowledge on the environment and needs to generate the trajectories on-line, based only on the information acquired during the mission through noisy measurements. For this reason, guaranteeing a global optimal solution of the problem is usually impossible. Furthermore, the presence of several constraints on the motion (collision avoidance, dynamics, etc.) as well as from limited energy and computational capabilities, makes this problem particularly challenging.

Depending on the sensing capabilities of the team (number of UAVs, range of on-board sensor, etc.) and the dimension of the environment to cover, different formulations of this problem can be considered. We firstly approached the deployment problem, where the goal is to find the optimal static UAVs configuration from which the visibility of a given region is maximized. A suitable way to tackle this problem is to adopt derivative-free optimization methods based on numerical approximations of the objective function. In 2012, Renzaglia et al. [82] proposed an approach based on a stochastic optimization algorithm to obtain a solution for arbitrary, initially unknown 3D terrains (see fig. 11.c). However, adopting this kind of approaches, the final configuration can be strongly dependent on the initial positions and the system can get stuck in local optima very far from the global solution. We identified that a way to overcome this problem can be found in initializing the optimization with a suitable starting configuration. An a priori partial knowledge on the environment is a fundamental source of information to exploit to this end. The main contribution of our work is thus to add another layer to the optimization scheme in order to exploit this information. This step, based on the concept of Centroidal Voronoi Tessellation, will then play the role of initialization for the on-line, measurement-based local optimizer. The resulting method, taking advantages of the complementary properties of geometric and stochastic optimization, significantly improves the result of the previous approach and notably reduces the probability of a far-to-optimal final configuration. Moreover, the number of iterations necessary for the convergence of the on-line algorithm is also reduced. This work led to a paper submitted to AAMAS 2019 [0], currently under review. The development of a realistic simulation environment based on Gazebo is an important on-going activity in Chroma and will allow us to further test the approach and to prepare the implementation of this algorithm on the real robotic platform of the team.

We are currently also investigating the dynamic version of this problem, where the information is collected along the trajectories and the environment reconstruction is obtained from the fusion of the total visual data.

7.5.3. Sequential decision-making

This research is the follow up of a group led by Jilles S. Dibangoye carried out during the last three years, which include foundations of sequential decision making by a group of cooperative or competitive robots or more generally artificial agents. To this end, we explore combinatorial, convex optimization and reinforcement learning methods.

[0] A. Renzaglia, J. Dibangoye, V. Le Doze and O. Simonin, "Multi-UAV Visual Coverage of Partially Known 3D Surfaces: Voronoi-based Initialization to Improve Local Optimizers", International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), 2019, *under review*.

7.5.3.1. Optimally solving cooperative and competitive games as continuous Markov decision processes

Participants: Jilles S. Dibangoye, Olivier Buffet [Inria Nancy], Vincent Thomas [Inria Nancy], Christopher Amato [Univ. New Hampshire], François Charpillet [Inria Nancy, Larsen team].

Our major findings this year include:

1. (Theoretical) – As an extension of [58] in the cooperative case [44], we characterize the optimal solution of partially observable stochastic games.
2. (Theoretical) – We further exhibit new underlying structures of the optimal solution for both cooperative and non-cooperative settings.
3. (Algorithmic) – We extend a non-trivial procedure in [27] for computing such optimal solutions when only an incomplete knowledge about the model is available.

This work proposes a novel theory and algorithms to optimally solving a two-person zero-sum POSGs (zs-POSGs). That is, a general framework for modeling and solving two-person zero-sum games (zs-Games) with imperfect information. Our theory builds upon a proof that the original problem is reducible to a zs-Game—but now with perfect information. In this form, we show that the dynamic programming theory applies. In particular, we extended Bellman equations [50] for zs-POSGs, and coined them maximin (resp. minimax) equations. Even more importantly, we demonstrated Von Neumann & Morgenstern's minimax theorem [99] [100] holds in zs-POSGs. We further proved that value functions—solutions of maximin (resp. minimax) equations—yield special structures. More specifically, the maximin value functions are convex whereas the minimax value functions are concave. Even more surprisingly, we prove that for a fixed strategy, the optimal value function is linear. Together these findings allow us to extend planning and learning techniques from simpler settings to zs-POSGs. To cope with high-dimensional settings, we also investigated low-dimensional (possibly non-convex) representations of the approximations of the optimal value function. In that direction, we extended algorithms that apply for convex value functions to lipschitz value functions [27].

7.5.3.2. Learning to act in (continuous) decentralized partially observable Markov decision process

Participants: Jilles S. Dibangoye, Olivier Buffet [Inria Nancy].

During the last year, we investigated deep and standard reinforcement learning for solving decentralized partially observable Markov decision processes. Our preliminary results include:

1. (Theoretical) Proofs that the optimal value function is linear in the occupancy-state space, the set of all possible distributions over hidden states and histories.
2. (Algorithmic) Value-based and policy-based (deep) reinforcement learning for common-payoff partially observable stochastic games.

This work addresses a long-standing open problem of Multi-Agent Reinforcement Learning (MARL) in decentralized stochastic control. MARL previously applied to finite decentralized decision making with a focus on team reinforcement learning methods, which at best lead to local optima. In this research, we build on our recent approach [44], which converts the original problem into a continuous-state Markov decision process, allowing knowledge transfer from one setting to the other. In particular, we introduce the first optimal reinforcement learning method for finite cooperative, decentralized stochastic control domains. We achieve significant scalability gains by allowing the latter to feed deep neural networks. Experiments show our approach can learn to act optimally in many finite decentralized stochastic control problems from the literature [43], [26].

7.5.3.3. Study of policy-gradient methods for decentralized stochastic control

Participants: Guillaume Bono, Jilles S. Dibangoye, Laëtitia Matignon, Olivier Simonin, Florian Peyreron [VOLVO Group, Lyon].

This work is part of the Ph.D. thesis in progress of Guillaume Bono, with VOLVO Group, in the context of the INSA-VOLVO Chair. The work aims at investigating an attractive family of reinforcement learning methods, namely policy-gradient and more generally actor-critic methods for solving decentralized partially observable Markov decision processes. Our preliminary results include:

1. (Theoretical) Proofs of the policy-gradient theorems for both total- and discounted-reward criteria in decentralized stochastic control.
2. (Algorithmic) (deep) actor-critic reinforcement learning methods for centralized and decentralized stochastic control.

Reinforcement Learning (RL) for decentralized partially observable Markov decision processes (Dec-POMDPs) is lagging behind the spectacular breakthroughs of single-agent RL. That is because assumptions that hold in single-agent settings are often obsolete in decentralized multi-agent systems. To tackle this issue, we investigate the foundations of policy gradient methods within the centralized training for decentralized control (CTDC) paradigm. In this paradigm, learning can be accomplished in a centralized manner while execution can still be independent. Using this insight, we establish policy gradient theorem and compatible function approximations for decentralized multi-agent systems. Resulting actor-critic methods preserve the decentralized control at the execution phase, but can also estimate the policy gradient from collective experiences guided by a centralized critic at the training phase. Experiments demonstrate our policy gradient methods compare favorably against standard RL techniques in benchmarks from the literature [42], [23]. Guillaume Bono also designed a simulator for urban logistic reinforcement learning, namely SULFR [39].

7.5.3.4. Towards efficient algorithms for two-echelon vehicle routing problems

Participants: Mohamad Hobballah, Jilles S. Dibangoye, Olivier Simonin, Elie Garcia [VOLVO Group, Lyon], Florian Peyreron [VOLVO Group, Lyon].

During the last year, Mohamad Hobballah (post-doc INSA VOLVO Chair) investigated efficient meta-heuristics for solving two-echelon vehicle routing problems (2E-VRPs) along with realistic logistic constraints. Algorithms for this problem are of interest in many real-world applications. Our short-term application targets goods delivery by a fleet of autonomous vehicles from a depot to the clients through an urban consolidation center using bikers. Preliminary results include:

1. (Methodological) Design of a novel meta-heuristic based on differential evolution algorithm [56] and iterative local search [97]. The former permits us to avoid being attracted by poor local optima whereas the latter performs the local solution improvement.
2. (Empirical) Empirical results on standard benchmarks available at http://www.vrp-rep.org/datasets.html show state-of-the-art performances on most VRP, MDVRP and 2E-VRP instances.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. VOLVO-Renault Trucks Group (2016-2019)

Participants: Olivier Simonin, Jilles Dibangoye, Guillaume Bono, Mohamad Hobballah, Laetitia Matignon.

This collaboration has been built inside the INSA-VOLVO Chair, led by Prof. Didier Remond (INSA). In this context, the Chair funds the PhD Thesis of Guillaume Bono (2016-19) in Chroma. The objective is to study how machine learning techniques can deal with optimization of goods distribution using a fleet of autonomous vehicles. In the following of the first results, VOLVO proposed to extend our collaboration by funding a Post-doc position concerning good distribution with platoons of autonomous vehicles. This is the Post-Doc of Mohamad Hobballah, started on February 2018.

8.1.2. Toyota Motor Europe (2006 - 2018)

Participants: Christian Laugier, David Sierra González, Özgür Erkent, Jilles Dibangoye, Christian Wolf.

The contract with Toyota Motors Europe is a joint collaboration involving Toyota Motors Europe, Inria and ProbaYes. It follows a first successful short term collaboration with Toyota in 2005. This contract aims at developing innovative technologies in the context of automotive safety. The idea is to improve road safety in driving situations by equipping vehicles with the technology to model on the fly the dynamic environment, to sense and identify potentially dangerous traffic participants or road obstacles, and to evaluate the collision risk. The sensing is performed using sensors commonly used in automotive applications such as cameras and lidar.

This collaboration has been extended in 2018 for 4 years (period 2018-2021) and Toyota provides us with an experimental vehicle Lexus equipped with various sensing and control capabilities. Several additional connected technical contracts have also been signed, and an exploitation licence for the *CMCDOT* software has been bought by Toyota in 2018.

8.2. Bilateral Grants with Industry

8.2.1. Renault (2015 - 2018)

Participants: Mathieu Barbier, Christian Laugier, Olivier Simonin.

This contract was linked to the PhD Thesis of Mathieu Barbier (Cifre Thesis). The objective is to develop technologies for collaborative driving as part of a Driving Assistance Systems for improving car safety in road intersections. Both vehicle perception and communications are considered in the scope of this study. Some additional short-term contracts (about 3 months) and an evaluation license for the team *CMCDOT* software have also been signed during this period. *We are on the process of signing a new PhD research agreement for the period 2019 – 2021, with objective to address the open problem of emergency obstacle avoidance in complex traffic situations (for ADAS or AD applications).*

8.2.2. IRT Nanoelec – Security of Autonomous Vehicles project (2018 - 2020)

Participants: Christian Laugier, Lukas Rummelhard, Jerome Lussereau, Jean-Alix David, Thomas Genevois, Nicolas Turro [SED].

Security of Autonomous Vehicles is a project supported by ANR in the scope of the program PULSE of IRT Nanoelec. The objective of this project is to integrate, develop and promote technological bricks of context capture, for the safety of the autonomous vehicle. Building on *Embedded Bayesian Perception for Dynamic Environment*, Bayesian data fusion and filtering technologies from sets of heterogeneous sensors, these bricks make it possible to secure the movements of vehicles, but also provide them with an enriched and useful representation for autonomy functions themselves. In this context, various demonstrators embedding those technology bricks are developed in cooperation with industrial partners.

8.2.3. FUI Tornado (2017 – 2020)

Participants: Rabbia Asghar, Anne Spalanzani, Christian Laugier, Olivier Simonin.

The project Tornado is coordinated by Renault. The academic partners of the project are Inria Grenoble-Rhône Alpes, UTC, Institut Pascal, University of Pau, IFSTTAR. The industrial and application partners are Renault, Easymile, Neavia, Exoskills, 4D-Virtualiz, MBPC and Rambouillet Territoires. The objective of the project is to demonstrate the feasibility of a mobility service systems operating in the commercial zone of Rambouillet and on some public roads located in its vicinity. Several autonomous cars (Autonomous Renault Zoe). The *IRT Nanoelec* is also involved in the project as a subcontractor, for testing the perception, decision-making, navigation and controls components developed in the project.

8.2.4. FUI STAR (2018 – 2021)

Participants: Andres Gomez Hernandez, Olivier Simonin, Christian Laugier.

The Project STAR is coordinated by IVECO. The academic partners of the projects are Inria Grenoble-Rhône, IFSTTAR, ISAE-Supaéro. The industrial and application partners are IVECO, Easymile, Transpolis, Transdev and Sector Groupe. The goal of the project is to build an autonomous bus that will operate on a safe from other vehicle lane but not from pedestrian. Inria is involved in helping design situation awarness perception, specialy in special case like docking at the bus stop and handling dynamicity of any obstacle. The *IRT Nanoelec* is also involved in the project as a subcontractor, for testing the perception, decision-making, navigation and controls components developed in the project.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. COMODYS project, FIL (Federation d'Informatique de Lyon), 2017-19

Participants: Laetitia Matignon, Olivier Simonin.

Project of the Informatics Federation of Lyon (FIL) between two teams of two laboratories: CHROMA (CITI) and SMA (LIRIS), entitled "COoperative Multi-robot Observation of DYnamic human poSes", 2017-2019. Leader : L. Matignon & O. Simonin.

This project funds materials, missions and internships and its objectives are the on-line adaptation of a team of robots that observe and must recognize human activities.

9.1.2. CORDES ADT Inria project, 2017-19

Participants: Olivier Simonin, Vincent Le Doze, Jilles Dibangoye, Alessandro Renzaglia.

The project CORDES (Coordination d'une Flotte de Drones Connectés pour la Cartographie 3D d'édifices) is an Inria ADT coordinated by Olivier Simonin. It funds an Inria expert engineer position in Chroma (Vincent Le Doze, 10/17-11/19) focusing on UAVs control and path-planning. The project aims to deploy a fleet of UAVs able to autonomously fly over an unknown infrastructure and to build a 3D map.

9.2. National Initiatives

9.2.1. ANR

9.2.1.1. ANR "Valet" (2016-18)

The ANR VALET, led by A. Spalanzani, proposes a novel approach for solving car-sharing vehicles redistribution problem using vehicle platoons guided by professional drivers. An optimal routing algorithm is in charge of defining platoons drivers' routes to the parking areas where the followers are parked in a complete automated mode. The consortium is made of 2 academic partners: Inria (RITS, Chroma, Prima) and Ircyyn Ecole Centrale de Nantes and the AKKA company. The phD student (Pavan Vashista) recruited in this project focus on integrating models of human behaviors to evaluate and communicate a risk to pedestrians that may encounter the trajectory of the VALET vehicle. His phD thesis started in february 2016 and is codirected by D. Vaufreydaz (Inria/PervasiveInteraction).

9.2.1.2. ANR "HIANIC" (2017-20)

The HIANIC project, led by A. Spalanzani, proposes to endow autonomous vehicles with smart behaviors (cooperation, negotiation, socially acceptable movements) that better suit complex SharedSpace situations. It will integrate models of human behaviors (pedestrian, crowds and passengers), social rules, as well as smart navigation strategies that will manage interdependent behaviors of road users and of cybercars. The consortium is made of 3 academic partners: Inria (RITS, Chroma, Pervasive Interaction teams), Lig Laboratory (Magma team) and LS2N laboratory (ARMEN and PACCE teams). A. Spalanzani is the leader of this project.

9.2.1.3. PIA Ademe "CAMPUS" (2017-20)

The CAMPUS project aims to identify, develop and deploy new functions for the autonomous cars in urban environments. In this project, Chroma will focus on finding solutions to navigate in complex situations such as crowded environments or dense traffic. The consortium is made of 1 academic partner: Inria (Rits and Chroma teams) and 3 companies: Safran electronics, Gemalto and Valeo.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

Program: ECSEL

Project acronym: ENABLE-S3

Project title: European Initiative to Enable Validation for Highly Automated Safe and Secure Systems

Duration: June 2016 – May 2019

Coordinator: AVL List GesmbH

Other partners: Major European Organizations, including academic partners (such as Inria or KIT) and a Large number of industrial partners from various application domains such as automotive industry or Aeronautics or Train industry

Abstract: ENABLE-S3 is *industry-driven* and therefore aims to foster the leading role of the European industry. This is also reflected in its *use case driven approach*. The main technical objectives are extracted from the use cases defined by the industrial partners, in order to validate the success of the developed methods and tools.

The ENABLE-S3 project will provide European industry with leading-edge technologies that *support the development of reliable, safe and secure functions for highly automated and/or autonomously operating systems* by enabling the *validation and verification at reduced time and costs*.

Enables-S3 is a large European consortium, involving a French consortium leaded by Renault and Inria Grenoble Rhône-Alpes. The Inria Tamis team (Rennes) is also involved in the project.

9.3.2. Collaborations in European Programs, Except FP7 & H2020

9.3.2.1. PHC DRONEM

Program: PHC franco-roumain "Brandusi"

Project acronym: DRONEM

Project title: Optimizing Data Delivery in Multi-robot Network Patrolling using Machine Learning

Duration: 01-2017 - 12-2018

Coordinator: O. Simonin, G. Czibula (University of Babes-Bolyai, Cluj-Napoca, Romania)

Abstract: The present research proposal is an interdisciplinary project that focuses on developing novel machine learning models and techniques for addressing the challenging problem of dynamic multi-robot network patrolling. This proposal brings together a team of researchers in the field of robotics (Chroma) with a team of researchers in the field of Machine Learning from Babe-Bolyai University, Cluj-Napoca (the MLyRE team) and aims to combine their expertise in autonomous robotics and machine learning, as well as to exploit the complementarity between the two fields. Deploying fleets of mobile robots in real scenarios/environments raises several scientific challenges. One of them concerns the ability of the robots to adapt to the complexity of their environment, i.e. its dynamics and uncertainty.

9.3.3. Collaborations with Major European Organizations

Partner 1 : ETHZ, Zurich, Autonomous System laboratory, (Switzerland) and University of Zurich, Robotics and Perception Group (Switzerland)

Subject 1 : Vision and IMU data Fusion for 3D navigation in GPS denied environment.

Partner 2 : Karlsruhe Institut fur Technologie (KIT, Germany)

Subject 2 :Autonomous Driving (student exchanges and common project).

Partner 3 : Vislab Parma (Italy)

Subject 3 : Embedded Perception & Autonomous Driving (visits, projects submissions, and book chapter in the new edition of the Handbook of Robotics).

9.4. International Initiatives

9.4.1. Inria International Labs

9.4.1.1. Informal International Partners

- UC Berkeley & Stanford University (CA, USA)
  Subject: Autonomous Driving (postdoc in the scope of Inria@SV, common publications and patent, visits).
- NUS Singapore & NTU Singapore.
  Subject: Autonomous Driving (visits, common ICT Asia projects, common organization of workshops, review of PhD students).
- Massachussetts Institute of Technology (MIT), Cambridge, MA (USA)
  Subject: Decentralized Control of Markov Decision Processes.
  Subject: Autonomous Driving (visits and common organization of a workshop).

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Visit of 3 researchers (Maria-Iuliana Bocicor, Vlad-Sebastian Ionescu, Ioan-Gabriel Mircea) from University Babes-Bolyai, Cluj-Napoca (Romania). In the context of our PHC project "DRONEM" (2017-18) we worked with them, in Lyon (CITI lab), on Sept. 11-14 2018.
- Jorge Villagra, Senior Scientist at the Center for Automation and Robotics (CSIC-UPM) in Madrid, visited us and given a seminar in novembre 2018. He also co-organized with C.Laugier an Autonomous Vehicle Demonstration event at IEEE IROS 2018 in Madrid (October 2018).

9.5.2. Visits to International Teams

- O. Simonin and J. Dibangoye visited the team of Prof. G. Czibula, at University Babes-Bolyai, Cluj-Napoca (Romania), on April 16-19. The visit was organized in the context of the PHC project "DRONEM" (2017-18). O. Simonin given a talk on the Chroma researches.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- C. laugier was Program Co-Chair of IEEE/RSJ IROS 2018. He also co-organized in the scope of IROS 2018, three interconnected events on Autonomous vehicles: a Workshop having attracted more than 360 people [0]; an Industrial Forum involving international companies(e.g. Renault, Nvidia, Baidu, EasyMile, Ambarella, etc) and having attracted about 80 people; an Autonomous Vehicles Demonstration which involved 5 international teams (including Chroma team with our Autonomous Renault Zoe car) [0].
- C. Laugier has been appointed as General co-Chair for IEEE/RSJ IROS 2019 (Makau).
- O. Simonin was general co-chair of PDIA 2018 "Perspective et Defis de l'IA" with Y. Demazeau (CNRS), organized in Paris Descartes, by AFIA, on October 11. The 2018 topic was "Véhicule Autonome et Intelligence Artificielle" ($\sim$ 80 people).
- O. Simonin is Chair of JFSMA [0] 2019 to be organized at the PFIA national platform, in Toulouse, July 2019.
- O. Simonin is co-chair with F. Chaprillet (Inria Larsen) of the National Conference on Robotics (JNRR), to be organized at Vittel, on October 2019.

[0] see : http://project.inria.fr/ppniv18
[0] see http://avdemo.car.upm-csic.es/
[0] Journées Francophone sur les Systèmes Multi-Agents https://www.irit.fr/pfia2019/jfsma/

10.1.1.2. Member of the Organizing Committees

- C. Laugier co-organized with F. Nashashibi (Inria), Ph. Martinet (Inria) and D. Wang (NTU Singapore) two special sessions on Mobile Robotics and Deep Learning at IEEE ICARCV 2018 (Singapore).
- C. Wolf co-organized the Workshop GDR ISIS & GDR IA on Machine Learning and Reasoning for Signal and Image Processing, October 4th ($\sim$ 90 people).
- F. Jumel is Member of Organizing Committees of the Robocup@Home league.
- F. Jumel is elected member of Technical Committees of the Robocup@Home league.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- O. Simonin is Chair of JFSMA 2019 to be organized at the PFIA national platform, in Toulouse, July 2019.

10.1.2.2. Member of the Conference Program Committees

- C. Laugier was Associate Editor for IEEE ICRA 2018 (Brisbane) and for IEEE ICRA 2019 (Monreal). He was also member of the Senior Program Committee of IEEE/RSJ IROS 2018 (Madrid).
- A. Martinelli was Associate Editor for IEEE ICRA 2019.
- Jilles S. Dibangoye served, in quality of program committee member, for the following conferences: AAAI, IJCAI
- O. Simonin served, in quality of program committee member, for the following conferences : AAMAS (Autonomous Agent and Multi-agent Systems International Conference) Track Robotics, ICAPS (International Conference on Automated Planning and Scheduling) Track Robotics.
- O. Simonin is Program Committee member of the JFSMA conference since 2008 (Journées Francophones sur les Systèmes Multi-Agents).
- C. Wolf served, in quality of program committee member, for the following conferences: IJCAI, BMVC, CVPR 2018 Workshop on Human Pose, Motion, Activities and Shape in 3D, ECCV 2018 Workshop on Hands in Action.

10.1.2.3. Reviewer

- Agostino Martinelli served, in quality of reviewer, for following conferences: ICRA, IROS.
- Jilles S. Dibangoye served, in quality of reviewer, for the following conferences: AAAI, IJCAI, ICRA.
- Olivier Simonin served, in quality of reviewer, for the following conferences: IROS, ICAPS, ACC.
- Anne Spalanzani served, in quality of reviewer, for the following conferences: IROS, RO-MAN.
- C. Wolf served, in quality of reviewer, for the following conferences: CVPR, NIPS, ICLR, ICML, IJCAI, BMVC.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- C. Laugier is Member of the Steering Committee of the journal IEEE Transaction on Intelligent Vehicles.
- C. Laugier is member of the Editorial Board of the journal IEEE ROBOMECH.
- O. Simonin is a member of the editorial board of RIA Revue d'Intelligence Artificielle.

10.1.3.2. Reviewer - Reviewing Activities

- C. Laugier is Senior Editor of the journal IEEE Transaction on Intelligent Vehicles.

- A. Martinelli served, in quality of reviewer, for the following journals: Transaction on Robotics, Transaction on Automatic Control, Robotics and Automation Letters.
- Jilles S. Dibangoye served, in quality of reviewer, for the following journals: Revue d'Intelligence Artificielle, Mathematics and Artificial Intelligence Journal
- O. Simonin served, in quality of reviewer, for the following journals : Autonomous Robots (AURO) and RIA (Revue d'Intelligence Artificielle).
- Anne Spalanzani served, in quality of reviewer, for the International Journal Robotica.

10.1.4. Invited Talks

- O. Simonin was invited for two talks at "Institut d'Automne en IA (IA^3) : Robotics" (GDR IA), on Ocotber 15th-19th, Toulouse.
- O. Simonin gave an invited talk at the "Journée sur les Systèmes Hors Equilibres" ENS Lyon, November 27 : "Coordination d'essaims de robots, vers des systèmes auto-organisés".
- C. Laugier gave an invited tutorial at "Institut d'Automne en IA (IA^3)", GDR IA, Toulouse, October 15-19 2018.*Title: Autonomous Vehicles Technologies for Perception and Decision-making.*
- C. Laugier gave an invited keynote talk at the ECCV 2018 Workshop on "Vision-based Navigation for Autonomous Vehicles", Munich, September 2018.*Title: Dynamic Traffic Scene Understanding using Bayesian Sensor Fusion and Motion Prediction.*
- C. Laugier gave an invited talk at Vedecom SMIV [0] 2018 conference, Versailles, November 2018.*Title: Dynamic Traffic Scene Understanding: Analysis, Prediction and Collision Risk Assessment.*
- C. Laugier gave an invited plenary talk at RWIA [0] 2018 conference, Guangzhou, December 7-10 2018.*Title: Dynamic Scene Understanding and Upcoming Collision Prediction to Improve Autonomous Driving Safety: A Bayesian Approach.*
- A. Spalanzani gave a talk at the "Journée Véhicule autonome et intelligence artificielle" organized by AFIA, Paris Descartes, October 11.
- C. Wolf gave an invited tutorial talk at the national "CORESA" conference, Poitiers, France, on November 12th.
- C. Wolf gave an invited talk at the Strasbourg doctoral school on mathematics, computer science and engineering, Strasbourg, France, on September 21st.
- C. Wolf gave an invited talk at the CVPR 2018 Workshop on "HUMAN 3D: HUman pose, Motion, Activities aNd Shape in 3D", Salt Lake City, USA, June 18th.
- C. Wolf gave an invited talk at the Heudiasyc Laboratory, Compiegne, France, June 5th.
- C. Wolf gave an Invited talk at Inria Stars, Nice, France, May 30th.
- C. Wolf gave an invited talk at the LABRI Laboratory, Bordeaux, France, March 5th.
- C. Wolf gave an invited talk at Inria THOTH, Grenoble, France, February 23rd.
- J.S. Dibangoye was an invited talk at "a NeurIPS-18 international workshop on Deep Reinforcement Learning in partially observable domains" (NeurIPS-18), on December, Montreal, Canada.

[0] Smart Mobility and Intelligent Vehicles
[0] International conference on Robotics Welding, Intelligence and Automation

10.1.5. Leadership within the Scientific Community

- A. Spalanzani served, in quality of Vice President, for the 2018 ANR project selection in Interaction and Robotics.
- C. Laugier is co-chair with Philippe Martinet and Christoph Stiller, of the IEEE RAS Technical Committee on “Autonomous Ground Vehicles and Intelligent Transportation Systems (AGVITS)”.
- C. Laugier is member of the Committee "safety of autonomous vehicles" (committee leaded by ARDI in the scope of the Innovation Regional Strategy).
- C. Laugier is member of the Scientific Committee of the French GDR Robotique.
- C. Laugier is member of several International Award Committee. In 2018, he was Chair of the Best Paper Award Committee of IEEE/RSJ IROS 2018 and Member of several IROS 2018 award committees (Fellow Award, Harashima Award for Innovative Technologies, Distinguished Service Award and Young Professional Award). In 2018, he was also a member of the Best Paper Award Committee of IEEE ICARCV 2018 and Member of the IEEE Chapter Award Committee.
- O. Simonin is an elected member of the Board of AFIA, the French Association for Artificial Intelligence.
- C. Wolf is part of the *Comité de direction* of the GDR ISIS "Information, Signal, Image et ViSion" and co-leads its theme "Machine Learning" (together with N. Thome, CNAM)
- C. Wolf is part of the *Comité scientifique* of the GDR IA "Aspects Formels et Algorithmiques de l’Intelligence Artificielle"
- C. Wolf co-leads the theme "Machine Learning and Robotics" of the GDR Robotique (together with D. Filiat, ENSTA)

10.1.6. Scientific Expertise

- C. Laugier is member of the Advisory Board of ISR University of Coimbra.
- C. Laugier is Scientific Advisor for the ProbaYes SA and for Baidu.
- O. Simonin served, in quality of reviewer, for ANR project submissions.
- J. Dibangoye served, in quality of reviewer, for ANR project submissions.
- C. Wolf was part of the ANR Committee CES 33 "Interaction and Robotics"

10.1.7. Research Administration

- C. Laugier is a member of several Ministerial and Regional French Committees on Robotics and Autonomous Cars.
- O. Simonin is an elected member of the AFIA Council (Association Française pour l’Intelligence Artificielle)
- O. Simonin is member of the Auvergne-Rhone-Alpes Robotics cluster (Coboteam), for Inria and INSA de Lyon entities.
- O. Simonin is member of the Scientific Council of the Digital League (Auvergne-Rhone-Alpes).
- F. Jumel is member of the International RoboCup competition (as Chief Referee for the RoboCup@Home league)
- F. Jumel is member of the board of IMAGINOVE cluster (digital content industry)
- F. Jumel is member of the Rhone-Alpes Robotics cluster (Coboteam)

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

CPE Lyon 4-5th year : F. Jumel, resp. of the Robotics option, 400h M1/ M2, Dept. SN CPE Lyon France.

CPE Lyon 4-5th year : F. Jumel, 250h (robotic vision, cognitive science, Interface robot machine, deeplearning, Robotic frameworks, robotic plateforms, Kalman Filter)

INSA Lyon 3rd year : Jilles S. Dibangoye, Algorithmics, 24h, L3, Dept. Telecom INSA de Lyon, France.

INSA Lyon 3rd year : Jilles S. Dibangoye, WEB, 42h, L3, Dept. Telecom INSA de Lyon, France.

INSA Lyon 3rd year : Jilles S. Dibangoye, Operating Systems, 56h, L3, Dept. Telecom INSA de Lyon, France.

INSA Lyon 4rd year : Jilles S. Dibangoye, Operating Systems, 16h, Master, Dept. Telecom INSA de Lyon, France.

INSA Lyon 5th year : Jilles S. Dibangoye, the Robotics option : AI for Robotics, Robotics projects, 8h, M2, Dept. Telecom INSA de Lyon, France.

M2R MoSIG: A. Martinelli, Autonomous Robotics, 12h, ENSIMAG Grenoble.

INSA Lyon 5th year : O. Simonin, Resp. of the Robotics option (25 students): AI for Robotics, Software and Hardware for robotics, Robotics projects, 90h, M2, Telecom Dept., France.

INSA Lyon 3rd year : O. Simonin, Resp. of Introduction to Algorithmics, 32h (100 students), L3, Telecom Dept., France.

INSA Lyon 5th year : A. Spalanzani, Navigation en environnement humain, 2h, M2, INSA de Lyon, France.

Master : Laetitia Matignon, Multi-Agents and Self-* Systems, 10h TD, M2 Artificial Intelligence, Lyon 1 University, France.

Master : Laetitia Matignon, Multi-Robot Systems, 20h TD, 5th year of engineer, Polytech Lyon Informatics Department, France.

10.2.2. Supervision

PhD in progress: David Sierra Gonzalez, Autonomous Driving (cooperation Toyota), 2014, C. Laugier, J. Dibangoye, E. Mazer (Inria Prima). Defense planned in March 2019.

PhD in progress: Mathieu Barbier, Decision making for Intelligent Vehicles (cooperation Renault), 2015, C. Laugier, O. Simonin and E. Mazer (Inria Pervasive Interaction). Defense planned in April 2019.

PhD in progress: Mihai Popescu, Robot fleet mobility under communication constraints, 2015, O. Simonin, A. Spalanzani, F. Valois (CITI/Inria Agora).

PhD in progress: Pavan Vasishta, Natural vision based perception and prediction, 2016, A. Spalanzani and D. Vaufreydaz (Inria Pervasive Interaction)

PhD in progress: Guillaume Bono, Global-local Optimization Under Uncertainty for Goods Distribution Using a Fleet of Autonomous Vehicles, 2016, O. Simonin, J. Dibangoye, L. Matignon.

PhD in progress: Remy Grunblatt, Mobilité contrôlée dans les réseaux de drones autonomes", 2017, I. Guerrin-Lassous (Inria Dante) and O. Simonin.

Starting PhD: Maria Kabtoul, Proactive Navigation in dense crowds, A. Spalanzani and P. Martinet (Inria Chorale).

Starting PhD: Manon Prédhumeau, Crowd simulation and autonomous vehicle, A. Spalanzani and J. Dugdale (LIG Hawai).

Starting PhD: Luiz Serafim-Guardini, Conduite Automobile Autonome : Utilisation de grilles d'occupation probabilistes dynamiques pour la planification contextualisée de trajectoire d'urgence à criticité minimale, A. Spalanzani, C. Laugier, P. Martinet (Inria Chorale).

Starting PhD: Benoit Renault, Navigation coopérative et sociale de robots mobiles en environnement modifiable, O. Simonin and J. Saraydaryan.

Starting PhD: Edward Beeching, Large-scale automatic learning of autonomous agent behavior with structured deep reinforcement learning, C. Wolf, O. Simonin and J. Dibangoye.

10.2.3. Juries

1. PhD thesis juries
   - C. Laugier was reviewer and president of the defense committee of the PhD thesis of Fernando Ireta Munoz, I3S Sophia-Antipolis, April 4th 2018.
   - C. Laugier was member of the defense committee of the PhD thesis of Tomasz Kucners, Örebro University (Sweden), September 20th 2018.
   - O. Simonin was reviewer and member of the defense committee of the PhD thesis of Nesrine Mahdoui Chedly, Compiegne (UTC), December 7th, 2018.
   - O. Simonin was reviewer and member of the defense committee of the PhD thesis of Thadeu Knychala Tucci, Université Franche Comté, Montbéliard, November 12th, 2018.
   - C. Wolf was reviewer and member of the defense committee of the PhD thesis of Michaël Blot, Sorbonne Université, Paris, November 11th.
   - C. Wolf was reviewer and member of the defense committee of the PhD thesis of Farhood Negin, Université de Nice, Inria Stars, October 15th.
   - C. Wolf was reviewer and member of the defense committee of the PhD thesis of Bruno Stuner, Université de Rouen, June 11th.
   - C. Wolf was member of the defense committee of the PhD thesis of Stéphane Lathuilière, Université de Grenoble, Inria Perception, May 22nd.
   - C. Wolf was reviewer and member of the defense committee of the PhD thesis of Nicolas Chesneau, Université de Grenoble, Inria Thoth, February 2nd.

10.3. Popularization

10.3.1. Articles and contents

- O. Simonin participated to the writing of "Livre blanc sur les Véhicules Autonomes et Connectés" [0]
- O. Simonin given several interviews in order to popularize AI & Robotics, and Autonomous Vehicles: JDN (Journal du Net), INSA letter, etc.
- C. Laugier, J. Lussereau and J.A. David have been interviewed in October 2018 by Radio France in order to popularize Autonomous Vehicles. The interview has been broadcasted by France Bleu Isere on October 11th 2018 at respectively 6:40 a.m and 8:15 a.m.

10.3.2. Education

- J. Dibangoye created a MOOC on deep reinforcement learning for drones, which will be launched in 2019 on OpenClassRooms.

10.3.3. Interventions

- C. Wolf gave a talk for developers at the journéé ARAMIS on May 24th.
- C. Wolf gave a talk for developers at the conference #MixIT, Lyon on April 20th.

10.3.4. Internal action

- Demo "Crazyflie micro-drones" at the Journées Recherche Inria & Industrie (O. Simonin and S. d'Alu), Paris, November 20th. Lien video : https://youtu.be/Tgwrb4xpJOA

[0] https://www.inria.fr/actualite/actualites-inria/livre-blanc-vehicules-autonomes-et-connectes

11. Bibliography

Major publications by the team in recent years

[1] M. ANDRIES, O. SIMONIN, F. CHARPILLET.*Localisation of humans, objects and robots interacting on load-sensing floors*, in "IEEE Sensors Journal", 2015, vol. PP, n^{o} 99, 12 [*DOI :* 10.1109/JSEN.2015.2493122], https://hal.inria.fr/hal-01196042

[2] A. BROGGI, A. ZELINSKY, U. OZGUNER, C. LAUGIER.*Handbook of Robotics 2nd edition, Chapter 62 on "Intelligent Vehicles"*, in "Handbook of Robotics 2nd Edition", B. SICILIANO, O. KHATIB (editors), Springer Verlag, July 2016, https://hal.inria.fr/hal-01260280

[3] J. S. DIBANGOYE, C. AMATO, O. BUFFET, F. CHARPILLET.*Optimally Solving Dec-POMDPs as Continuous-State MDPs*, in "Journal of Artificial Intelligence Research", February 2016, vol. 55, p. 443-497 [*DOI :* 10.1613/JAIR.4623], https://hal.inria.fr/hal-01279444

[4] Ö. ERKENT, C. WOLF, C. LAUGIER, D. SIERRA GONZÁLEZ, V. R. CANO.*Semantic Grid Estimation with a Hybrid Bayesian and Deep Neural Network Approach*, in "IROS 2018 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Madrid, Spain, IEEE, October 2018, p. 1-8, https://hal.inria.fr/hal-01881377

[5] A. MARTINELLI.*State Observability in Presence of Disturbances: the Analytic Solution and its Application in Robotics*, in "IROS 2017 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Vancouver, Canada, September 2017, p. 1-8, https://hal.inria.fr/hal-01669046

[6] L. MATIGNON, O. SIMONIN.*Multi-Robot Simultaneous Coverage and Mapping of Complex Scene - Comparison of Different Strategies*, in "AAMAS 2018 - 17th International Conference on Autonomous Agents and Multiagent Systems - Robotics Track", Stockholm, Sweden, M. DASTANI, G. SUKTHANKAR, E. ANDRE, S. KOENIG (editors), ACM, July 2018, p. 559-567, https://hal.archives-ouvertes.fr/hal-01726120

[7] T. RAKOTOVAO, J. MOTTIN, D. PUSCHINI, C. LAUGIER.*Multi-sensor fusion of occupancy grids based on integer arithmetic*, in "In 2016 IEEE International Conference on Robotics and Automation (ICRA)", 2016

[8] J. RIOS-MARTINEZ, A. SPALANZANI, C. LAUGIER.*From Proxemics Theory to Socially-Aware Navigation: A Survey*, in "International Journal of Social Robotics", April 2015 [*DOI :* 10.1007/S12369-014-0251-1], https://hal.inria.fr/hal-01067278

[9] D. SIERRA GONZÁLEZ, V. ROMERO-CANO, J. STEEVE DIBANGOYE, C. LAUGIER.*Interaction-Aware Driver Maneuver Inference in Highways Using Realistic Driver Models*, in "Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC 2017)", Yokohama, Japan, October 2017, https://hal.inria.fr/hal-01589493

[10] P. VASISHTA, D. VAUFREYDAZ, A. SPALANZANI.*Building Prior Knowledge: A Markov Based Pedestrian Prediction Model Using Urban Environmental Data*, in "ICARCV 2018 - 15th International Conference on Control, Automation, Robotics and Vision (BEST Student Paper)", Singapore, Singapore, November 2018, p. 1-12, https://hal.inria.fr/hal-01875147

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] S.-G. CHITIC.*Middleware and programming models for multi-robot systems*, INSA Lyon, March 2018, https://hal.archives-ouvertes.fr/tel-01809505

Articles in International Peer-Reviewed Journal

[12] A. MARTINELLI.*Nonlinear Unknown Input Observability: Extension of the Observability Rank Condition*, in "IEEE Transactions on Automatic Control", 2018, https://hal.archives-ouvertes.fr/hal-01966303

[13] A. MARTINELLI, A. OLIVA, B. MOURRAIN.*Cooperative Visual-Inertial Sensor Fusion: the Analytic Solution*, in "IEEE Robotics and Automation Letters", 2018, https://hal.archives-ouvertes.fr/hal-01966542

Invited Conferences

[14] Ö. ERKENT, C. WOLF, C. LAUGIER.*Semantic Grid Estimation with Occupancy Grids and Semantic Segmentation Networks*, in "ICARCV 2018 - 15th International Conference on Control, Automation, Robotics and Vision", Singapore, Singapore, November 2018, p. 1-6, https://hal.inria.fr/hal-01933939

[15] C. LAUGIER.*Dynamic Scene Understanding and Upcoming Collision Prediction to improve Autonomous Driving Safety: A Bayesian Approach*, in "RWIA 2018 - International Conference on Robotic Welding, Intelligence and Automation", Guangzhou, China, December 2018, p. 1-26, https://hal.inria.fr/hal-01970468

[16] C. LAUGIER.*Dynamic Scene Understanding for Autonomous vehicles: Analysis, Prediction and Collision Risk Assessment (Invited talk)*, in "SMIV 2018 - Smart Mobility and Intelligent Vehicles Conference", Versailles, France, Vedecom, November 2018, The publication of the presented slides has not been authorized, https://hal.inria.fr/hal-01970436

[17] C. LAUGIER.*Dynamic Traffic Scene Understanding using Bayesian Sensor Fusion and Motion Prediction*, in "Workshop “Vision-based Navigation for Autonomous Driving” - ECCV 2018", Munich, Germany, September 2018, https://hal.inria.fr/hal-01968786

[18] C. LAUGIER.*Improving Autonomous Driving Safety through a better Understanding of Traffic Scenes and of Potential Upcoming Collisions : A Bayesian & Machine Learning Approach (Invited Plenary Speech)*, in "ICARCV 2018 - 15th International Conference on Control, Automation, Robotics and Vision", Singapore, Singapore, November 2018, p. 1-15, Plenary Forum on the “Impact of AI on Robotics and Computer Vision”, https://hal.inria.fr/hal-01969802

International Conferences with Proceedings

[19] F. BARADEL, N. NEVEROVA, C. WOLF, J. MILLE, G. MORI.*Object Level Visual Reasoning in Videos*, in "ECCV 2018 - European Conference on Computer Vision", Munich, Germany, September 2018, p. 1-17, https://hal.inria.fr/hal-01828872

[20] F. BARADEL, C. WOLF, J. MILLE.*Human Activity Recognition with Pose-driven Attention to RGB*, in "BMVC 2018 - 29th British Machine Vision Conference", Newcastle, United Kingdom, September 2018, p. 1-14, https://hal.inria.fr/hal-01828083

[21] F. BARADEL, C. WOLF, J. MILLE, G. W. TAYLOR.*Glimpse Clouds: Human Activity Recognition from Unstructured Feature Points*, in "CVPR 2018 - Computer Vision and Pattern Recognition", Salt Lake City, United States, IEEE, June 2018, p. 1-10, https://hal.inria.fr/hal-01713109

[22] M. BARBIER, C. LAUGIER, O. SIMONIN, J. IBAÑEZ-GUZMÁN.*Probabilistic Decision-Making at Road Intersections: Formulation and Quantitative Evaluation*, in "ICARCV 2018 - 15 th International Conference on Control, Automation, Robotics and Vision", Singapour, Singapore, November 2018, p. 1-8, https://hal.inria.fr/hal-01940392

[23] G. BONO, J. S. DIBANGOYE, L. MATIGNON, F. PEREYRON, O. SIMONIN.*Cooperative Multi-Agent Policy Gradient*, in "ECML-PKDD 2018 - European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases", Dublin, Ireland, M. BERLINGERIO, F. BONCHI, E. CURRY, T. GÄRTNER, G. IFRIM, A. MARASCU (editors), LNCS Proceedings of ECML-PKDD, Springer, September 2018, https://hal.inria.fr/hal-01847775

[24] R. CAMBUZAT, F. ELISEI, G. BAILLY, O. SIMONIN, A. SPALANZANI.*Immersive Teleoperation of the Eye Gaze of Social Robots Assessing Gaze-Contingent Control of Vergence, Yaw and Pitch of Robotic Eyes*, in "ISR 2018 - 50th International Symposium on Robotics", Munich, Germany, VDE, June 2018, p. 232-239, https://hal.archives-ouvertes.fr/hal-01779633

[25] Q. DEBARD, C. WOLF, S. CANU, J. ARNÉ.*Learning to recognize touch gestures: recurrent vs. convolutional features and dynamic sampling*, in "FG 2018 - 13th IEEE International Conference on Automatic Face & Gesture Recognition", Xi'an, China, IEEE, May 2018, p. 114-121, https://arxiv.org/abs/1802.09901 - 9 pages, 4 figures, accepted at the 13th IEEE Conference on Automatic Face and Gesture Recognition (FG2018). Dataset available at http://itekube7.itekube.com [*DOI :* 10.1109/FG.2018.00026], https://hal.inria.fr/hal-01720716

[26] J. S. DIBANGOYE, O. BUFFET.*Learning to Act in Decentralized Partially Observable MDPs*, in "ICML 2018 - 35th International Conference on Machine Learning", Stockholm, Sweden, Proceedings of Machine Learning Research, PMLR, July 2018, vol. 80, p. 1233-1242, https://hal.inria.fr/hal-01851806

[27] J. S. DIBANGOYE, M. FEHR, O. BUFFET, V. THOMAS, J. DIBANGOYE.*ρ-POMDPs have Lipschitz-Continuous ϵ-Optimal Value Functions*, in "NIPS 2018 - Thirty-second Conference on Neural Information Processing Systems", Montréal, Canada, December 2018, https://hal.inria.fr/hal-01903685

[28] Ö. ERKENT, C. WOLF, C. LAUGIER, D. SIERRA GONZÁLEZ, V. R. CANO.*Semantic Grid Estimation with a Hybrid Bayesian and Deep Neural Network Approach*, in "IROS 2018 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Madrid, Spain, IEEE, October 2018, p. 1-8, https://hal.inria.fr/hal-01881377

[29] E. FERRIER-BARBUT, D. VAUFREYDAZ, J.-A. DAVID, J. LUSSEREAU, A. SPALANZANI.*Personal space of autonomous car's passengers sitting in the driver's seat*, in "IV'2018 - The 29th IEEE Intelligent Vehicles Symposium", Changshu, Suzhou, China, IEEE, June 2018, p. 2022-2029, https://arxiv.org/abs/1805.03563 [*DOI :* 10.1109/IVS.2018.8500648], https://hal.inria.fr/hal-01786006

[30] M. GARZÓN, A. SPALANZANI.*An hybrid simulation tool for autonomous cars in very high traffic scenarios*, in "ICARCV 2018 - 15th International Conference on Control, Automation, Robotics and Vision", Singapore, Singapore, November 2018, p. 1-6, https://hal.inria.fr/hal-01872095

[31] F. JUMEL, J. SARAYDARYAN, R. LEBER, L. MATIGNON, E. LOMBARDI, C. WOLF, O. SIMONIN.*Context Aware Robot Architecture, Application to the RoboCup@Home Challenge*, in "RoboCup symposium", Montreal, Canada, June 2018, p. 1-12, https://hal.archives-ouvertes.fr/hal-01832613

[32] L. MATIGNON, O. SIMONIN.*Multi-Robot Simultaneous Coverage and Mapping of Complex Scene - Comparison of Different Strategies*, in "AAMAS 2018 - 17th International Conference on Autonomous Agents and Multiagent Systems - Robotics Track", Stockholm, Sweden, M. DASTANI, G. SUKTHANKAR, E. ANDRE, S. KOENIG (editors), ACM, July 2018, p. 559-567, https://hal.archives-ouvertes.fr/hal-01726120

[33] L. MATIGNON, O. SIMONIN.*Multi-Robot Simultaneous Coverage and Mapping of Complex Scene*, in "AAMAS 2018 - 17th International Conference on Autonomous Agents and Multiagent Systems - Demonstration", Stockholm, Sweden, Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems, July 2018, p. 559-567, https://hal.archives-ouvertes.fr/hal-01774799

[34] M.-I. POPESCU, O. SIMONIN, A. SPALANZANI, F. VALOIS.*DMRR: Dynamic Multi-Robot Routing for Evolving Missions*, in "DMAP 2018 - 6th Workshop on Distributed and Multi-Agent Planning", Delft, Netherlands, June 2018, https://hal.inria.fr/hal-01942845

[35] D. SIERRA GONZÁLEZ, Ö. ERKENT, V. ROMERO-CANO, J. DIBANGOYE, C. LAUGIER.*Modeling Driver Behavior From Demonstrations in Dynamic Environments Using Spatiotemporal Lattices*, in "ICRA 2018 - Proceedings of the 2018 IEEE International Conference on Robotics and Automation", Brisbane, Australia, May 2018, p. 3384-3390 [*DOI :* 10.1109/ICRA.2018.8460208], https://hal.inria.fr/hal-01729960

[36] *Best Paper*
P. VASISHTA, D. VAUFREYDAZ, A. SPALANZANI.*Building Prior Knowledge: A Markov Based Pedestrian Prediction Model Using Urban Environmental Data*, in "ICARCV 2018 - 15th International Conference on Control, Automation, Robotics and Vision", Singapore, Singapore, November 2018, p. 1-12, https://arxiv.org/abs/1809.06045 , https://hal.inria.fr/hal-01875147.

National Conferences with Proceeding

[37] O. BUFFET, J. S. DIBANGOYE, A. SAFFIDINE, V. THOMAS.*Recherche heuristique pour jeux stochastiques (à somme nulle)*, in "JFPDA 2018 - Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes", Nancy, France, July 2018, p. 1-8, https://hal.inria.fr/hal-01840591

[38] J. S. DIBANGOYE, O. BUFFET.*Learning to Act in Continuous Dec-POMDPs*, in "JFPDA 2018 - Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes", Nancy, France, July 2018, p. 1-10, https://hal.inria.fr/hal-01840602

Conferences without Proceedings

[39] G. BONO, J. S. DIBANGOYE, L. MATIGNON, F. PEREYRON, O. SIMONIN.*SULFR: Simulation of Urban Logistic For Reinforcement*, in "PGMRL 2018Workshop on Prediction and Generative Modeling in Reinforcement Learning", Stockholm, Sweden, July 2018, p. 1-5, https://hal.inria.fr/hal-01847773

[40] G. BONO, J. S. DIBANGOYE, L. MATIGNON, F. PEREYRON, O. SIMONIN.*Sur le Gradient de la Politique pour les Systèmes Multi-Agents Coopératifs*, in "JFPDA 2018 - Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes", Nancy, France, July 2018, p. 1-13, https://hal.inria.fr/hal-01840852

[41] J. QUILBEUF, M. BARBIER, L. RUMMELHARD, C. LAUGIER, A. LEGAY, B. BAUDOUIN, T. GENEVOIS, J. IBAÑEZ-GUZMÁN, O. SIMONIN.*Statistical Model Checking Applied on Perception and Decision-making*

Systems for Autonomous Driving, in "PPNIV 2018 - 10th Workshop on Planning, Perception and Navigation for Intelligent Vehicles", Madrid, Spain, October 2018, p. 1-8, https://hal.inria.fr/hal-01888556

Research Reports

[42] G. BONO, J. S. DIBANGOYE, L. MATIGNON, F. PEREYRON, O. SIMONIN.*On the Study of Cooperative Multi-Agent Policy Gradient*, INSA Lyon ; Inria, June 2018, n^o RR-9188, p. 1-27, https://hal.inria.fr/hal-01821677

[43] J. S. DIBANGOYE, O. BUFFET.*Learning to Act in Decentralized Partially Observable MDPs*, Inria Grenoble - Rhone-Alpes - CHROMA Team ; Inria Nancy, équipe LARSEN, June 2018, n^o RR-9179, https://hal.inria.fr/hal-01809897

[44] J. S. DIBANGOYE.*On Markov Policies For Decentralized POMDPs*, Inria Grenoble - Rhone-Alpes - CHROMA Team ; CITI - CITI Centre of Innovation in Telecommunications and Integration of services ; INSA Lyon, August 2018, n^o RR-9202, https://hal.inria.fr/hal-01860060

Other Publications

[45] C. LAUGIER.*Tutorial on Autonomous Vehicles Technologies for Perception & Decision-Making*, October 2018, Lecture, https://hal.inria.fr/cel-01969720

[46] L. RUMMELHARD, J.-A. DAVID, J. LUSSEREAU, T. GENEVOIS, C. LAUGIER.*Application of the CMCDOT framework to autonomous car navigation*, October 2018, https://hal.inria.fr/medihal-01963296

[47] L. N. RUMMELHARD, T. GENEVOIS, J.-A. DAVID, A. NÈGRE, C. LAUGIER.*Real-time Ground Estimation and Point Cloud Segmentation*, March 2018, 1, GTC 2018 - GPU Technology Conference, Poster, https://hal.inria.fr/hal-01903668

References in notes

[48] G. AOUDE, B. LUDERS, K. LEE, D. LEVINE, J. HOW.*Threat assessment design for driver assistance system at intersections*, in "Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on", sept. 2010, p. 1855 -1862

[49] L. ARMESTO, J. TORNERO, M. VINCZE.*Fast Ego-motion Estimation with Multi-rate Fusion of Inertial and Vision*, in "The Int. J. of Rob. Research", 2007, vol. 26, n^o 6, p. 577–589

[50] R. E. BELLMAN.*Dynamic Programming*, Dover Publications, Incorporated, 1957

[51] F. A. W. BELO, P. SALARIS, A. BICCHI.*3 Known Landmarks are Enough for Solving Planar Bearing SLAM and Fully Reconstruct Unknown Inputs*, in "International Conference on Intelligent Robots and Systems (IROS)", IEEE, December 2010

[52] M. BENNEWITZ, W. BURGARD, G. CIELNIAK, S. THRUN.*Learning motion patterns of people for compliant robot motion*, in "Int. J. of Robot. Res.", 2005, vol. 24, p. 31–48

[53] G. BONO, J. S. DIBANGOYE, L. MATIGNON, F. PEREYRON, O. SIMONIN.*Classification des problèmes stochastiques et dynamiques de collectes et de livraisons par des véhicules intelligents*, in "Journées Franco-

phones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes (JFPDA 2017)", Caen, France, Actes des Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes (JFPDA 2017), July 2017, https://hal.archives-ouvertes.fr/hal-01576351

[54] S.-G. CHITIC, J. PONGE, O. SIMONIN.*SDfR -Service discovery for multi-robot systems*, in "ICAART 2016 The 8th International Conference on Agents and Artificial Intelligence", Rome, Italy, February 2016, https://hal.archives-ouvertes.fr/hal-01286895

[55] C. COUÉ, C. PRADALIER, C. LAUGIER, T. FRAICHARD, P. BESSIÈRE.*Bayesian Occupancy Filtering for Multi-Target Tracking: an Automotive Application*, in "Int. Journal of Robotics Research", January 2006, vol. 25, n^{o} 1, p. 19–30

[56] S. DAS, P. N. SUGANTHAN.*Differential evolution: a survey of the state-of-the-art*, in "IEEE transactions on evolutionary computation", 2011, vol. 15, n^{o} 1, p. 4–31

[57] J. S. DIBANGOYE, C. AMATO, O. BUFFET, F. CHARPILLET.*Exploiting Separability in Multiagent Planning with Continuous-State MDPs (Extended Abstract)*, in "24th International Joint Conference on Artificial Intelligence (IJCAI)", 2015, p. 4254–4260

[58] J. S. DIBANGOYE, C. AMATO, O. BUFFET, F. CHARPILLET.*Optimally Solving Dec-POMDPs as Continuous-State MDPs*, in "Journal of Artificial Intelligence Research", February 2016, vol. 55, p. 443-497 [*DOI :* 10.1613/JAIR.4623], https://hal.inria.fr/hal-01279444

[59] G. FERRER, A. SANFELIU.*Bayesian Human Motion Intentionality Prediction in urban environments*, in "Pattern Recognition Letters", 2014, vol. 44, n^{o} 0, p. 134 - 140, Pattern Recognition and Crowd Analysis [*DOI :* 10.1016/J.PATREC.2013.08.013], http://www.sciencedirect.com/science/article/pii/S0167865513003176

[60] C. FORSTER, L. CARLONE, F. DELLAERT, D. SCARAMUZZA.*IMU Preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation*, in "RSS", 2015

[61] C. FORSTER, M. PIZZOLI, D. SCARAMUZZA.*SVO: Fast Semi-Direct Monocular Visual Odometry*, in "ICRA", 2014

[62] C. FULGENZI, C. TAY, A. SPALANZANI, C. LAUGIER.*Probabilistic navigation in dynamic environment using Rapidly-exploring Random Trees and Gaussian Processes*, in "IEEE/RSJ 2008 International Conference on Intelligent RObots and Systems", France Nice, 2008, http://hal.inria.fr/inria-00332595/en/

[63] J. HESCH, D. KOTTAS, S. BOWMAN, S. ROUMELIOTIS.*Consistency analysis and improvement of vision-aided inertial navigation*, in "Trans. on Rob.", 2014, vol. 30, n^{o} 1, p. 158–176

[64] G. HUANG, M. KAESS, J. J. LEONARD.*Towards consistent visual-inertial navigation*, in "Int. Conf. Rob. Aut.", 2015

[65] G. HUANG, A. MOURIKIS, S. ROUMELIOTIS.*An observability-constrained sliding window filter for SLAM*, in "Intelligent Robots and Systems (IROS)", IEEE, 2011, p. 65–72

[66] F. JUMEL, J. SARAYDARYAN, O. SIMONIN.*Mapping likelihood of encountering humans: application to path planning in crowded environment*, in "The European Conference on Mobile Robotics (ECMR)", Paris, France, Proceedings of ECMR 2017, September 2017, https://hal.archives-ouvertes.fr/hal-01588815

[67] S. LEFÈVRE, R. BAJCSY, C. LAUGIER.*Probabilistic Decision Making for Collision Avoidance Systems: Postponing Decisions*, in "IEEE/RSJ International Conference on Intelligent Robots and Systems", Tokyo, Japan, 2013, https://hal.inria.fr/hal-00880440

[68] S. LEUTENEGGER, P. FURGALE, V. RABAUD, M. CHLI, K. KONOLIGE, R. SIEGWART.*Keyframe-based visual-inertial odometry using nonlinear optimization*, in "Int. J. of Rob. Res.", 2014

[69] M. LI, A. I. MOURIKIS.*High-precision, Consistent EKF-based Visual-inertial Odometry*, in "Int. J. Rob. Res.", May 2013, vol. 32, n° 6, p. 690–711, http://dx.doi.org/10.1177/0278364913481251

[70] T. LUPTON, S. SUKKARIEH.*Visual-inertial-aided navigation for high-dynamic motion in built environments without initial conditions*, in "Trans. on Rob.", 2012, vol. 28, n° 1, p. 61–76

[71] A. MARTINELLI.*State Estimation based on the Concept of Continuous Symmetry and Observability Analysis: The Case of Calibration*, in "IEEE Transactions on Robotics", May 2011, http://hal.inria.fr/hal-00578795/en

[72] A. MARTINELLI.*Vision and IMU Data Fusion: Closed-Form Solutions for Attitude, Speed, Absolute Scale and Bias Determination*, in "Transaction on Robotics", 2012, vol. 28, n° 1

[73] A. MARTINELLI.*Observabilty Properties and Deterministic Algorithms in Visual-Inertial Structure from Motion*, in "Foundations and Trends in Robotics (FnTROB)", December 2013, p. 1–75 [*DOI :* 10.1088/1742-5468/2014/03/P03003], https://hal.inria.fr/hal-01096948

[74] A. MARTINELLI.*Closed-form solution of visual-inertial structure from motion*, in "International Journal of Computer Vision", January 2014, vol. 106, n° 2, p. 138–152, https://hal.archives-ouvertes.fr/hal-00905881

[75] A. MARTINELLI.*Nonlinear Unknown Input Observability: Extension of the Observability Rank Condition*, in "IEEE Transactions on Automatic Control, in Press", 2019

[76] A. MARTINELLI, A. RENZAGLIA, A. OLIVA.*Cooperative Visual-Inertial Sensor Fusion: Fundamental Equations and State Determination in Closed-Form*, in "Autonomous Robots, Conditionally accepted", 2019

[77] L. MATIGNON, S. D'ALU, O. SIMONIN.*Multi-robot human scene observation based on hybrid metric-topological mapping*, in "European Conference on Mobile Robotics", Paris, France, September 2017, https://hal.archives-ouvertes.fr/hal-01553064

[78] A. MOURIKIS, S. ROUMELIOTIS.*A multi-state constraint Kalman filter for vision-aided inertial navigation*, in "Int. Conf. on Rob. and Aut. (ICRA)", IEEE, 2007, p. 3565–3572

[79] A. NÈGRE, L. RUMMELHARD, C. LAUGIER.*Hybrid Sampling Bayesian Occupancy Filter*, in "IEEE Intelligent Vehicles Symposium (IV)", Dearborn, United States, June 2014, https://hal.inria.fr/hal-01011703

[80] A. PANDEY, R. ALAMI.*A framework towards a socially aware Mobile Robot motion in Human-Centered dynamic environment*, in "IEEE/RSJ International Conference on Intelligent Robots and Systems", oct. 2010, p. 5855 -5860

[81] M.-I. POPESCU, H. RIVANO, O. SIMONIN.*Multi-robot Patrolling in Wireless Sensor Networks using Bounded Cycle Coverage*, in "ICTAI 2016 28th International Conference on Tools with Artificial Intelligence", San Jose, United States, IEEE, November 2016, https://hal.archives-ouvertes.fr/hal-01357866

[82] A. RENZAGLIA, L. DOITSIDIS, A. MARTINELLI, E. KOSMATOPOULOS.*Multi-Robot 3D Coverage of Unknown Areas*, in "International Journal of Robotics Research", May 2012, vol. 31, n^{o} 6, p. 738-752 [*DOI :* 10.1177/0278364912439332], http://hal.inria.fr/hal-00751309

[83] L. RUMMELHARD, A. NÈGRE, C. LAUGIER.*Conditional Monte Carlo Dense Occupancy Tracker*, in "18th IEEE International Conference on Intelligent Transportation Systems", Las Palmas, Spain, September 2015, https://hal.inria.fr/hal-01205298

[84] L. N. RUMMELHARD, A. NÈGRE, A. N. PAIGWAR, C. N. LAUGIER.*Ground Estimation and Point Cloud Segmentation using SpatioTemporal Conditional Random Field*, in "IEEE Intelligent Vehicles Symposium (IV)", Redondo Beach, United States, June 2017, p. 1105 - 1110 [*DOI :* 10.1109/IVS.2017.7995861], https://hal.inria.fr/hal-01579095

[85] L. RUMMELHARD, A. NÈGRE, M. PERROLLAZ, C. LAUGIER.*Probabilistic Grid-based Collision Risk Prediction for Driving Application*, in "ISER", Marrakech/Essaouira, Morocco, June 2014, https://hal.inria.fr/hal-01011808

[86] J. SARAYDARYAN, F. JUMEL, O. SIMONIN.*Robots Delivering Services to Moving People : Individual vs. Group Patrolling Strategies*, in "The 2015 IEEE International Workshop on Advanced Robotics and its Social Impacts (ARSO 2015)", Lyon, France, IEEE, July 2015, https://hal.archives-ouvertes.fr/hal-01191457

[87] D. SIERRA GONZÁLEZ, J. S. DIBANGOYE, C. LAUGIER.*High-Speed Highway Scene Prediction Based on Driver Models Learned From Demonstrations*, in "Proceedings of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC 2016)", Rio de Janeiro, Brazil, November 2016 [*DOI :* 10.1109/ITSC.2016.7795546], https://hal.inria.fr/hal-01396047

[88] D. SIERRA GONZÁLEZ, V. ROMERO-CANO, J. STEEVE DIBANGOYE, C. LAUGIER.*Interaction-Aware Driver Maneuver Inference in Highways Using Realistic Driver Models*, in "Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC 2017)", Yokohama, Japan, October 2017, https://hal.inria.fr/hal-01589493

[89] O. SIMONIN, F. CHARPILLET, E. THIERRY.*Revisiting wavefront construction with collective agents: an approach to foraging*, in "Swarm Intelligence", June 2014, vol. 8, n^{o} 2, p. 113-138 [*DOI :* 10.1007/s11721-014-0093-3], https://hal.inria.fr/hal-00974068

[90] A. SPALANZANI, J. RIOS-MARTINEZ, C. LAUGIER, S. LEE.*Risk Based Navigation Decisions*, in "Handbook of Intelligent Vehicles", A. ESKANDARIAN (editor), Springer Verlag, February 2012, vol. 1, http://hal.inria.fr/hal-00743336

[91] S. THRUN, W. BURGARD, D. FOX.*Probabilistic Robotics*, The MIT Press, September 2005, 1

[92] G. D. TIPALDI, K. O. ARRAS.*I want my coffee hot! Learning to find people under spatio-temporal constraints*, in "ICRA", IEEE, 2011, p. 1217–1222

[93] M. TLIG, O. BUFFET, O. SIMONIN.*Stop-Free Strategies for Traffic Networks: Decentralized On-line Optimization*, in "ECAI 2014 - 21st European Conference on Artificial Intelligence, 18-22 August 2014, Prague, Czech Republic - Including Prestigious Applications of Intelligent Systems (PAIS 2014)", 2014, p. 1191–1196, http://dx.doi.org/10.3233/978-1-61499-419-0-1191

[94] P. VASISHTA, D. VAUFREYDAZ, A. SPALANZANI.*Natural Vision Based Method for Predicting Pedestrian Behaviour in Urban Environments*, in "IEEE 20th International Conference on Intelligent Transportation Systems", Yokohama, Japan, October 2017, https://hal.inria.fr/hal-01561029

[95] P. VASISHTA, D. VAUFREYDAZ, A. SPALANZANI.*Urban Pedestrian Behaviour Modelling using Natural Vision andPotential Fields*, in "9th Workshop on Planning, Perception and Navigation for Intelligent Vehicles at the IEEE International Conference on Intelligent Robots and Systems", Vancouver, Canada, September 2017, https://hal.inria.fr/hal-01578741

[96] D. VASQUEZ, Y. YU, S. KUMAR, C. LAUGIER.*An open framework for human-like autonomous driving using Inverse Reinforcement Learning*, in "IEEE Vehicle Power and Propulsion Conference", 2014

[97] T. VIDAL, T. G. CRAINIC, M. GENDREAU, C. PRINS.*A Hybrid Genetic Algorithm with Adaptive Diversity Management for a Large Class of Vehicle Routing Problems with Time-windows*, in "Comput. Oper. Res.", January 2013, vol. 40, n^{o} 1, p. 475–489

[98] N. VIGNARD, Ö. ERKENT, C. LAUGIER, C. WOLF.*Semantic Occupancy Grid Segmentation*, 2018, Patent PCT

[99] J. VON NEUMANN, O. MORGENSTERN.*Theory of games and economic behavior*, Princeton University Press, 2007

[100] J. VON NEUMANN, O. MORGENSTERN.*Theory of Games and Economic Behavior*, Princeton University Press, 1944

Project-Team CONVECS

Construction of verified concurrent systems

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

IN PARTNERSHIP WITH:
Institut polytechnique de Grenoble

Université Joseph Fourier (Grenoble)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Proofs and Verification

Table of contents

Project-Team CONVECS

Creation of the Team: 2012 January 01, updated into Project-Team: 2014 January 01

Keywords:

Computer Science and Digital Science:

A1.3. - Distributed Systems
A1.3.5. - Cloud
A1.3.6. - Fog, Edge
A2.1.1. - Semantics of programming languages
A2.1.7. - Distributed programming
A2.4.1. - Analysis
A2.4.2. - Model-checking
A2.5. - Software engineering
A2.5.1. - Software Architecture & Design
A2.5.4. - Software Maintenance & Evolution
A2.5.5. - Software testing
A7.1.1. - Distributed algorithms
A7.1.3. - Graph algorithms
A7.2. - Logic in Computer Science
A8.9. - Performance evaluation

Other Research Topics and Application Domains:

B6.1.1. - Software engineering
B6.3.2. - Network protocols
B6.4. - Internet of things
B6.6. - Embedded systems
B8.1. - Smart building/home

1. Team, Visitors, External Collaborators

Research Scientists

Radu Mateescu [Team leader, Inria, Senior Researcher, HDR]
Hubert Garavel [Inria, Senior Researcher]
Frédéric Lang [Inria, Researcher]
Wendelin Serwe [Inria, Researcher]

Faculty Member

Gwen Salaün [UGA, Professor, HDR]

External Collaborator

Viet Anh Nguyen [IRT Saint-Exupéry]

Technical Staff

Lian Apostol [Inria, until Nov. 2018]

PhD Students

Gianluca Barbon [UGA]
Lina Marsso [Inria]

Ajay Muroor Nadumane [Inria]
Umar Ozeer [Orange Labs]

Administrative Assistant
Myriam Etienne [Inria]

2. Overall Objectives

2.1. Overview

The CONVECS project-team addresses the rigorous design of concurrent asynchronous systems using formal methods and automated analysis. These systems comprise several activities that execute simultaneously and autonomously (i.e., without the assumption about the existence of a global clock), synchronize, and communicate to accomplish a common task. In computer science, asynchronous concurrency arises typically in hardware, software, and telecommunication systems, but also in parallel and distributed programs.

Asynchronous concurrency is becoming ubiquitous, from the micro-scale of embedded systems (asynchronous logic, networks-on-chip, GALS – *Globally Asynchronous, Locally Synchronous* systems, multi-core processors, etc.) to the macro-scale of grids and cloud computing. In the race for improved performance and lower power consumption, computer manufacturers are moving towards asynchrony. This increases the complexity of the design by introducing nondeterminism, thus requiring a rigorous methodology, based on formal methods assisted by analysis and verification tools.

There exist several approaches to formal verification, such as theorem proving, static analysis, and model checking, with various degrees of automation. When dealing with asynchronous systems involving complex data types, verification methods based on state space exploration (reachability analysis, model checking, equivalence checking, etc.) are today the most successful way to detect design errors that could not be found otherwise. However, these verification methods have several limitations: they are not easily accepted by industry engineers, they do not scale well while the complexity of designs is ever increasing, and they require considerable computing power (both storage capacity and execution speed). These are the challenges that CONVECS seeks to address.

To achieve significant impact in the design and analysis of concurrent asynchronous systems, several research topics must be addressed simultaneously. There is a need for user-friendly, intuitive, yet formal specification languages that will be attractive to designers and engineers. These languages should provide for both functional aspects (as needed by formal verification) and quantitative ones (to enable performance evaluation and architecture exploration). These languages and their associated tools should be smoothly integrated into large-scale design flows. Finally, verification tools should be able to exploit the parallel and distributed computing facilities that are now ubiquitous, from desktop to high-performance computers.

3. Research Program

3.1. New Formal Languages and their Concurrent Implementations

We aim at proposing and implementing new formal languages for the specification, implementation, and verification of concurrent systems. In order to provide a complete, coherent methodological framework, two research directions must be addressed:

- *Model-based specifications*: these are operational (i.e., constructive) descriptions of systems, usually expressed in terms of processes that execute concurrently, synchronize together and communicate. Process calculi are typical examples of model-based specification languages. The approach we promote is based on LOTOS NT (LNT for short), a formal specification language that incorporates most constructs stemming from classical programming languages, which eases its acceptance by students and industry engineers. LNT [5] is derived from the ISO standard E-LOTOS (2001), of which it represents the first successful implementation, based on a source-level translation from LNT to the former ISO standard LOTOS (1989). We are working both on the semantic foundations of LNT (enhancing the language with module interfaces and timed/probabilistic/stochastic features, compiling the m among n synchronization, etc.) and on the generation of efficient parallel and distributed code. Once equipped with these features, LNT will enable formally verified asynchronous concurrent designs to be implemented automatically.

- *Property-based specifications*: these are declarative (i.e., non-constructive) descriptions of systems, which express *what* a system should do rather than *how* the system should do it. Temporal logics and μ-calculi are typical examples of property-based specification languages. The natural models underlying value-passing specification languages, such as LNT, are Labeled Transition Systems (LTSs or simply *graphs*) in which the transitions between states are labeled by actions containing data values exchanged during handshake communications. In order to reason accurately about these LTSs, temporal logics involving data values are necessary. The approach we promote is based on MCL (*Model Checking Language*) [55], which extends the modal μ-calculus with data-handling primitives, fairness operators encoding generalized Büchi automata, and a functional-like language for describing complex transition sequences. We are working both on the semantic foundations of MCL (extending the language with new temporal and hybrid operators, translating these operators into lower-level formalisms, enhancing the type system, etc.) and also on improving the MCL on-the-fly model checking technology (devising new algorithms, enhancing ergonomy by detecting and reporting vacuity, etc.).

We address these two directions simultaneously, yet in a coherent manner, with a particular focus on applicable concurrent code generation and computer-aided verification.

3.2. Parallel and Distributed Verification

Exploiting large-scale high-performance computers is a promising way to augment the capabilities of formal verification. The underlying problems are far from trivial, making the correct design, implementation, fine-tuning, and benchmarking of parallel and distributed verification algorithms long-term and difficult activities. Sequential verification algorithms cannot be reused as such for this task: they are inherently complex, and their existing implementations reflect several years of optimizations and enhancements. To obtain good speedup and scalability, it is necessary to invent new parallel and distributed algorithms rather than to attempt a parallelization of existing sequential ones. We seek to achieve this objective by working along two directions:

- *Rigorous design:* Because of their high complexity, concurrent verification algorithms should themselves be subject to formal modeling and verification, as confirmed by recent trends in the certification of safety-critical applications. To facilitate the development of new parallel and distributed verification algorithms, we promote a rigorous approach based on formal methods and verification. Such algorithms will be first specified formally in LNT, then validated using existing model checking algorithms of the CADP toolbox. Second, parallel or distributed implementations of these algorithms will be generated automatically from the LNT specifications, enabling them to be experimented on large computing infrastructures, such as clusters and grids. As a side-effect, this “bootstrapping” approach would produce new verification tools that can later be used to self-verify their own design.
- *Performance optimization:* In devising parallel and distributed verification algorithms, particular care must be taken to optimize performance. These algorithms will face concurrency issues at several levels: grids of heterogeneous clusters (architecture-independence of data, dynamic load balancing), clusters of homogeneous machines connected by a network (message-passing communication, detection of stable states), and multi-core machines (shared-memory communication, thread synchronization). We will seek to exploit the results achieved in the parallel and distributed computing field to improve performance when using thousands of machines by reducing the number of connections and the messages exchanged between the cooperating processes carrying out the verification task. Another important issue is the generalization of existing LTS representations (explicit, implicit, distributed) in order to make them fully interoperable, such that compilers and verification tools can handle these models transparently.

3.3. Timed, Probabilistic, and Stochastic Extensions

Concurrent systems can be analyzed from a *qualitative* point of view, to check whether certain properties of interest (e.g., safety, liveness, fairness, etc.) are satisfied. This is the role of functional verification, which produces Boolean (yes/no) verdicts. However, it is often useful to analyze such systems from a *quantitative* point of view, to answer non-functional questions regarding performance over the long run, response time, throughput, latency, failure probability, etc. Such questions, which call for numerical (rather than binary) answers, are essential when studying the performance and dependability (e.g., availability, reliability, etc.) of complex systems.

Traditionally, qualitative and quantitative analyzes are performed separately, using different modeling languages and different software tools, often by distinct persons. Unifying these separate processes to form a seamless design flow with common modeling languages and analysis tools is therefore desirable, for both scientific and economic reasons. Technically, the existing modeling languages for concurrent systems need to be enriched with new features for describing quantitative aspects, such as probabilities, weights, and time. Such extensions have been well-studied and, for each of these directions, there exist various kinds of automata, e.g., discrete-time Markov chains for probabilities, weighted automata for weights, timed automata for hard real-time, continuous-time Markov chains for soft real-time with exponential distributions, etc. Nowadays, the next scientific challenge is to combine these individual extensions altogether to provide even more expressive models suitable for advanced applications.

Many such combinations have been proposed in the literature, and there is a large amount of models adding probabilities, weights, and/or time. However, an unfortunate consequence of this diversity is the confuse landscape of software tools supporting such models. Dozens of tools have been developed to implement theoretical ideas about probabilities, weights, and time in concurrent systems. Unfortunately, these tools do not interoperate smoothly, due both to incompatibilities in the underlying semantic models and to the lack of common exchange formats.

To address these issues, CONVECS follows two research directions:

- *Unifying the semantic models*. Firstly, we will perform a systematic survey of the existing semantic models in order to distinguish between their essential and non-essential characteristics, the goal being to propose a unified semantic model that is compatible with process calculi techniques for specifying and verifying concurrent systems. There are already proposals for unification either theoretical (e.g., Markov automata) or practical (e.g., PRISM and MODEST modeling languages), but these languages focus on quantitative aspects and do not provide high-level control structures and data handling features (as LNT does, for instance). Work is therefore needed to unify process calculi and quantitative models, still retaining the benefits of both worlds.
- *Increasing the interoperability of analysis tools*. Secondly, we will seek to enhance the interoperability of existing tools for timed, probabilistic, and stochastic systems. Based on scientific exchanges with developers of advanced tools for quantitative analysis, we plan to evolve the CADP toolbox as follows: extending its perimeter of functional verification with quantitative aspects; enabling deeper connections with external analysis components for probabilistic, stochastic, and timed models; and introducing architectural principles for the design and integration of future tools, our long-term goal being the construction of a European collaborative platform encompassing both functional and non-functional analyzes.

3.4. Component-Based Architectures for On-the-Fly Verification

On-the-fly verification fights against state explosion by enabling an incremental, demand-driven exploration of LTSs, thus avoiding their entire construction prior to verification. In this approach, LTS models are handled implicitly by means of their *post* function, which computes the transitions going out of given states and thus serves as a basis for any forward exploration algorithm. On-the-fly verification tools are complex software artifacts, which must be designed as modularly as possible to enhance their robustness, reduce their development effort, and facilitate their evolution. To achieve such a modular framework, we undertake research in several directions:

- *New interfaces for on-the-fly LTS manipulation.* The current application programming interface (API) for on-the-fly graph manipulation, named OPEN/CAESAR [41], provides an "opaque" representation of states and actions (transitions labels): states are represented as memory areas of fixed size and actions are character strings. Although appropriate to the pure process algebraic setting, this representation must be generalized to provide additional information supporting an efficient construction of advanced verification features, such as: handling of the types, functions, data values, and parallel structure of the source program under verification, independence of transitions in the LTS, quantitative (timed/probabilistic/stochastic) information, etc.
- *Compositional framework for on-the-fly LTS analysis.* On-the-fly model checkers and equivalence checkers usually perform several operations on graph models (LTSs, Boolean graphs, etc.), such as exploration, parallel composition, partial order reduction, encoding of model checking and equivalence checking in terms of Boolean equation systems, resolution and diagnostic generation for Boolean equation systems, etc. To facilitate the design, implementation, and usage of these functionalities, it is necessary to encapsulate them in software components that could be freely combined and replaced. Such components would act as graph transformers, that would execute (on a sequential machine) in a way similar to coroutines and to the composition of lazy functions in functional programming languages. Besides its obvious benefits in modularity, such a component-based architecture will also make it possible to take advantage of multi-core processors.
- *New generic components for on-the-fly verification.* The quest for new on-the-fly components for LTS analysis must be pursued, with the goal of obtaining a rich catalog of interoperable components serving as building blocks for new analysis features. A long-term goal of this approach is to provide an increasingly large catalog of interoperable components covering all verification and analysis functionalities that appear to be useful in practice. It is worth noticing that some components can be very complex pieces of software (e.g., the encapsulation of an on-the-fly model checker for a rich temporal logic). Ideally, it should be possible to build a novel verification or analysis tool by assembling on-the-fly graph manipulation components taken from the catalog. This would provide a flexible means of building new verification and analysis tools by reusing generic, interoperable model manipulation components.

3.5. Real-Life Applications and Case Studies

We believe that theoretical studies and tool developments must be confronted with significant case studies to assess their applicability and to identify new research directions. Therefore, we seek to apply our languages, models, and tools for specifying and verifying formally real-life applications, often in the context of industrial collaborations.

4. Application Domains

4.1. Application Domains

The theoretical framework we use (automata, process algebras, bisimulations, temporal logics, etc.) and the software tools we develop are general enough to fit the needs of many application domains. They are applicable to virtually any system or protocol that consists of distributed agents communicating by asynchronous messages. The list of recent case studies performed with the CADP toolbox (see in particular § 6.5) illustrates the diversity of applications:

- *Bioinformatics:* genetic regulatory networks, nutritional stress response, metabolic pathways,
- *Component-based systems:* Web services, peer-to-peer networks,
- *Cloud computing:* self-deployment protocols, dynamic reconfiguration protocols,
- *Fog and IoT:* stateful IoT applications in the fog,

- *Databases:* transaction protocols, distributed knowledge bases, stock management,
- *Distributed systems:* virtual shared memory, dynamic reconfiguration algorithms, fault tolerance algorithms, cloud computing,
- *Embedded systems:* air traffic control, avionic systems, medical devices,
- *Hardware architectures:* multiprocessor architectures, systems on chip, cache coherency protocols, hardware/software codesign,
- *Human-machine interaction:* graphical interfaces, biomedical data visualization, plasticity,
- *Security protocols:* authentication, electronic transactions, cryptographic key distribution,
- *Telecommunications:* high-speed networks, network management, mobile telephony, feature interaction detection.

5. New Software and Platforms

5.1. CADP Pro

Construction and Analysis of Distributed Processes
KEYWORDS: Formal methods - Verification
FUNCTIONAL DESCRIPTION: CADP (*Construction and Analysis of Distributed Processes* – formerly known as *CAESAR/ALDEBARAN Development Package*) [4] is a toolbox for protocols and distributed systems engineering.

In this toolbox, we develop and maintain the following tools:

- CAESAR.ADT [40] is a compiler that translates LOTOS abstract data types into C types and C functions. The translation involves pattern-matching compiling techniques and automatic recognition of usual types (integers, enumerations, tuples, etc.), which are implemented optimally.
- CAESAR [46], [45] is a compiler that translates LOTOS processes into either C code (for rapid prototyping and testing purposes) or finite graphs (for verification purposes). The translation is done using several intermediate steps, among which the construction of a Petri net extended with typed variables, data handling features, and atomic transitions.
- OPEN/CAESAR [41] is a generic software environment for developing tools that explore graphs on the fly (for instance, simulation, verification, and test generation tools). Such tools can be developed independently of any particular high level language. In this respect, OPEN/CAESAR plays a central role in CADP by connecting language-oriented tools with model-oriented tools. OPEN/CAESAR consists of a set of 16 code libraries with their programming interfaces, such as:
  - CAESAR_GRAPH, which provides the programming interface for graph exploration,
  - CAESAR_HASH, which contains several hash functions,
  - CAESAR_SOLVE, which resolves Boolean equation systems on the fly,
  - CAESAR_STACK, which implements stacks for depth-first search exploration, and
  - CAESAR_TABLE, which handles tables of states, transitions, labels, etc.

A number of on-the-fly analysis tools have been developed within the OPEN/CAESAR environment, among which:

  - BISIMULATOR, which checks bisimulation equivalences and preorders,
  - CUNCTATOR, which performs steady-state simulation of continuous-time Markov chains,
  - DETERMINATOR, which eliminates stochastic nondeterminism in normal, probabilistic, or stochastic systems,
  - DISTRIBUTOR, which generates the graph of reachable states using several machines,
  - EVALUATOR, which evaluates MCL formulas,
  - EXECUTOR, which performs random execution,
  - EXHIBITOR, which searches for execution sequences matching a given regular expression,
  - GENERATOR, which constructs the graph of reachable states,
  - PROJECTOR, which computes abstractions of communicating systems,
  - REDUCTOR, which constructs and minimizes the graph of reachable states modulo various equivalence relations,
  - SIMULATOR, XSIMULATOR, and OCIS, which enable interactive simulation, and
  - TERMINATOR, which searches for deadlock states.

- BCG (*Binary Coded Graphs*) is both a file format for storing very large graphs on disk (using efficient compression techniques) and a software environment for handling this format. BCG also plays a key role in CADP as many tools rely on this format for their inputs/outputs. The BCG environment consists of various libraries with their programming interfaces, and of several tools, such as:
  - BCG_CMP, which compares two graphs,
  - BCG_DRAW, which builds a two-dimensional view of a graph,
  - BCG_EDIT, which allows the graph layout produced by BCG_DRAW to be modified interactively,
  - BCG_GRAPH, which generates various forms of practically useful graphs,
  - BCG_INFO, which displays various statistical information about a graph,
  - BCG_IO, which performs conversions between BCG and many other graph formats,
  - BCG_LABELS, which hides and/or renames (using regular expressions) the transition labels of a graph,
  - BCG_MIN, which minimizes a graph modulo strong or branching equivalences (and can also deal with probabilistic and stochastic systems),
  - BCG_STEADY, which performs steady-state numerical analysis of (extended) continuous-time Markov chains,
  - BCG_TRANSIENT, which performs transient numerical analysis of (extended) continuous-time Markov chains, and
  - XTL (*eXecutable Temporal Language*), which is a high level, functional language for programming exploration algorithms on BCG graphs. XTL provides primitives to handle states, transitions, labels, *successor* and *predecessor* functions, etc.

For instance, one can define recursive functions on sets of states, which allow evaluation and diagnostic generation fixed point algorithms for usual temporal logics (such as HML [49], CTL [37], ACTL [38], etc.) to be defined in XTL.

- PBG (*Partitioned BCG Graph*) is a file format implementing the theoretical concept of *Partitioned LTS* [44] and providing a unified access to a graph partitioned in fragments distributed over a set of remote machines, possibly located in different countries. The PBG format is supported by several tools, such as:
  - PBG_CP, PBG_MV, and PBG_RM, which facilitate standard operations (copying, moving, and removing) on PBG files, maintaining consistency during these operations,
  - PBG_MERGE (formerly known as BCG_MERGE), which transforms a distributed graph into a monolithic one represented in BCG format,
  - PBG_INFO, which displays various statistical information about a distributed graph.
- The connection between explicit models (such as BCG graphs) and implicit models (explored on the fly) is ensured by OPEN/CAESAR-compliant compilers, e.g.:
  - BCG_OPEN, for models represented as BCG graphs,
  - CAESAR.OPEN, for models expressed as LOTOS descriptions,
  - EXP.OPEN, for models expressed as communicating automata,
  - FSP.OPEN, for models expressed as FSP [53] descriptions,
  - LNT.OPEN, for models expressed as LNT descriptions, and
  - SEQ.OPEN, for models represented as sets of execution traces.

The CADP toolbox also includes TGV (*Test Generation based on Verification*), which has been developed by the VERIMAG laboratory (Grenoble) and Inria Rennes – Bretagne-Atlantique.

The CADP tools are well-integrated and can be accessed easily using either the EUCALYPTUS graphical interface or the SVL [42] scripting language. Both EUCALYPTUS and SVL provide users with an easy and uniform access to the CADP tools by performing file format conversions automatically whenever needed and by supplying appropriate command-line options as the tools are invoked.

- Participants: Hubert Garavel, Frédéric Lang, Radu Mateescu and Wendelin Serwe
- Contact: Hubert Garavel
- URL: http://cadp.inria.fr/

5.2. TRAIAN

KEYWORDS: Compilation - LOTOS NT
FUNCTIONAL DESCRIPTION: TRAIAN is a compiler for translating LOTOS NT descriptions into C programs, which will be used for simulation, rapid prototyping, verification, and testing.

The current version of TRAIAN, which handles LOTOS NT types and functions only, has useful applications in compiler construction [43], being used in all recent compilers developed by CONVECS.

- Participants: Hubert Garavel, Frédéric Lang and Wendelin Serwe
- Contact: Hubert Garavel
- URL: http://convecs.inria.fr/software/traian/

6. New Results

6.1. New Formal Languages and their Implementations

6.1.1. LOTOS and LNT Specification Languages

Participants: Hubert Garavel, Frédéric Lang, Wendelin Serwe.

LNT [5] [36] is a next-generation formal description language for asynchronous concurrent systems. The design of LNT at CONVECS is the continuation of the efforts undertaken in the 80s to define sound languages for concurrency theory and, indeed, LNT is derived from the ISO standards LOTOS (1989) and E-LOTOS (2001). In a nutshell, LNT attempts to combine the best features of imperative programming languages, functional languages, and value-passing process calculi.

LNT is not a frozen language: its definition started in 2005, as part of an industrial project. Since 2010, LNT has been systematically used by CONVECS for numerous case studies (many of which being industrial applications — see § 6.5). LNT is also used as a back-end by other research teams who implement various languages by translation to LNT. It is taught in university courses, e.g., at University Grenoble Alpes and ENSIMAG, where it is positively accepted by students and industry engineers. Based on the feedback acquired by CONVECS, LNT is continuously improved.

In 2018, the CADP tools that translate LNT to LOTOS have been enhanced in various ways. In the warning and error messages emitted by LNT2LOTOS, line numbers have been made more precise. In addition to a bug fix, the LNT_DEPEND tool, which computes dependencies between LNT modules has been entirely rewritten and made much faster. Also, the LNT language has been simplified by removing "!external" pragmas for constructors, as "!external" pragmas for types are sufficient.

We also continued improving the TRAIAN compiler for the LOTOS NT language (a predecessor of LNT), which is used for the construction of most CADP compilers and translators.

In February 2018, we released version 2.9 of TRAIAN. We scrutinized the source code of TRAIAN, deleting all parts of code corresponding to those features of the LOTOS NT language that were either not fully implemented or seldom used in practice. This reduced the source code of TRAIAN by 40% and the binaries by 50%. External LOTOS NT functions are now allowed to return a non-void result. Support for 64-bit macOS executables was added. A few bugs have been fixed and the reference manual of TRAIAN was entirely revised.

The main limitation of TRAIAN 2.x is that it is a 20-year-old compiler that is increasingly difficult to maintain. It consists in a large collection of attribute grammars and is built using the FNC-2 compiler generation system, which is no longer supported. For this reason, TRAIAN only exists in 32-bit version, and sometimes hits the 3–4 GB RAM limit when dealing with large compiler specifications, such as those of LNT2LOTOS or EVALUATOR 5.

For this reason, we undertook a complete rewrite of TRAIAN to get rid of FNC-2. Two main design decisions behind TRAIAN 3.0 are the following: (i) it supports (most of) the LOTOS NT language currently accepted by TRAIAN 2.9, but also extensions belonging to LNT, so as to allow a future migration from LOTOS NT to LNT; and (ii) TRAIAN 3.0 is currently written in LOTOS NT and compiled using TRAIAN 2.9, but should be ultimately capable of bootstrapping itself.

So far, a lexer and parser for LOTOS NT have been developed using the SYNTAX compiler-generation system [0] developed at Inria Paris. This work triggered an in-depth reexamination of the programming interfaces offered by SYNTAX and led to enhancements of these interfaces (see § 6.1.6).

The abstract syntax tree of LOTOS NT, and the library of predefined LOTOS NT types and functions have been redesigned; previously specified as FNC-2 attribute grammars, they are now themselves written in LOTOS NT, so as to allow bootstrap, using the current version of TRAIAN to build the next one. The construction of the abstract syntax tree has also been completed. Finally, we set several non-regression test bases gathered all available programs written in LOTOS NT.

[0] http://syntax.gforge.inria.fr

6.1.2. NUPN

Participant: Hubert Garavel.

Nested-Unit Petri Nets (NUPNs) is an upward-compatible extension of P/T nets, which are enriched with structural information on their concurrent structure. Such additional information can easily be produced when NUPNs are generated from higher-level specifications (e.g., process calculi); quite often, such information allows logarithmic reductions in the number of bits required to represent states, thus enabling verification tools to perform better. The principles of NUPNs are exposed in [39] and its PNML representation is described here [0].

The NUPN model has been adopted by the Model Checking Contest and the Rigorous Examination of Reactive Systems challenge. It has been so far implemented in thirteen different tools developed in four countries.

In 2018, a journal article (to appear in 2019) has been written to formalize the complete theory of NUPNs. The CAESAR.BDD tool for NUPNs has been extended with twelve new options. A new tool named NUPN_INFO has been added to CADP to perform three normalizing transformations of NUPNs.

6.1.3. MCL and XTL Property Specification Languages

Participants: Hubert Garavel, Radu Mateescu.

CADP provides two different languages, named MCL and XTL, for expressing data-handling temporal properties of concurrent systems. MCL is an extension of alternation-free modal μ-calculus with data values, programming language constructs, generalized regular formulas on transition sequences, and fairness operators. XTL is a functional-like programming language interpreted on Labeled Transition Systems, enabling the definition of temporal operators by computing their interpretation using fixed point iterations over sets of states and transitions.

In 2018, we enhanced these languages and their associated tools as follows:

- The MCL v4 language was enhanced with a new operator "**loop**" on regular formulas over transition sequences. This general iteration operator parameterized by data variables is able to characterize complex (recursively definable) sequences in an LTS. Two auxiliary regular operators "**continue**" and "**exit**" carrying data values were also introduced to express the repetition and the termination of a loop regular formula, respectively. These operators are particularly useful for specifying transition sequences having a particular cumulated cost (e.g., number of transitions, sum of weights associated to actions, etc.) in the context of probabilistic verification (see § 6.3.2).
- The MCL v3 language was modified and aligned on MCL v4 by removing syntactic differences that existed between both languages concerning the infinite repetition operator ("@") and the respective precedences of the concatenation (".") and choice ("|") operators in regular formulas. MCL v3 has also been enriched with the option operator ("?") on regular formulas already present in MCL v4.
- Consequently, the two versions of MCL_EXPAND for MCL v3 and MCL v4 have been unified in one single tool, which is now invoked by both EVALUATOR 3 and EVALUATOR 4. The corresponding manual pages have been simplified accordingly, with the introduction of two overarching manual pages ("mcl" and "evaluator"). In addition to five bug fixes, the memory footprint of MCL_EXPAND has been reduced. The error messages displayed by MCL_EXPAND, EVALUATOR 3, and EVALUATOR 4 have been improved in terms of accuracy and explanatory contents.
- In addition to four bug fixes, the XTL model checker now performs consistency checks on the C identifiers specified by the pragmas "!implementedby", "!comparedby", "!enumeratedby", and "!printedby".
- Two new options were added to the EVALUATOR and XTL model checkers: "-depend", which displays the libraries transitively included in an MCL or XTL file, and "-source", which is used by SVL to display correct file names and line numbers for MCL or XTL formulas embedded in SVL scenarios.

[0] http://mcc.lip6.fr/nupn.php

6.1.4. Translation of Term Rewrite Systems

Participant: Hubert Garavel.

We pursued the development undertaken in 2015 of a software platform for systematically comparing the performance of rewrite engines and pattern-matching implementations in algebraic specification and functional programming languages. Our platform reuses the benchmarks of the three Rewrite Engine Competitions (2006, 2009, and 2010). Such benchmarks are term-rewrite systems expressed in a simple formalism named REC, for which we developed automated translators that convert REC benchmarks into many languages, among which AProVE, Clean, Haskell, LNT, LOTOS, Maude, mCRL, MLTON, OCAML, Opal, Rascal, Scala, SML-NJ, Stratego/XT, and Tom.

In 2018, we corrected and/or enhanced several of the existing REC translators and finalized experiments. The results of this study have been presented during an invited talk at WRLA'2018 (*12th International Workshop on Rewriting Logic and its Applications*) and an article [15] was published in the WRLA post-proceedings.

6.1.5. Formal Modeling and Analysis of BPMN

Participant: Gwen Salaün.

A business process is a set of structured activities that provide a certain service or product. Business processes can be modeled using the BPMN standard, and several industrial platforms have been developed for supporting their design, modeling, and simulation.

In collaboration with Francisco Durán and Camilo Rocha (University of Málaga, Spain), we proposed a rewriting logic executable specification of BPMN with time and extended with probabilities. Duration times and delays for tasks and flows can be specified as stochastic expressions, while probabilities are associated to various forms of branching behavior in gateways. These quantities enable discrete-event simulation and automatic stochastic verification of properties such as expected processing time, expected synchronization time at merge gateways, and domain-specific quantitative assertions. The mechanization of the stochastic analysis tasks is done with Maude's statistical model checker PVeStA. These results led to a publication in an international journal [10].

We also worked on an extension of BPMN with data, which is convenient for describing real-world processes involving complex behavior and data descriptions. By considering this level of expressiveness due to the new features, challenging questions arise regarding the choice of the semantic framework for specifying such an extension of BPMN, as well as how to carry out the symbolic simulation, validation, and assess the correctness of the process models. These issues were addressed first by providing a symbolic executable rewriting logic semantics of BPMN using the rewriting modulo SMT framework, where the execution is driven by rewriting modulo axioms and by querying SMT decision procedures for data conditions. Second, reachability properties, such as deadlock freedom and detection of unreachable states with data exhibiting certain values, can be specified and automatically checked with the help of Maude, thanks to its support for rewriting modulo SMT. These results led to a publication in an international conference [21].

6.1.6. Other Language Developments

Participants: Hubert Garavel, Frédéric Lang, Wendelin Serwe.

The ability to compile and verify formal specifications with complex, user-defined operations and data structures is a key feature of the CADP toolbox since its very origins.

In 2018, we enhanced the SYNTAX compiler generator [0] in various ways: (i) The "string manager" has been generalized to allow several symbol tables to be handled simultaneously; (ii) The "source manager" has been extended with new relocation primitives that enable the caller to specify alternative file names and line numbers for the source file being parsed; for instance, this is typically useful for implementing the "#line" pragma of the C preprocessor; this mechanism has been extended to transparently handle multiple relocations (triggered by the lexer) while recognizing the right-hand side of a syntax rule in the grammar; (iii) The "include manager"

[0] http://syntax.gforge.inria.fr

has been modified to store file names in a distinct symbol table than the table of identifiers, and to provide the list of all files transitively included from the principal module; (iv) Finally, the main programming interface of SYNTAX has been extended with new primitives, so that at present only 5 calls (rather than 9–13 calls, formerly) are required to launch a compiler written using SYNTAX.

All the CADP compilers have been modified to take advantage of the improvements of the SYNTAX library.

Also, a master student started to study an automated translation from Event-B to LNT. He reviewed the syntax and semantics of Event-B and proposed a pencil-paper translation of most Event-B operators. He applied it to a small example consisting of a bank system, where accounts can be created and closed, and money can be deposited or withdrawn. This was a preliminary work that did not lead to a full implementation, due to lack of time. However, this work is a solid basis for a later implementation.

6.2. Parallel and Distributed Verification

6.2.1. Distributed State Space Manipulation

Participant: Wendelin Serwe.

For distributed verification, CADP provides the PBG format, which implements the theoretical concept of *Partitioned LTS* [44] and provides a unified access to an LTS distributed over a set of remote machines.

In 2018, we improved the usability of distributed state space manipulation tools. In particular:

- A memory shortage error that occurs on a computing node now triggers a distributed termination of the computation, producing proper error messages in the log file of that node.
- A similar naming scheme for log files produced by computing nodes was enforced for all distributed verification tools, which prevents interferences between different invocations of the tools.

6.2.2. Debugging of Concurrent Systems using Counterexample Analysis

Participants: Gianluca Barbon, Gwen Salaün.

Model checking is an established technique for automatically verifying that a model satisfies a given temporal property. When the model violates the property, the model checker returns a counterexample, which is a sequence of actions leading to a state where the property is not satisfied. Understanding this counterexample for debugging the specification is a complicated task for several reasons: (i) the counterexample can contain hundreds of actions, (ii) the debugging task is mostly achieved manually, (iii) the counterexample does not explicitly highlight the source of the bug that is hidden in the model, (iv) the most relevant actions are not highlighted in the counterexample, and (v) the counterexample does not give a global view of the problem.

We proposed an approach that improves the usability of model checking by simplifying the comprehension of counterexamples. Our solution aims at keeping only actions in counterexamples that are relevant for debugging purposes. This is achieved by detecting in the models some specific choices between transitions leading to a correct behaviour or falling into an erroneous part of the model. These choices, which we call “neighbourhoods”, provide key information for understanding the bug behind the counterexample. To extract such choices, we proposed a first method for debugging the counterexamples of safety property violations. To do so, it builds a new model from the original one containing all the counterexamples, and then compares the two models to identify neighbourhoods.

In 2018, we proposed a different method for debugging the counterexamples of liveness property violations. Given a liveness property, it extends the model with prefix and suffix information w.r.t. that property. This enriched model is then analysed to identify neighbourhoods. A set of abstraction techniques we developed exploit the enriched model annotated with neighbourhoods to extract relevant actions from counterexamples, which makes their comprehension easier. This work led to a publication in an international conference [16].

Both approaches are fully automated by a tool we implemented and that has been validated on real-world case studies from various application areas. We extended the methodology and tool with 3D visualization techniques to visualize the erroneous part of the model with a specific focus on neighbourhoods, in order to have a global view of the bug behaviour. This work led to a publication to appear in an international conference.

A detailed description of the proposed methodology is available in G. Barbon's PhD thesis [8].

6.3. Timed, Probabilistic, and Stochastic Extensions

6.3.1. Tools for Probabilistic and Stochastic Systems

Participants: Hubert Garavel, Frédéric Lang.

Formal models and tools dealing with quantitative aspects (such as time, probabilities, and other continuous physical quantities) have become unavoidable for a proper study and computer-aided verification of functional and non-functional properties of cyber-physical systems. The wealth of such formal models is sometimes referred to as a quantitative "zoo" [48].

The CADP toolbox already implements some of these probabilistic/stochastic models, namely DTMCs and CTMCs (*Discrete-Time* and *Continuous-Time Markov Chains*), and IMCs (*Interactive Markov Chains*) [50]. Our long-term goal is to increase the capability and flexibility of the CADP tools, so as to support other quantitative models more easily.

In 2018, BCG_STEADY and BCG_TRANSIENT were enhanced along the following lines:

- They were extended to handle single-state Markov chains and to properly compute state solution vectors and transition throughputs on such models.
- Their command-line options were simplified and warnings are emitted when the input Markov chain contains no stochastic transition.
- A problem which caused correct Markov chains to be rejected was corrected. This problem was due to floating point conversion and rounding errors.
- A confusion between state numbers and matrix indices was fixed in the output and error messages.
- Models containing probabilistic self-loops are now rejected, as was already the case of longer circuits of probabilistic transitions, as both represent similar "timelock" situations.

6.3.2. On-the-fly Model Checking for Extended Regular Probabilistic Operators

Participant: Radu Mateescu.

Specifying and verifying quantitative properties of concurrent systems requires expressive and user-friendly property languages combining temporal, data-handling, and quantitative aspects. In collaboration with José Ignacio Requeno (Univ. Zaragoza, Spain), we undertook the quantitative analysis of concurrent systems modeled as PTSs (*Probabilistic Transition Systems*), whose actions contain data values and probabilities. We proposed a new regular probabilistic operator that extends naturally the Until operators of PCTL (*Probabilistic Computation Tree Logic*) [47], by specifying the probability measure of a path characterized by a generalized regular formula involving arbitrary computations on data values. We integrated the regular probabilistic operator into MCL, we devised an associated on-the-fly model checking method based on a combined local resolution of linear and Boolean equation systems, and we implemented the method in a prototype extension of the EVALUATOR model checker.

In 2018, we continued improving and using the extended model checker as follows:

- The model checker now determinizes the dataless regular formulas contained in regular probabilistic operators, ensuring automatically that the linear equation systems produced by the verification of these operators have a unique solution.
- For nondeterministic data-handling regular formulas contained in regular probabilistic operators, the model checker now produces a warning message informing the user that the determinization has to be done manually.
- We carried out further experiments to analyze the quantitative behaviour of the Bounded Retransmission Protocol, namely the variation of the probability of transmission failure w.r.t. the total number of retransmissions attempts.

A paper describing the probabilistic extension of MCL and of the on-the-fly model checker was published in an international journal [13].

6.4. Component-Based Architectures for On-the-Fly Verification

6.4.1. Compositional Verification

Participants: Hubert Garavel, Frédéric Lang.

The CADP toolbox contains various tools dedicated to compositional verification, among which EXP.OPEN, BCG_MIN, BCG_CMP, and SVL play a central role. EXP.OPEN explores on the fly the graph corresponding to a network of communicating automata (represented as a set of BCG files). BCG_MIN and BCG_CMP respectively minimize and compare behavior graphs modulo strong or branching bisimulation and their stochastic extensions. SVL (*Script Verification Language*) is both a high-level language for expressing complex verification scenarios and a compiler dedicated to this language.

In 2018, we improved these tools along the following lines:

- SVL now invokes EVALUATOR 3, EVALUATOR 4, and XTL with their new "-source" option, so that error and warning messages regarding temporal logic formulas now display line numbers in the SVL file itself, rather than in the temporary files generated to contain the temporal logic formulas, making it easier for users to modify incorrect MCL and XTL formulas contained in SVL files.
- SVL has been modified so that both EVALUATOR 3 and EVALUATOR 4 can now be used to compute "deadlock" and "livelock" statements.
- SVL does not require anymore that every "property" statement contains at least one verification statement, namely "comparison", "verify", "deadlock", "livelock", or a shell-line command with an "expected" clause.
- In addition to a bug fix, the EXP.OPEN tool was enhanced with a new option "-depend", displaying both the list of EXP files included (directly or transitively) in the input EXP file, and the list of automata, hide, rename, and cut files used (directly or transitively) in the input EXP file.

A paper containing both a tutorial and a survey on compositional verification was published in an international conference [14].

6.4.2. On-the-Fly Test Generation

Participants: Lina Marsso, Radu Mateescu, Wendelin Serwe.

The CADP toolbox provides support for conformance test case generation by means of the TGV tool. Given a formal specification of a system and a test purpose described as an input-output LTS (IOLTS), TGV automatically generates test cases, which assess using black box testing techniques the conformance of a system under test w.r.t. the formal specification. A test purpose describes the goal states to be reached by the test and enables one to indicate parts of the specification that should be ignored during the testing process. TGV does not generate test cases completely on the fly (i.e., *online*), because it first generates the complete test graph (CTG) and then traverses it backwards to produce controllable test cases.

To address these limitations, we developed the prototype tool TESTOR [0] to extract test cases completely on the fly. TESTOR presents several advantages w.r.t. TGV: (i) it has a more modular architecture, based on generic graph transformation components taken from the OPEN/CAESAR libraries (τ-compression, τ-confluence, τ-closure, determinization, resolution of Boolean equation systems); (ii) it is capable of extracting a test case completely on the fly, by exploiting the diagnostic generation features of the Boolean equation system resolution algorithms; (iii) it enables a more flexible expression of test purposes, taking advantage of the multiway rendezvous, a primitive to express communication and synchronization among a set of distributed processes.

[0] http://convecs.inria.fr/software/testor

In 2018, we improved TESTOR and TGV as follows:

- TESTOR has been ported to the Windows operating system.
- TESTOR can now be directly connected (by means of Unix pipes) to a system under test (SUT), executing the test case, rather than generating an abstract test-case that has to be connected to the SUT.
- We revised the architecture of TESTOR, so that the interface for the user is more similar to the one of TGV. This enables a user to easily switch between both tools.
- Taking advantage of the similar interfaces, we merged the non-regression test bases of TESTOR and TGV.
- We also fixed a bug and added a new option "-self" to TGV, reducing the number of warning messages.

These activities led to a new version 3.0 of TESTOR and two publications in international conferences [24], [18].

6.4.3. Other Component Developments

Participants: Pierre Bouvier, Hubert Garavel, Frédéric Lang, Radu Mateescu, Wendelin Serwe.

In 2018, several components of CADP have been improved as follows:

- The CADP toolbox now contains a new tool named SCRUTATOR for pruning Labeled Transition Systems on the fly.
- The OPEN/CAESAR environment was enriched with a new SOLVE_2 library for solving linear equation systems on the fly.
- Two manual pages ("bes" and "seq") have been added, which provide standalone definitions of CADP's BES format for Boolean Equation Systems and SEQ format for execution traces. The OPEN/CAESAR manual pages have been enhanced to give full prototypes for function parameters.
- The CADP toolbox has been ported to Solaris 11 and to SunOS 5.11 OpenIndiana "Hipster". CADP has also been ported to macOS 10.14 "Mojave" and a 64-bit version of CADP is now available for macOS.
- We also designed new C functions for handling path names in order to replace the traditional POSIX primitives basename(), dirname(), and realpath(), which suffer from limitations and ambiguities.

6.5. Real-Life Applications and Case Studies

6.5.1. Autonomous Resilience of Distributed IoT Applications in a Fog Environment

Participants: Umar Ozeer, Gwen Salaün.

Fog computing provides computing, storage and communication resources (and devices) at the edge of the network, near the physical world (PW). These end-devices nearing the physical world can have interesting properties such as short delays, responsiveness, optimized communications and privacy, which are especially appealing to IoT (Internet of Things) applications. However, IoT devices in the fog have low stability and are prone to failures.

In the framework of the collaboration with Orange Labs (see § 7.1.1), we are working on the key challenge of providing reliable services. This may be critical in this context since the non-containment of failures may impact the physical world. For instance, the failure of a smoke detector or a lamp in a smart home for elderly/medicated people may be hazardous. The design of such resilience solutions is complex due to the specificities of the environment, i.e., (i) dynamic infrastructure, where entities join and leave without synchronization; (ii) high heterogeneity in terms of functions, communication models, network, processing and storage capabilities; and (iii) cyber-physical interactions, which introduce non-deterministic and physical world's space and time dependent events.

In 2018, our work focused on proposing an end-to-end resilience approach for stateful IoT applications in the fog taking into account the three specificities mentioned above. The resilience protocol is functionally divided into four phases: (i) state-saving; (ii) monitoring and failure detection; (iii) failure notification and reconfiguration; and (iv) decision and recovery. The protocol implements a combination of different state-saving techniques based on rules and policies to cope with the heterogeneous nature of the environment and recover from failures in a consistent way, including PW-consistency. This work led to a publication in an international conference [25].

To illustrate our protocol at work, we mounted a smart home testbed with objects that can be found in real-life smart homes to test our solution. Our resilience approach was also implemented as a framework and deployed onto the testbed. The empirical results showed that multiple failures are recovered in an acceptable time in regard to end users. This work led to a publication to appear in an international conference.

6.5.2. Verified Composition and Deployment of IoT Applications

Participants: Radu Mateescu, Ajay Muroor Nadumane, Gwen Salaün.

The Internet of Things (IoT) is an interconnection of physical devices and software entities that can communicate and perform meaningful tasks largely without human intervention. The design and development of IoT applications is an interesting problem as these applications are typically dynamic, distributed and, more importantly, heterogeneous in nature.

In the framework of the collaboration with Nokia Bell Labs (see § 7.1.2), we proposed to build and deploy reliable IoT applications using a series of steps: (i) IoT objects and compositions are described using an interface-based behavioural model; (ii) the correctness of the composition is ensured by checking a behavioural compatibility notion that we proposed for IoT systems; and (iii) finally, a deployment plan respecting the dependencies between the objects is generated to facilitate automated deployment and execution of the application.

Regarding implementation, behavioural models and composition are specified in LNT and we take advantage of the CADP toolbox to perform compatibility checks. The deployment is automated using the Majord'Home platform developed by Nokia Bell Labs. The entire implementation is packaged as a Web tool available for end-users. This work led to a publication to appear in an international conference.

6.5.3. Memory Protection Unit

Participants: Hubert Garavel, Radu Mateescu, Wendelin Serwe.

Asynchronous circuits have key advantages in terms of low energy consumption, robustness, and security. However, the absence of a global clock makes the design prone to deadlock, livelock, synchronization, and resource-sharing errors. Formal verification is thus essential for designing such circuits, but it is not widespread enough, as many hardware designers are not familiar with it and few verification tools can cope with asynchrony on complex designs. In the framework of the SECURIOT-2 project (see § 8.2.2.1), we are interested in the rigorous design of asynchronous circuits used in the secure elements for IoT devices developed in the project.

In collaboration with Aymane Bouzafour and Marc Renaudin (Tiempo Secure), we suggested an extension of Tiempo's industrial design flow for asynchronous circuits, based upon the standard Hardware Description Language SystemVerilog (SV), with the formal verification capabilities provided by CADP. This was achieved by translating SV descriptions into LNT, expressing correctness properties in MCL, and verifying them using the EVALUATOR model checker of CADP. It turned out that the constructs of SV and LNT are in close correspondence, and that the synthesizable SV subset can be entirely translated into LNT. The MCL language was also shown adequate for expressing all property patterns relevant for asynchronous circuits.

The practicality of the approach was demonstrated on an asynchronous circuit (4000 lines of SV) implementing a memory protection unit (MPU). The MPU block exhibits a high degree of internal concurrency, comprising 660 parallel execution flows and 250 internal communication channels. The corresponding state space was generated compositionally, by identifying a suitable minimization and composition strategy described in

SVL (the largest intermediate state space had more than 116 million states and 862 million transitions). A set of 184 MCL properties were successfully verified on the state space, expressing the correct initialization of the MPU configuration registers, the mutual exclusion of read and write operations on registers, the correct responses to stimuli, and the security requirements related to the many access-control policies enforced by the MPU. This work led to a publication in an international conference [17].

6.5.4. TLS 1.3 Handshake Protocol

Participants: Lina Marsso, Radu Mateescu.

Security services are extensively used in fields like online banking, e-government, online shopping, etc. To ensure a secure communication between peers in terms of authenticity, privacy, and data integrity, cryptographic protocols are applied to regulate the data transfer. These protocols provide a standardized set of rules and methods for the interaction between peers. The Transport Layer Security (TLS) is a widely used security protocol, encompassing a set of rules for the communication between clients and servers, and relying on public-key cryptography to ensure integrity of exchanged data. However, despite multiple prevention measurements, several vulnerabilities (such as Heartbleed and DROWN), have been discovered recently. Therefore, testing the implementations of security protocols is still a crucial issue.

In the framework of the RIDINGS PHC project (see § 8.3.1), we are interested in testing protocols and distributed systems. In collaboration with Josip Bozic and Franz Wotawa (TU Graz, Austria), we undertook the formal modelling of the draft TLS 1.3 handshake protocol [0]. Taking as input the informal description of TLS 1.3 in the draft standard, we developed a formal model (1293 lines of LNT) specifying the handshake messages and client-server interactions. As far as we are aware, this is the first formal model of the draft TLS 1.3 handshake.

We used our LNT model for conformance testing with the OpenSSL version 1.0.1e implementation of the TLS protocol [0]. We defined three test purposes specifying requirements from the draft TLS 1.3 handshake, and applied the newly developed TESTOR tool (see § 6.4.2) to generate the test cases from the LNT model and each test purpose. The execution of these test cases on the OpenSSL implementation spotted a discrepancy of the server's response to a client certificate request w.r.t. the draft TLS 1.3 standard. This work led to a publication in an international workshop [18].

6.5.5. Message Authenticator Algorithm

Participants: Hubert Garavel, Lina Marsso.

The Message Authenticator Algorithm (MAA) is one of the first cryptographic functions for computing a Message Authentication Code. Between 1987 and 2001, the MAA was adopted in international standards (ISO 8730 and ISO 8731-2) to ensure the authenticity and integrity of banking transactions. The MAA also played a role in the history of formal methods, as the National Physical Laboratory (NPL, United Kingdom) developed, in the early 90s, three formal, yet non-executable, specifications of the MAA in VDM, Z, and LOTOS abstract data types.

In 2018, we examined how the new generation of formal methods can cope with the MAA case study. We specified the MAA in both LOTOS and LNT and checked these specifications using the CADP tools. The C code generated by the CADP compilers was executed w.r.t. a set of reference MAA test vectors, as well as supplementary test vectors devised to improve the coverage of byte permutations and message segmentation. This enabled us to detect and correct several errors in the reference test vectors given in the ISO 8730 and ISO 8731-2 standards. This work led to a publication in an international workshop [22].

6.5.6. Other Case Studies

Participants: Hubert Garavel, Frédéric Lang, Lina Marsso, Radu Mateescu, Wendelin Serwe.

Based on the work described above, the demo examples of the CADP toolbox have been enriched. Two new demo examples have been added: demo_06 (Transport Layer Security v1.3 handshake protocol specified in

[0] https://tools.ietf.org/html/draft-ietf-tls-tls13-24
[0] https://www.openssl.org/

LNT), and demo_11 (a hardware block implementing a Dynamic Task Dispatcher). The demo_12 (Message Authenticator Algorithm) is now documented in a publication [22]. The demo_17 (distributed leader election protocol) has been converted from LOTOS to LNT. Finally, most existing demo examples have been updated to reflect the evolution of the MCL v3 and SVL languages.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants with Industry

7.1.1. Orange Labs

Participants: Umar Ozeer, Gwen Salaün.

Umar Ozeer is supported by a PhD grant (from November 2016 to November 2019) from Orange Labs (Grenoble) on detecting and repairing failures of data-centric applications distributed in the cloud and the IoT (see § 6.5.1), under the supervision of Loïc Letondeur (Orange Labs), Gwen Salaün (CONVECS), François Gaël Ottogalli (Orange Labs), and Jean-Marc Vincent (POLARIS project-team).

7.1.2. Nokia Bell Labs

Participants: Radu Mateescu, Ajay Muroor Nadumane, Gwen Salaün.

Ajay Muroor Nadumane is supported by a PhD grant (from October 2017 to October 2020) from Nokia Bell Labs (Nozay) on IoT service composition supported by formal methods, under the supervision of Gwen Salaün (CONVECS), Radu Mateescu (CONVECS), Ludovic Noirie, and Michel Le Pallec (Nokia Bell Labs).

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. ARC6 Programme

Participants: Lina Marsso, Radu Mateescu [correspondent], Wendelin Serwe.

ARC6 is an academic research community funded by the Auvergne Rhône-Alpes region, whose objective is to foster the scientific collaborations between different academic institutions of the region working in the domain of information and communication technologies. ARC6 organizes various scientific animations (conferences, working groups, summer schools, etc.) and issues a yearly call for PhD and post-doctorate research project proposals.

Lina Marsso is supported by an ARC6 grant (from October 2016 to October 2019) on formal methods for testing networks of programmable logic controllers, under the supervision of Radu Mateescu and Wendelin Serwe (CONVECS), Ioannis Parissis and Christophe Deleuze (LCIS, Valence).

8.2. National Initiatives

8.2.1. PIA (Programme d'Investissements d'Avenir)

8.2.1.1. CAPHCA

Participants: Frédéric Lang, Radu Mateescu [correspondent], Wendelin Serwe.

CAPHCA (*Critical Applications on Predictable High-Performance Computing Architectures*) is a project funded by the PIA. The project, led by IRT Saint-Exupéry (Toulouse), involves a dozen of industrial partners (among which Airbus, CS Systèmes d'Information, Synopsis, and Thalès Avionics), the University Paul Sabatier (Toulouse), and Inria Grenoble – Rhône-Alpes (CONVECS and SPADES project-teams). CAPHCA addresses the dual problem of achieving performance and determinism when using new, high performance, multicore System-on-Chip (SoC) platforms for the deployment of real-time, safety-critical applications. The methodology adopted by CAPHCA consists in building a pragmatic combination of methods, tools, design constraints and patterns deployable at a short-term horizon in the industrial domains targeted in the project.

CAPHCA started in December 2017 for four years. The main contributions of CONVECS to CAPHCA are the detection of concurrency errors in parallel applications by means of formal methods and verification techniques.

8.2.2. Competitivity Clusters

8.2.2.1. SECURIOT-2

Participants: Lian Apostol, Hubert Garavel [correspondent], Radu Mateescu, Wendelin Serwe.

SECURIOT-2 is a project funded by the FUI (*Fonds Unique Interministériel*) within the *Pôle de Compétitivité* Minalogic. The project, led by Tiempo Secure (Grenoble), involves the SMEs (*Small and Medium Enterprises*) Alpwise, Archos, Sensing Labs, and Trusted Objects, the Institut Fourier and the VERIMAG laboratories of Université Grenoble Alpes, and CONVECS. SECURIOT-2 aims at developing a secure micro-controller unit (SMCU) that will bring to the IoT a high level of security, based on the techniques used for smart cards or electronic passports. The SMCU will also include an original power management scheme adequate with the low power consumption constraints of the IoT.

SECURIOT-2 started in September 2017 for three years. The main contributions of CONVECS to SECURIOT-2 are the formal modeling and verification of the asynchronous hardware implementing the secure elements developed by the project partners.

8.2.3. Other National Collaborations

We had sustained scientific relations with the following researchers:

- Xavier Etchevers (Orange Labs, Meylan),
- Fabrice Kordon and Lom Messan Hillah (LIP6, Paris),
- Eric Jenn and Viet Anh Nguyen (IRT Saint-Exupéry, Toulouse),
- Ioannis Parissis and Oum-El-Kheir Aktouf (LCIS, Valence),
- Pascal Poizat (LIP6, Paris).

8.3. European Initiatives

8.3.1. Collaborations in European Programs, Except FP7 & H2020

Program: PHC Amadeus
Project acronym: RIDINGS
Project title: Rigourous Development of GALS Systems
Duration: January 2017 – December 2018
Coordinator: Inria Grenoble – Rhône-Alpes / CONVECS
Other partners: TU Graz, Institute of Software Technology (Austria)
Abstract: GALS systems, composed of synchronous components (driven by local clocks) that communicate through a network, are increasingly spreading with the development of the IoT. GALS systems are intrinsically complex due to the interplay of synchronous and asynchronous aspects, which make their development and debugging difficult. Therefore, it is necessary to adopt rigorous design methodologies, based on formal methods assisted by efficient validation tools. The RIDINGS project aims at enhancing the design flow of a GALS system by integrating the automatic generation of conformance tests from the formal model and the temporal properties used for verifying the system. This yields a double benefit for the designer: (i) it makes possible to check that a physical implementation conforms to the verified model; (ii) the development cost of the model and properties is distributed on the verification and testing phases of the design process, therefore increasing the return on investment.

8.3.2. Collaborations with Major European Organizations

The CONVECS project-team is member of the FMICS (*Formal Methods for Industrial Critical Systems*) working group of ERCIM [0]. H. Garavel and R. Mateescu are members of the FMICS board, H. Garavel being in charge of dissemination actions.

[0] http://fmics.inria.fr

8.4. International Initiatives

H. Garavel is a member of IFIP (*International Federation for Information Processing*) Technical Committee 1 (*Foundations of Computer Science*) Working Group 1.8 on Concurrency Theory chaired successively by Luca Aceto and Jos Baeten.

8.4.1. Inria International Partners

8.4.1.1. Informal International Partners

Saarland University (Germany): we collaborate on a regular basis with the DEPEND (*Dependable Systems and Software*) research group headed by Holger Hermanns, who received an ERC Advanced Grant ("POWVER") in 2016.

8.4.2. Other International Collaborations

In 2018, we had scientific relations with several universities and institutes abroad, including:

- University of Málaga, Spain (Francisco Durán),
- University of Cali, Colombia (Camilo Rocha),
- University of Zaragoza, Spain (José Ignacio Requeno),
- ISTI/CNR, Pisa, Italy (Franco Mazzanti),
- RWTH Aachen, Germany (Joost-Pieter Katoen),
- Saarland University, Germany (Holger Hermanns),
- Eindhoven University of Technology, The Netherlands (Anton Wijs and Sander de Putter).

8.5. International Research Visitors

8.5.1. Visits of International Scientists

- H. Garavel is an invited professor at Saarland University (Germany) as a holder of the Gay-Lussac Humboldt Prize.
- Josip Bozic, Birgit Hofer, Hermann Felbinger, and Franz Wotawa (TU Graz, Austria) visited us from March 5 to March 9, 2018 in the framework of the RIDINGS PHC project (see § 8.3.1).
- G. Salaün visited the University of Málaga (Spain) from May 30 to June 13 and from December 16 to December 22, 2018.
- L. Marsso and R. Mateescu visited TU Graz (Austria) from August 20 to August 24, 2018 in the framework of the PHC RIDINGS project.

The annual CONVECS seminar was held in Dullin (France) on July 10–12, 2018. The following invited scientists attended the seminar:

- Eric Jenn (IRT Saint-Exupéry / Thales Avionics) gave on July 10, 2018 a talk entitled "*The CAPHCA Project, or How to Be Fast and Reasonable*".
- Viet Anh Nguyen (IRT Saint-Exupéry) gave on July 12, 2018 a talk entitled "*Cache-conscious Off-line Real-time Scheduling for Multi-core Platforms: Algorithms and Implementation*".
- Yliès Falcone (CORSE project-team) gave on July 11, 2018 a talk entitled "*Some Recent Work on the Runtime Monitoring of Systems*".

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Member of the Organizing Committees

- H. Garavel is a member of the model board [0] of MCC (*Model Checking Contest*). In 2018, he helped preparing new models (especially those in the NUPN format) and verified, using the CÆSAR.BDD tool of CADP, the forms describing all benchmark models submitted by the contest participants; this revealed a number of inconsistencies. The results of MCC'2018 have been published online [51].
- Together with Peter Höfner (Data61, CSIRO, Sydney, Australia), H. Garavel set up a model repository (hosted on the Gforge of Inria) to collect and archive formal models of real systems; this infrastructure is used by the series of MARS workshops [0]. This repository currently contains 21 models, one of which (a Transport Layer Security protocol) was deposited in 2018 by CONVECS.
- G. Salaün is member of the steering committee of the ACM SAC-SVT (*Symposium of Applied Computing – Software Verification and Testing Track*) conference series since 2018.
- G. Salaün is member of the steering committee of the SEFM (*International Conference on Software Engineering and Formal Methods*) conference series since 2014.
- G. Salaün is member of the steering committee of the FOCLASA (*International Workshop on Foundations of Coordination Languages and Self-Adaptative Systems*) workshop series since 2011.

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

- G. Salaün was co-chair of ACM SAC-SOAP'2018 (*33rd ACM Symposium of Applied Computing – Service-Oriented Architectures and Programming Track*), Pau, France, April 9–13, 2018.
- G. Salaün was workshops co-chair at STAF'2018 (*Software Technologies: Applications and Foundations*), Toulouse, France, June 25-29, 2018.
- W. Serwe was co-chair of MARS'2018 (*3rd Workshop on Models for Formal Analysis of Real Systems*) affiliated with ETAPS'2018 (*European Joint Conferences on Theory and Practice of Software*), Thessaloniki, Greece, April 20, 2018.

9.1.2.2. Member of the Conference Program Committees

- H. Garavel was program committee member of FVPS'2018 (*International Workshop on Formal Verification of Physical Systems*), Hagenberg, Austria, August 17, 2018.
- F. Lang was program committee member of SPIN'2018 (*25th International SPIN Symposium on Model Checking of Software*), Málaga, Spain, June 20–22, 2018.
- R. Mateescu was program committee member of ICTSS'2018 (*30th IFIP International Conference on Testing Software and Systems*), Cádiz, Spain, October 1–3, 2018.
- R. Mateescu was program committee member of FMICS'2018 (*23rd International Conference on Formal Methods for Industrial Critical Systems*), Maynooth, Ireland, September 3–4, 2018.
- G. Salaün was program committee member of CAL'2018 (*11ème Conférence francophone sur les Architectures Logicielles*), Grenoble, France, June 14–15, 2018.
- G. Salaün was program committee member of COMPSAC'2018 (*IEEE International Conference on Computers, Software, and Applications*), Tokyo, Japan, July 23–27, 2018.
- G. Salaün was program committee member of DATAMOD'2018 (*7th International Symposium "From Data to Models and Back"*), Toulouse, France, June 25–26, 2018.

[0] http://mcc.lip6.fr/models.php
[0] http://www.mars-workshop.org/

- G. Salaün was program committee member of FACS'2018 (*15th International Conference on Formal Aspects of Component Software*), Pohang, Korea, October 10–12, 2018.
- G. Salaün was program committee member of FOCLASA'2018 (*16th International Workshop on Foundations of Coordination Languages and Self-Adaptative Systems*), Toulouse, France, June 26, 2018.
- G. Salaün was program committee member of HPCS-4PAD'2018 (*5th International Symposium on Formal Approaches to Parallel and Distributed Systems*), Orléans, France, July 16–20, 2018.
- G. Salaün was program committee member of SEFM'2018 (*16th International Conference on Software Engineering and Formal Methods*), Toulouse, France, June 27–29, 2018.
- G. Salaün was program committee member of SAC-SVT'2018 (*33rd ACM Symposium on Applied Computing - Software Verification and Testing Track*), Pau, France, April 9–13, 2018.

9.1.2.3. Reviewer

- G. Barbon was a reviewer for COMPSAC'2018, DATAMOD'2018, and SEFM'2018.
- F. Lang was a reviewer for FoSSaCS'2018 (*21st International Conference on Foundations of Software Science and Computation Structures*).
- A. Muroor Nadumane was a reviewer for COMPSAC'2018, FACS'2018, FOCLASA'2018, SAC-SVT'2018, and SEFM'2018.
- U. Ozeer was a reviewer for COMPSAC'2018.
- R. Mateescu and W. Serwe were reviewers for the Festschrift in honor of Bernhard Steffen.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- H. Garavel is an editorial board member of STTT (*Springer International Journal on Software Tools for Technology Transfer*).

9.1.3.2. Reviewer - Reviewing Activities

- F. Lang was a reviewer for FAoC (*Formal Aspects of Computing*) and ToCL (*ACM Transactions on Computational Logic*).
- R. Mateescu was a reviewer for STTT, ToCL, and TOR (*IEEE Transactions on Reliability*).
- A. Muroor Nadumane was a reviewer for STTT.
- G. Salaün was a reviewer for FAoC, JCC (*Journal of Computer and Communications*), JLAMP (*Journal of Logical and Algebraic Methods in Programming*), SCP (*Science of Computer Programming*), TSE (*IEEE Transactions on Software Engineering*), TSI (*Technique et Science Informatiques*).

9.1.4. Software Dissemination and Internet Visibility

The CONVECS project-team distributes several software tools, among which the CADP toolbox.

In 2018, the main facts are the following:

- We prepared and distributed twelve successive versions (2018-a to 2018-l) of CADP.
- We were requested to grant CADP licenses for 381 different computers in the world.

The CONVECS Web site [0] was updated with scientific contents, announcements, publications, etc.

By the end of December 2018, the CADP forum [0], opened in 2007 for discussions regarding the CADP toolbox, had over 426 registered users and over 1847 messages had been exchanged.

Also, for the 2018 edition of the Model Checking Contest, 4 families of models generated using CADP (totalling 101 Nested-Unit Petri Nets) were provided.

[0] http://convecs.inria.fr
[0] http://cadp.inria.fr/forum.html

We contributed to Wikipedia as follows:

- We created a new page about the Message Authenticator Algorithm (https://en.wikipedia.org/wiki/Message_Authenticator_Algorithm)
- We added the three last paragraphs of the section about constructing integer numbers (https://en.wikipedia.org/wiki/Integer#Construction)

Other research teams took advantage of the software components provided by CADP (e.g., the BCG and OPEN/CAESAR environments) to build their own research software. We can mention the following developments:

- The RichTest Tool for Message-Passing Concurrent Programs [33]
- The REFINER Tool for Verifying Behavioural Model-to-Model Transformations [63]
- The ALVIS Tool for Modelling and Verification of Real-Time Systems [60]
- The COSTO Tool for Component-Based Software [30]
- The IDCM Tool for Analyzing UML Architectures [52]
- The OCARINA Tool and its Extension AADL2LNT for Analysing AADL Descriptions [58]
- The aZiZa Tool for Heterogeneous Behavioural Models [31]
- The Papyrus-RT Tool for Model-driven Engineering with UML-RT [62]
- Formal Analysis of Distributed Reactive Applications [35], [34]

Other teams also used the CADP toolbox for various case studies:

- Formal Modelling and Verification of an Automatic Train Supervision System [56], [57]
- Verification of Highly-Optimized Concurrent Data Structures [61]
- Detection of Data Breaches in Banking Transaction Processes [54]
- Verification of Visibility-Based Properties on Multiple Moving Robots [59]
- Experimental Analysis of Compositional State Space Generation Strategies [64]
- Product-Line for Families of Program Translators [32]

9.1.5. Invited Talks

- H. Garavel participated in the workshop "*Safety of Future Systems: Science meets Industry*" organized by the Lorentz Center (Leiden, The Netherlands) on April 9–13, 2018. He gave a lecture entitled "*Concurrency Theory Meets IoT*".
- H. Garavel gave an invited talk entitled "*Benchmarking Implementations of Term Rewriting and Pattern Matching in Algebraic, Functional, and Object-Oriented Languages - The 4th Rewrite Engines Competition*" at WRLA'2018 (*12th International Workshop on Rewriting Logic and its Applications*), Thessaloniki, Greece, April 14–15, 2018.
- L. Marsso gave a talk and presented a poster entitled "*Automated Test Generation for GALS Systems*" on March 8, 2018 at the 2nd year PhD LIG Day.
- L. Marsso gave a talk entitled "*Generation with CADP of Relevant Scenarios for Testing Autonomous Cars*" at the seminar of the group TransForm (*Méthodes formelles pour les systèmes de transport*) held in Villeneuve d'Ascq on November 22, 2018.
- L. Marsso gave a talk entitled "*Automated Test Generation for GALS Systems*" at the Scientific day of ARC 6 held in Lyon on November 29, 2018.
- R. Mateescu participated to the Kobe-Grenoble workshop organized by UGA in Grenoble on February 26–27, 2018. He gave a talk entitled "*Rigorous Design of PLC Networks using Formal Methods*" on February 26.
- A. Muroor Nadumane gave a talk entitled "*Building Reliable IoT Application and Beyond*" at the Inria-Nokia Bell Labs meeting held in Paris on November 27, 2018.

- U. Ozeer gave a talk and presented a poster entitled "*Autonomous Resilience of Distributed IoT Applications in a Fog Environment*" on March 8, 2018 at the 2nd year PhD LIG Day.
- U. Ozeer gave a talk entitled "*Autonomous Resilience of Distributed IoT Applications in a Fog Environment*" at the seminar on IoT research projects held at Orange Labs, Meylan, on March 29–30, 2018.
- U. Ozeer presented a poster entitled "*Autonomous Resilience of Distributed IoT Applications in a Fog Environment*" at the LIG seminar "*Regards sur le futur de l'informatique*" held in Grenoble on April 6, 2018.
- U. Ozeer gave a talk entitled "*Resilience of Distributed IoT Applications in a Dynamic Fog Environment*" at the IO Labs seminar held in Paris on October 30–31, 2018.
- G. Salaün gave a keynote talk entitled "*Safe Composition of Software Services*" at DATAMOD'2018, Toulouse, France, on June 26, 2018.

9.1.6. Research Administration

- H. Garavel was appointed to the Executive Commission in charge of International Relations at COMUE Université Grenoble Alpes.
- F. Lang is chair of the "*Commission du développement technologique*", which is in charge of selecting R&D projects for Inria Grenoble – Rhône-Alpes.
- R. Mateescu is the scientific correspondent of the European and International Partnerships for Inria Grenoble – Rhône-Alpes.
- R. Mateescu is a member of the *Comité d'orientation scientifique* for Inria Grenoble – Rhône-Alpes.
- R. Mateescu is a member of the "*Bureau*" of the LIG laboratory.
- G. Salaün is a member of the Scientific Committee of the PCS (*Pervasive Computing Systems*) action of the PERSYVAL Labex.
- W. Serwe is (together with Laurent Lefèvre from the AVALON Inria project-team) correspondent in charge of the 2017 Inria activity reports at Inria Grenoble – Rhône-Alpes.
- W. Serwe is a member of the "*Comité de Centre*" at Inria Grenoble – Rhone-Alpes.
- W. Serwe is "*chargé de mission*" for the scientific axis *Formal Methods, Models, and Languages* of the LIG laboratory.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

CONVECS is a host team for the computer science master entitled "*Mathématiques, Informatique, spécialité : Systèmes et Logiciels*", common to Grenoble INP and Université Grenoble Alpes (UGA).

In 2018, we carried out the following teaching activities:

G. Barbon and W. Serwe supervised each a group of six teams in the context of the "*projet Génie Logiciel*" (55 hours "*équivalent TD*", consisting in 16 hours of lectures, plus supervision and evaluation), ENSIMAG, January 2018.

F. Lang and R. Mateescu gave a lecture on "*Modeling and Analysis of Concurrent Systems: Models and Languages for Model Checking*" (27 hours "*équivalent TD*") to third year students of ENSIMAG and second year students of the MOSIG (*Master of Science in Informatics at Grenoble*).

F. Lang gave a course on "*Formal Software Development Methods*" (7.5 hours "*équivalent TD*") in the framework of the "*Software Engineering*" lecture given to first year students of the MOSIG.

F. Lang and W. Serwe provided a 6-hour training about the CADP toolbox to Eric Jenn, Nicolas Hili, and Sun Wei Tsun (IRT Saint-Exupéry, Toulouse, France) on June 28, 2018.

L. Marsso gave a course on “*Algorithms and Web Programming*” (64 hours “*équivalent TD*”) at the department MMI of IUT1 (UGA).

A. Muroor Nadumane gave a course on “*Object Oriented Programming*” (42 hours “*équivalent TD*”) at the department MMI of IUT1 (UGA).

G. Salaün taught about 230 hours of classes (algorithmics, Web development, object-oriented programming, iOS programming) at the department MMI of IUT1 (UGA). He is also headmaster of the “*Services Mobiles et Interface Nomade*” (SMIN) professional licence (3rd year of university) at IUT1/UGA.

9.2.2. Supervision

PhD in progress: L. Marsso, “*Formal Methods for Testing Networks of Controllers*, Université Grenoble Alpes, since October 2016, R. Mateescu, W. Serwe, I. Parissis, and Ch. Deleuze

PhD in progress: A. Muroor Nadumane, “*Softwarization of Everything: IoT Service Composition*, Université Grenoble Alpes, since October 2017, G. Salaün, R. Mateescu, L. Noirie, and M. Le Pallec

PhD in progress: U. Ozeer, “*Autonomous Resilience of Applications in a Largely Distributed Cloud Environment*, Université Grenoble Alpes, since November 2016, L. Letondeur, G. Salaün, F.-G. Ottogalli, and J.-M. Vincent

9.2.3. Juries

- G. Salaün was PhD committee member for Gustavo García Pascual’s PhD thesis, entitled “*Optimizing Mobile Applications by Exploiting Variability Models at Runtime*”, defended at University of Málaga (Spain) on December 18, 2018.

10. Bibliography

Major publications by the team in recent years

[1] X. ETCHEVERS, G. SALAÜN, F. BOYER, T. COUPAYE, N. DE PALMA.*Reliable Self-deployment of Distributed Cloud Applications*, in "Software: Practice and Experience", 2017, vol. 47, n^o 1, p. 3-20 [*DOI :* 10.1002/SPE.2400], https://hal.inria.fr/hal-01290465

[2] H. EVRARD, F. LANG.*Automatic Distributed Code Generation from Formal Models of Asynchronous Processes Interacting by Multiway Rendezvous*, in "Journal of Logical and Algebraic Methods in Programming", March 2017, vol. 88, 33 [*DOI :* 10.1016/J.JLAMP.2016.09.002], https://hal.inria.fr/hal-01412911

[3] H. GARAVEL, F. LANG, R. MATEESCU.*Compositional Verification of Asynchronous Concurrent Systems using CADP*, in "Acta Informatica", June 2015, vol. 52, n^o 4, 56 [*DOI :* 10.1007/S00236-015-0226-1], https://hal.inria.fr/hal-01247507

[4] H. GARAVEL, F. LANG, R. MATEESCU, W. SERWE.*CADP 2011: A Toolbox for the Construction and Analysis of Distributed Processes*, in "International Journal on Software Tools for Technology Transfer", 2013, vol. 15, n^o 2, p. 89-107 [*DOI :* 10.1007/S10009-012-0244-Z], http://hal.inria.fr/hal-00715056

[5] H. GARAVEL, F. LANG, W. SERWE.*From LOTOS to LNT*, in "ModelEd, TestEd, TrustEd - Essays Dedicated to Ed Brinksma on the Occasion of His 60th Birthday", J.-P. KATOEN, R. LANGERAK, A. RENSINK (editors), Lecture Notes in Computer Science, Springer, October 2017, vol. 10500, p. 3 - 26 [*DOI :* 10.1007/978-3-319-68270-9_1], https://hal.inria.fr/hal-01621670

[6] F. JEBALI, F. LANG, R. MATEESCU.*Formal Modelling and Verification of GALS Systems Using GRL and CADP*, in "Formal Aspects of Computing", April 2016, vol. 28, n^{o} 5, p. 767–804 [*DOI :* 10.1007/S00165-016-0373-3], https://hal.inria.fr/hal-01290449

[7] R. MATEESCU, W. SERWE.*Model Checking and Performance Evaluation with CADP Illustrated on Shared-Memory Mutual Exclusion Protocols*, in "Science of Computer Programming", February 2012 [*DOI :* 10.1016/J.SCICO.2012.01.003], http://hal.inria.fr/hal-00671321

Publications of the year

Doctoral Dissertations and Habilitation Theses

[8] G. BARBON.*Debugging of Behavioural Models using Counterexample Analysis*, Université Grenoble Alpes, December 2018

Articles in International Peer-Reviewed Journal

[9] L. AKROUN, G. SALAÜN.*Automated verification of automata communicating via FIFO and bag buffers*, in "Formal Methods in System Design", June 2018, vol. 52, n^{o} 3, p. 260 - 276 [*DOI :* 10.1007/S10703-017-0285-8], https://hal.inria.fr/hal-01898159

[10] F. DURÁN, C. ROCHA, G. SALAÜN.*Stochastic Analysis of BPMN with Time in Rewriting Logic*, in "Science of Computer Programming", December 2018, vol. 168, p. 1-17 [*DOI :* 10.1016/J.SCICO.2018.08.007], https://hal.inria.fr/hal-01866289

[11] F. KORDON, H. GARAVEL, L. HILLAH, E. PAVIOT-ADET, L. JEZEQUEL, F. HULIN-HUBARD, E. AMPARORE, M. BECCUTI, B. BERTHOMIEU, H. EVRARD, P. G. JENSEN, D. LE BOTLAN, T. LIEBKE, J. MEIJER, J. SRBA, Y. THIERRY-MIEG, J. VAN DE POL, K. WOLF.*MCC'2017-The Seventh Model Checking Contest*, in "LNCS Transactions on Petri Nets and Other Models of Concurrency", 2018, vol. 11090, p. 181-209 [*DOI :* 10.1007/978-3-662-58381-4_9], https://hal.inria.fr/hal-01917492

[12] A. KRISHNA, P. POIZAT, G. SALAÜN.*Checking Business Process Evolution*, in "Science of Computer Programming", January 2019, vol. 170, p. 1-26 [*DOI :* 10.1016/J.SCICO.2018.09.007], https://hal.inria.fr/hal-01920273

[13] R. MATEESCU, J. I. REQUENO.*On-the-Fly Model Checking for Extended Action-Based Probabilistic Operators*, in "International Journal on Software Tools for Technology Transfer", October 2018, vol. 20, n^{o} 5, p. 563–587 [*DOI :* 10.1007/S10009-018-0499-0], https://hal.inria.fr/hal-01862754

Invited Conferences

[14] H. GARAVEL, F. LANG, L. MOUNIER.*Compositional Verification in Action*, in "FMICS 2018 - 23rd International Conference on Formal Methods for Industrial Critical Systems", Maynooth, Ireland, LNCS, Springer, September 2018, vol. 11119, p. 189-210 [*DOI :* 10.1007/978-3-030-00244-2_13], https://hal.inria.fr/hal-01890246

[15] H. GARAVEL, M.-A. TABIKH, I.-S. ARRADA.*Benchmarking Implementations of Term Rewriting and Pattern Matching in Algebraic, Functional, and Object-Oriented Languages - The 4th Rewrite Engines Competition*, in "Proceedings of the 12th International Workshop on Rewriting Logic and its Applications (WRLA'18)", Thessaloniki, Greece, 2018, https://hal.inria.fr/hal-01883212

International Conferences with Proceedings

[16] G. BARBON, V. LEROY, G. SALAÜN.*Counterexample Simplification for Liveness Property Violation*, in "SEFM 2018 - 16th International Conference on Software Engineering and Formal Methods", Toulouse, France, LNCS, June 2018, vol. 10886, p. 173-188 [*DOI :* 10.1007/978-3-319-92970-5_11], https://hal.inria.fr/hal-01818790

[17] A. BOUZAFOUR, M. RENAUDIN, H. GARAVEL, R. MATEESCU, W. SERWE.*Model-checking Synthesizable SystemVerilog Descriptions of Asynchronous Circuits*, in "ASYNC'18 - 24th IEEE International Symposium on Asynchronous Circuits and Systems", Vienne, Austria, May 2018, https://hal.inria.fr/hal-01777093

[18] J. BOZIC, L. MARSSO, R. MATEESCU, F. WOTAWA.*A Formal TLS Handshake Model in LNT*, in "MARS/VPT 2018 - 3nd Workshop on Models for Formal Analysis of Real Systems and 6th International Workshop on Verification and Program Transformation", Thessaloniki, Greece, April 2018, vol. 268, p. 1 - 40 [*DOI :* 10.4204/EPTCS.268.1], https://hal.inria.fr/hal-01779151

[19] M. CORTES-CORNAX, A. KRISHNA, A. MOS, G. SALAÜN.*Automated Analysis of Industrial Workflow-based Models*, in "SAC 2018 - 33rd Annual ACM Symposium on Applied Computing", Pau, France, Proceedings of the 33rd Annual ACM Symposium on Applied Computing, SAC 2018, Pau, France, April 09-13, 2018, ACM, April 2018, p. 120-127 [*DOI :* 10.1145/3167132.3167142], https://hal.inria.fr/hal-01781315

[20] F. DURÁN, C. ROCHA, G. SALAÜN.*Computing the Parallelism Degree of Timed BPMN Processes*, in "FOCLASA 2018 - 16th International Workshop on Foundations of Coordination Languages and Self-Adaptative Systems", Toulouse, France, June 2018, p. 1-16 [*DOI :* 10.1007/978-3-030-04771-9_24], https://hal.inria.fr/hal-01961952

[21] F. DURÁN, C. ROCHA, G. SALAÜN.*Symbolic Specification and Verification of Data-aware BPMN Processes using Rewriting Modulo SMT*, in "WRLA 2018 : 12th International Workshop on Rewriting Logic and its Applications", Thessaloniki, Greece, April 2018, p. 1-20, https://hal.inria.fr/hal-01866268

[22] H. GARAVEL, L. MARSSO.*Comparative Study of Eight Formal Specifications of the Message Authenticator Algorithm*, in "MARS/VPT 2018 - 3nd Workshop on Models for Formal Analysis of Real Systems and the 6th International Workshop on Verification and Program Transformation", Thessaloniki, Greece, April 2018, vol. 268, p. 41 - 87 [*DOI :* 10.4204/EPTCS.268.2], https://hal.inria.fr/hal-01775332

[23] B. HOFER, R. MATEESCU, W. SERWE, F. WOTAWA.*Using LNT Formal Descriptions for Model-Based Diagnosis*, in "DX 2018 - 29th International Workshop on Principles of Diagnosis", Warsaw, Poland, August 2018, p. 1-8, https://hal.inria.fr/hal-01877693

[24] L. MARSSO, R. MATEESCU, W. SERWE.*TESTOR: A Modular Tool for On-the-Fly Conformance Test Case Generation*, in "TACAS 2018 - 24th International Conference on Tools and Algorithms for the Construction and Analysis of Systems", Thessaloniki, Greece, Lecture Notes in Computer Science, Springer, April 2018, vol. 10806, p. 211-228 [*DOI :* 10.1007/978-3-319-89963-3_13], https://hal.inria.fr/hal-01777861

[25] U. OZEER, X. ETCHEVERS, L. LETONDEUR, F.-G. OTTOGALLI, G. SALAÜN, J.-M. VINCENT.*Resilience of Stateful IoT Applications in a Dynamic Fog Environment*, in "EAI International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous '18)", New York, United States, November 2018, p. 1-10 [*DOI :* 10.1145/3286978.3287007], https://hal.archives-ouvertes.fr/hal-01927286

Conferences without Proceedings

[26] U. OZEER, L. LETONDEUR, F.-G. OTTOGALLI, G. SALAÜN, J.-M. VINCENT.*Designing and Implementing Resilient IoT Applications in the Fog: A Smart Home Use Case*, in "ICIN 2019 2019 - 22nd Conference on Innovation in Clouds, Internet and Networks", Paris, France, February 2019, p. 1-3, https://hal.archives-ouvertes.fr/hal-01979686

Books or Proceedings Editing

[27] C. CANAL, G. SALAÜN (editors). *Preface: Special issue on Foundations of Coordination Languages and Self-adaptive Systems*, Elsevier, December 2018, vol. 168, p. 169 - 170 [*DOI :* 10.1016/J.SCICO.2018.09.003], https://hal.inria.fr/hal-01898179

[28] R. MATEESCU (editor). *Recent advances in interactive and automated analysis*, Springer Verlag, April 2018, vol. 20, n° 2, p. 119 - 123 [*DOI :* 10.1007/S10009-017-0477-Y], https://hal.inria.fr/hal-01766570

[29] M. MAZZARA, I. OBER, G. SALAÜN (editors). *Software Technologies: Applications and Foundations (STAF 2018)*, 2018 [*DOI :* 10.1007/978-3-030-04771-9], https://hal.inria.fr/hal-01961961

References in notes

[30] P. ANDRÉ, O. CARDIN.*Trusted Services for Cyber Manufacturing Systems*, in "Proceedings of the 7th Workshop on Service Orientation in Holonic and Multi-agent Manufacturing (SOHOMA'2017), Nantes, France", T. BORANGIU, D. TRENTESAUX, A. THOMAS, O. CARDIN (editors), Springer International Publishing, October 2018, p. 359–370

[31] C. ATTIOGBÉ.*Mastering Heterogeneous Behavioural Models*, in "Proceedings of the 7th International Conference on Model and Data Engineering (MEDI'2017), Barcelona, Spain", Y. OUHAMMOU, M. IVANOVIC, A. ABELLÓ, L. BELLATRECHE (editors), Lecture Notes in Computer Science, Springer Verlag, October 2017, vol. 10563, p. 291–299

[32] D. A. O. CAMACHO.*A Product-Line for Families of Program Translators: A Grammar-Based Approach*, Université Catholique de Louvain, Belgium, August 2010

[33] R. CARVER, Y. LEI.*Stateless Techniques for Generating Global and Local Test Oracles for Message-Passing Concurrent Programs*, in "Journal of Systems and Software", February 2018, vol. 136, p. 237–265

[34] S. CHABANE, R. AMEUR-BOULIFA, M. MEZGHICHE.*Formal Framework for Automated Analysis and Verification of Distributed Reactive Applications*, in "Proceedings of the 1st International Conference on Embedded and Distributed Systems (EDiS'2017), Oran, Algeria", IEEE Computer Society Press, December 2017

[35] S. CHABANE, R. AMEUR-BOULIFA, M. MEZGHICHE.*Rethinking of I/O-Automata Composition*, in "Proceedings of the 2017 Forum on Specification and Design Languages (FDL'2017), Verona, Italy", F. FUMMI, H. D. PATEL, S. CHAKRABORTY (editors), IEEE Computer Society Press, September 2017, p. 1–7

[36] D. CHAMPELOVIER, X. CLERC, H. GARAVEL, Y. GUERTE, C. MCKINTY, V. POWAZNY, F. LANG, W. SERWE, G. SMEDING.*Reference Manual of the LNT to LOTOS Translator (Version 6.8)*, January 2019, Inria, Grenoble, France

[37] E. M. CLARKE, E. A. EMERSON, A. P. SISTLA.*Automatic Verification of Finite-State Concurrent Systems using Temporal Logic Specifications*, in "ACM Transactions on Programming Languages and Systems", April 1986, vol. 8, n^{o} 2, p. 244–263

[38] R. DE NICOLA, F. W. VAANDRAGER.*Action versus State Based Logics for Transition Systems*, Lecture Notes in Computer Science, Springer Verlag, 1990, vol. 469, p. 407–419

[39] H. GARAVEL.*Nested-Unit Petri Nets: A Structural Means to Increase Efficiency and Scalability of Verification on Elementary Nets*, in "Proceedings of the 36th International Conference on Application and Theory of Petri Nets and Concurrency (PETRI NETS'15), Brussels, Belgium", R. R. DEVILLERS, A. VALMARI (editors), Lecture Notes in Computer Science, Springer Verlag, June 2015, vol. 9115, p. 179–199

[40] H. GARAVEL.*Compilation of LOTOS Abstract Data Types*, in "Proceedings of the 2nd International Conference on Formal Description Techniques FORTE'89 (Vancouver B.C., Canada)", S. T. VUONG (editor), North Holland, December 1989, p. 147–162

[41] H. GARAVEL.*OPEN/CÆSAR: An Open Software Architecture for Verification, Simulation, and Testing*, in "Proceedings of the First International Conference on Tools and Algorithms for the Construction and Analysis of Systems TACAS'98 (Lisbon, Portugal)", Berlin, B. STEFFEN (editor), Lecture Notes in Computer Science, Springer Verlag, March 1998, vol. 1384, p. 68–84, Full version available as Inria Research Report RR-3352

[42] H. GARAVEL, F. LANG.*SVL: a Scripting Language for Compositional Verification*, in "Proceedings of the 21st IFIP WG 6.1 International Conference on Formal Techniques for Networked and Distributed Systems FORTE'2001 (Cheju Island, Korea)", M. KIM, B. CHIN, S. KANG, D. LEE (editors), Kluwer Academic Publishers, August 2001, p. 377–392, Full version available as Inria Research Report RR-4223

[43] H. GARAVEL, F. LANG, R. MATEESCU.*Compiler Construction using LOTOS NT*, in "Proceedings of the 11th International Conference on Compiler Construction CC 2002 (Grenoble, France)", N. HORSPOOL (editor), Lecture Notes in Computer Science, Springer Verlag, April 2002, vol. 2304, p. 9–13

[44] H. GARAVEL, R. MATEESCU, I. SMARANDACHE-STURM.*Parallel State Space Construction for Model-Checking*, in "Proceedings of the 8th International SPIN Workshop on Model Checking of Software SPIN'2001 (Toronto, Canada)", Berlin, M. B. DWYER (editor), Lecture Notes in Computer Science, Springer Verlag, May 2001, vol. 2057, p. 217–234, Revised version available as Inria Research Report RR-4341 (December 2001)

[45] H. GARAVEL, W. SERWE.*State Space Reduction for Process Algebra Specifications*, in "Theoretical Computer Science", February 2006, vol. 351, n^{o} 2, p. 131–145

[46] H. GARAVEL, J. SIFAKIS.*Compilation and Verification of LOTOS Specifications*, in "Proceedings of the 10th International Symposium on Protocol Specification, Testing and Verification (Ottawa, Canada)", L. LOGRIPPO, R. L. PROBERT, H. URAL (editors), North Holland, June 1990, p. 379–394

[47] H. HANSSON, B. JONSSON.*A Logic for Reasoning about Time and Reliability*, in "Formal Aspects of Computing", 1994, vol. 6, n^{o} 5, p. 512–535

[48] A. HARTMANNS, H. HERMANNS.*In the Quantitative Automata Zoo*, in "Science of Computer Programming", 2015, vol. 112, p. 3–23

[49] M. HENNESSY, R. MILNER.*Algebraic Laws for Nondeterminism and Concurrency*, in "Journal of the ACM", 1985, vol. 32, p. 137–161

[50] H. HERMANNS.*Interactive Markov Chains and the Quest for Quantified Quality*, Lecture Notes in Computer Science, Springer Verlag, 2002, vol. 2428

[51] F. KORDON, H. GARAVEL, L. M. HILLAH, F. HULIN-HUBARD, E. AMPARORE, M. BECCUTI, B. BERTHOMIEU, G. CIARDO, S. DAL ZILIO, T. LIEBKE, A. LINARD, J. MEIJER, A. MINER, J. SRBA, Y. THIERRY-MIEG, J. VAN DE POL, K. WOLF.*Complete Results for the 2018 Edition of the Model Checking Contest*, June 2018, http://mcc.lip6.fr/2018/results.php

[52] T. LAMBOLAIS, A.-L. COURBIS.*Development and Verification of UML Architectures by Refinement and Extension Techniques*, in "Proceedings of the 9th European Congress on Embedded Real Time Software and Systems (ERTS2'2018), Toulouse, France", January 2018

[53] J. MAGEE, J. KRAMER.*Concurrency: State Models and Java Programs*, 2006, Wiley, April 2006

[54] F. MARTINELLI, F. MERCALDO, V. NARDONE, A. ORLANDO, A. SANTONE, G. VAGLINI.*Safety Critical Systems Formal Verification Using Execution Traces*, in "Proceedings of the 27th IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE'2018), Paris, France", IEEE Computer Society Press, June 2018, p. 247–250

[55] R. MATEESCU, D. THIVOLLE.*A Model Checking Language for Concurrent Value-Passing Systems*, in "Proceedings of the 15th International Symposium on Formal Methods FM'08 (Turku, Finland)", J. CUELLAR, T. MAIBAUM, K. SERE (editors), Lecture Notes in Computer Science, Springer Verlag, May 2008, vol. 5014, p. 148–164

[56] F. MAZZANTI, A. FERRARI.*Ten Diverse Formal Models for a CBTC Automatic Train Supervision System*, in "Proceedings of the 3rd Workshop on Models for Formal Analysis of Real Systems and 6th International Workshop on Verification and Program Transformation (MARS/VPT@ETAPS'2018), Thessaloniki, Greece", J. P. GALLAGHER, R. VAN GLABBEEK, W. SERWE (editors), EPTCS, April 2018, vol. 268, p. 104–149

[57] F. MAZZANTI, A. FERRARI, G. O. SPAGNOLO.*Towards Formal Methods Diversity in Railways: an Experience Report with Seven Frameworks*, in "Springer International Journal on Software Tools for Technology Transfer (STTT)", June 2018, vol. 20, n^{o} 3, p. 263–288

[58] H. MKAOUAR, B. ZALILA, J. HUGUES, M. JMAIEL.*An Ocarina Extension for AADL Formal Semantics Generation*, in "Proceedings of the 33rd Annual ACM Symposium on Applied Computing (SAC'2018), Pau, France", ACM Computer Society Press, April 2018, p. 1402–1409

[59] A. N. SHESHKALANI, R. KHOSRAVI.*Verification of Visibility-Based Properties on Multiple Moving Robots in an Environment with Obstacles*, in "International Journal of Advanced Robotic Systems", July 2018, vol. 15, n^{o} 4, p. 1–13

[60] M. SZPYRKA, Ł. PODOLSKI, M. WYPYCH.*Modelling and Verification of Real-Time Systems with Alvis*, in "Towards a Synergistic Combination of Research and Practice in Software Engineering", P. KOSIUCZENKO, L. MADEYSKI (editors), Springer International Publishing, 2018, p. 165–178

[61] X. YANG, J. KATOEN, H. LIN, G. LIU, H. WU.*Branching Bisimulation and Concurrent Object Verification*, in "Proceedings of the 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN'2018), Luxembourg City, Luxembourg", IEEE Computer Society Press, June 2018, p. 267–278

[62] R. A. DE OLIVEIRA, J. DINGEL.*Supporting Model Refinement with Equivalence Checking in the Context of Model-driven Engineering with UML-RT*, in "Proceedings of the 15th Workshop on Model-Driven Engineering, Verification and Validation (MoDeVVa'2017) co-located with ACM/IEEE 20th International Conference on Model Driven Engineering Languages and Systems (MODELS'2017), Austin, TX, USA", CEUR Workshop Proceedings, September 2017, vol. 2019, p. 307–314

[63] S. DE PUTTER, A. WIJS.*A Formal Verification Technique for Behavioural Model-to-Model Transformations*, in "Formal Aspects of Computing", January 2018, vol. 30, n^{o} 1, p. 3–43

[64] S. DE PUTTER, A. WIJS.*To Compose, or Not to Compose, That Is the Question: An Analysis of Compositional State Space Generation*, in "Proceedings of the 22nd International Symposium on Formal Methods (FM'2018), Oxford, UK", K. HAVELUND, J. PELESKA, B. ROSCOE, E. DE VINK (editors), Lecture Notes in Computer Science, Springer Verlag, July 2018, vol. 10951, p. 485–504

Project-Team CORSE

compiler optimization and run-time systems

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

IN PARTNERSHIP WITH:
Institut polytechnique de Grenoble

Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Architecture, Languages and Compilation

Table of contents

Project-Team CORSE

Creation of the Team: 2014 November 01, updated into Project-Team: 2016 July 01

Keywords:

Computer Science and Digital Science:

A1.1.1. - Multicore, Manycore
A1.1.3. - Memory models
A1.6. - Green Computing
A2.1.6. - Concurrent programming
A2.1.7. - Distributed programming
A2.1.10. - Domain-specific languages
A2.2. - Compilation
A2.2.1. - Static analysis
A2.2.2. - Memory models
A2.2.4. - Parallel architectures
A2.2.5. - Run-time systems
A2.2.6. - GPGPU, FPGA...
A6.2.7. - High performance computing
A7.1. - Algorithms
A8.2. - Optimization
A8.2.1. - Operations research
A8.4. - Computer Algebra
A8.7. - Graph theory

Other Research Topics and Application Domains:

B4.5. - Energy consumption
B4.5.1. - Green computing
B5.3. - Nanotechnology
B6.1.2. - Software evolution, maintenance
B6.6. - Embedded systems
B6.7. - Computer Industry (harware, equipments...)
B9.1. - Education
B9.8. - Reproducibility

1. Team, Visitors, External Collaborators

Research Scientists

Fabrice Rastello [Team leader, Inria, Senior Researcher, HDR]
Farah Ait Salaht [Inria, Starting Research Position]
Frederic Desprez [Inria, Senior Researcher, HDR]

Faculty Members

Florent Bouchez Tichadou [Univ Grenoble Alpes, Associate Professor]
Francois Broquedis [Institut polytechnique de Grenoble, Associate Professor]
Yliès Falcone [Univ Grenoble Alpes, Associate Professor]

Jean Francois Mehaut [Univ Grenoble Alpes, Professor, until Aug 2018, HDR]

Technical Staff

Ali Kassem [Inria, from Sep 2018]
Nicolas Tollenaere [Inria, from Apr 2018 until Jun 2018]

PhD Students

Georgios Christodoulis [Univ Grenoble Alpes]
Pedro Henrique de Mello Morado Penna [Univ Grenoble Alpes, until Aug 2018]
Antoine El Hokayem [Univ Grenoble Alpes]
Luis Felipe Garlet Millani [Univ Grenoble Alpes, until Aug 2018]
Fabian Gruber [Univ Grenoble Alpes]
Raphael Jakse [Univ Grenoble Alpes]
Mathieu Stoffel [Bull, from Feb 2018]
Philippe Virouleau [Inria, until Feb 2018]
Ye Xia [Orange Labs, until Oct 2018]

Post-Doctoral Fellows

Laercio Lima Pilla [Inria, until Sep 2018]
Manuel Selva [Univ Grenoble Alpes, from Feb 2018]
Francieli Zanon Boito [Inria, until Oct 2018]

Administrative Assistant

Maria Immaculada Presseguer [Inria]

2. Overall Objectives

2.1. Overall Objectives

Languages, compilers, and run-time systems are some of the most important components to bridge the gap between applications and hardware. With the continuous increasing power of computers, expectations are evolving, with more and more ambitious, *computational intensive and complex applications*. As desktop PCs are becoming a niche and servers mainstream, three categories of computing impose themselves for the next decade: mobile, cloud, and super-computing. Thus *diversity, heterogeneity* (even on a single chip) and thus also *hardware virtualization* is putting more and more pressure both on compilers and run-time systems. However, because of the energy wall, *architectures* are becoming more and more *complex* and *parallelism ubiquitous* at every level. Unfortunately, the memory-CPU gap continues to increase and energy consumption remains an important issue for future platforms. To address the challenge of *performance and energy consumption* raised by silicon companies, compilers and run-time systems must *evolve* and, in particular, interact, *taking into account the complexity of the target architecture*.

The overall objective of CORSE is to address this challenge by *combining static and dynamic compilation* techniques, with more interactive *embedding of programs and compiler environment in the run-time system*.

3. Research Program

3.1. Scientific Foundations

One of the characteristics of CORSE is to base our researches on diverse advanced mathematical tools. Compiler optimization requires the usage of the several tools around discrete mathematics: combinatorial optimization, algorithmic, and graph theory. The aim of CORSE is to tackle optimization not only for general purpose but also for domain specific applications. We believe that new challenges in compiler technology design and in particular for split compilation should also take advantage of graph labeling techniques. In addition to run-time and compiler techniques for program instrumentation, hybrid analysis and compilation advances will be mainly based on polynomial and linear algebra.

The other specificity of CORSE is to address technical challenges related to compiler technology, run-time systems, and hardware characteristics. This implies mastering the details of each. This is especially important as any optimization is based on a reasonably accurate model. Compiler expertise will be used in modeling applications (e.g. through automatic analysis of memory and computational complexity); Run-time expertise will be used in modeling the concurrent activities and overhead due to contention (including memory management); Hardware expertise will be extensively used in modeling physical resources and hardware mechanisms (including synchronization, pipelines, etc.).

The core foundation of the team is related to the combination of static and dynamic techniques, of compilation, and run-time systems. We believe this to be essential in addressing high-performance and low energy challenges in the context of new important changes shown by current application, software, and architecture trends.

Our project is structured along two main directions. The first direction belongs to the area of run-time systems with the objective of developing strong relations with compilers. The second direction belongs to the area of compiler analysis and optimization with the objective of combining dynamic analysis and optimization with static techniques. The aim of CORSE is to ground those two research activities on the development of the end-to-end optimization of some specific domain applications.

4. Application Domains

4.1. Transfer

The main industrial sector related to the research activities of CORSE is the one of semi-conductor (programmable architectures spanning from embedded systems to servers). Obviously any computing application which has the objective of exploiting as much as possible the resources (in terms of high-performance but also low energy consumption) of the host architecture is intended to take advantage of advances in compiler and run-time technology. These applications are based over numerical kernels (linear algebra, FFT, convolution...) that can be adapted on a large spectrum of architectures. Members of CORSE already maintain fruitful and strong collaborations with several companies such as STMicroelectronics, Atos/Bull, Kalray.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Christodoulis, G., Broquedis, F., Muller, O., Selva, M., Desprez, F., *An FPGA target for the StarPU heterogeneous runtime system*. ReCoSoC 2018

BEST PAPERS AWARDS :

[25]

G. CHRISTODOULIS, M. SELVA, F. BROQUEDIS, F. DESPREZ, O. MULLER.*An FPGA target for the StarPU heterogeneous runtime system*, in "13th International Symposium on Reconfigurable Communication-centric Systems-on-Chip (RECOSOC 2018)", Lille, France, IEEE, July 2018, p. 1-8, http://hal.univ-grenoble-alpes.fr/hal-01858951

6. New Software and Platforms

6.1. Verde

KEYWORDS: Debug - Verification

FUNCTIONAL DESCRIPTION: Interactive Debugging with a traditional debugger can be tedious. One has to manually run a program step by step and set breakpoints to track a bug.

i-RV is an approach to bug fixing that aims to help developpers during their Interactive Debbugging sessions using Runtime Verification.

Verde is the reference implementation of i-RV.

- Participants: Kevin Pouget, Ylies Falcone, Raphael Jakse and Jean-François Méhaut
- Contact: Raphael Jakse
- Publication: Interactive Runtime Verification - When Interactive Debugging meets Runtime Verification
- URL: https://gitlab.inria.fr/monitoring/verde

6.2. Mickey

KEYWORDS: Dynamic Analysis - Performance analysis - Profiling - Polyhedral compilation
FUNCTIONAL DESCRIPTION: Mickey is a set of tools for profiling based performance debugging for compiled binaries. It uses a dynamic binary translator to instrument arbitrary programs as they are being run to reconstruct the control flow and track data dependencies. This information is then fed to a polyhedral optimizer that proposes structured transformations for the original code.

Mickey can handle both inter- and intra-procedural control and data flow in a unified way, thus enabling inter-procedural structured transformations. It is based on QEMU to allow for portability, both in terms of targeted CPU architectures, but also in terms of programming environment and the use of third-party libraries for which no source code is available.

- Partner: STMicroelectronics
- Contact: Fabian Gruber

6.3. Platforms

6.3.1. Grid'5000

Grid'5000 is a large-scale and versatile testbed for experiment-driven research in all areas of computer science, with a focus on parallel and distributed computing including Cloud, HPC and Big Data. It provides access to a large amount of resources: 12000 cores, 800 compute-nodes grouped in homogeneous clusters, and featuring various technologies (GPU, SSD, NVMe, 10G and 25G Ethernet, Infiniband, Omni-Path) and advanced monitoring and measurement features for traces collection of networking and power consumption, providing a deep understanding of experiments. It is highly reconfigurable and controllable. Researchers can experiment with a fully customized software stack thanks to bare-metal deployment features, and can isolate their experiment at the networking layer advanced monitoring and measurement features for traces collection of networking and power consumption, providing a deep understanding of experiments designed to support Open Science and reproducible research, with full traceability of infrastructure and software changes on the testbed. Frédéric Desprez is director of the GIS GRID5000.

6.3.2. SILECS

Frédéric Desprez is co-PI with Serge Fdida (Université Sorbonne) of the SILECS infrastructure (IR ministère) which goal is to provide an experimental platform for experimental computer Science (Internet of things, clouds, hpc, big data, ...). This new infrastructure is based on two existing infrastructures, Grid'5000 and FIT.

7. New Results

7.1. Profiling Feedback based Optimizations and Performance Debugging

Participants: Fabrice Rastello, Diogo Sampaio, Fabian Gruber, Christophe Guillon [STMicroelectronics], Antoine Moynault [STMicroelectronics], Changwan Hong [OSU, USA], Aravind Sukumaran-Rajam [OSU, USA], Jinsung Kim [OSU, USA], Prashant Singh Rawat [OSU, USA], Sriram Krishnamoorthy [PNNL, USA], Louis-Noël Pouchet [CSU, USA], P. Sadayappan [OSU, USA].

Profiling feedback is an important technique used by developers for performance debugging, where it is usually used to pinpoint performance bottlenecks and also to find optimization opportunities. Our contributions in this area are twofold: (1) we developed a new technique that combines abstract simulation and sensitive analysis that allows to pinpoint performance bottleneck; (2) we developed a new technique to build a polyhedral representation out of an execution trace that allows to provide feedback on possible missed transformations.

7.1.1. Compiler Optimization for GPUs Using Bottleneck Analysis

Optimizing compilers generally use highly simplified performance models due to the significant challenges in developing accurate analytical performance models for complex computer systems. In this work, we develop an alternate approach to performance modeling using abstract execution of GPU kernel binaries. We use the performance model to predict the bottleneck resource for a given kernel's execution through differential analysis by performing multiple abstract executions with varying machine parameters. The bottleneck analysis is then used to develop an automated search through a configuration space of different grid reshaping, thread/block coarsening, and loop unrolling factors. Experimental results using a number of benchmarks from the Parboil/Rodinia/SHOC suites demonstrate the effectiveness of the approach. The bottleneck analysis is also shown to be useful in assisting high-level domain-specific code generators for GPUs.

This work is the fruit of the collaboration 9.4.1.1 with OSU. It has been presented at the ACM/SIGPLAN conference on Programming Language Design and Implementation, PLDI 2018.

7.1.2. Data-Flow/Dependence Profiling for Structured Transformations

Profiling feedback is an important technique used by developers for performance debugging, where it is usually used to pinpoint performance bottlenecks and also to find optimization opportunities. Assessing the validity and potential benefit of a program transformation requires accurate knowledge of the data flow and data dependencies, which can be uncovered by profiling a particular execution of the program.

In this work we develop Mickey, an end-to-end infrastructure for dynamic binary analysis, which produces feedback about the potential to apply structured transformations to uncover non-trivial parallelism and data locality via complex program rescheduling. Our tool can handle both inter- and intraprocedural aspects of the program in a unified way, thus providing structured interprocedural transformation feedback.

This work is the fruit of the collaboration 9.4.1.1 with CSU and the past collaboration Nano2017 with STMicroelectronics. It has been submitted for presentation at the ACM conference on Principles and Practice of Parallel Programming, PPoPP 2019.

7.2. Combined Scheduling and Register Allocation

Participants: Prashant Singh Rawah [OSU, USA], Aravind Sukumaran-Rajam [OSU, USA], Atanas Rountev [OSU, USA], Fabrice Rastello, Louis-Noël Pouchet [CSU, USA], Atanas Rountev [OSU, USA], P. Sadayappan [OSU, USA].

Register allocation is one of the most studied compiler optimization but its impact on performance is highly coupled with scheduling. Recent advances on computer simulation and artificial intelligence lead to application kernels with very high register pressure. Our contributions in this area consist in developing new scheduling schemes that both expose SIMD parallelism and register reuse.

7.2.1. Register Optimizations for Stencils on GPUs

The recent advent of compute-intensive GPU architecture has allowed application developers to explore high-order 3D stencils for better computational accuracy. A common optimization strategy for such stencils is to expose sufficient data reuse by means such as loop unrolling, with the hope of register-level reuse. However, the resulting code is often highly constrained by register pressure. While the current state-of-the-art register allocators are satisfactory for most applications, they are unable to effectively manage register pressure for such complex high-order stencils, resulting in a sub-optimal code with a large number of register spills. In this work, we develop a statement reordering framework that models stencil computations as DAG of trees with shared leaves, and adapts an optimal scheduling algorithm for minimizing register usage for expression trees. The effectiveness of the approach is demonstrated through experimental results on a range of stencils extracted from application codes.

This work is the fruit of the collaboration 9.4.1.1 with OSU. It has been presented at the ACM/SIGPLAN Symposium on Principles and Practice of Parallel Programming, PPoPP 2018.

7.2.2. Associative instruction reordering to alleviate register pressure

Register allocation is generally considered a practically solved problem. For most applications, the register allocation strategies in production compilers are very effective in controlling the number of loads/stores and register spills. However, existing register allocation strategies are not effective and result in excessive register spilling for computation patterns with a high degree of many-to-many data reuse, e.g., high-order stencils and tensor contractions. We develop a source-to-source instruction reordering strategy that exploits the flexibility of reordering associative operations to alleviate register pressure. The developed transformation module implements an adaptable strategy that can appropriately control the degree of instruction-level parallelism, while relieving register pressure. The effectiveness of the approach is demonstrated through experimental results using multiple production compilers (GCC, Clang/LLVM) and target platforms (Intel Xeon Phi, and Intel x86 multi-core).

This work is the fruit of the collaboration 9.4.1.1 with OSU. It has been presented at ACM/IEEE International Conference for High Performance Computing, Networking, Storage, and Analysis, SC 2018.

7.3. Runtime Verification and Monitoring

Participants: Raphael Jakse, Yliès Falcone, Jean Francois Mehaut, Srdan Krstic, Giles Reger, Dmitriy Traytel, Hosein Nazarpour, Mohamad Jaber, Marius Bozga, Saddek Bensalem, Salwa Kobeissi, Adnan Utayim.

We report on several contributions related with the runtime verification and monitoring of systems. We address several aspects such as the instrumentation, the understanding and classification of existing concepts and tools, the definition of frameworks for monitoring distributed systems and a case study on monitoring smart homes.

7.3.1. Interactive Runtime Verification: Formal Models, Algorithms, and Implementation

Interactive runtime verification (i-RV) combines runtime verification and interactive debugging. Runtime verification consists in studying a system at runtime, looking for input and output events to discover, check or enforce behavioral properties. Interactive debugging consists in studying a system at runtime in order to discover and understand its bugs and fix them, inspecting interactively its internal state. We define an efficient and convenient way to check behavioral properties automatically on a program using a debugger. We aim at helping bug discovery and understanding by guiding classical interactive debugging techniques using runtime verification.

In this work, we provide a formal model for a program execution under a debugger, which we compose with a general model of a monitor and a scenario to model the interactively verified program. We provide guarantees on the verdicts issued by the monitor using the instrumentation provided by the debugger. We provide an algorithmic view of this model suitable for producing implementations, and we present Verde, an implementation based on GDB to interactively verify C programs. We built a set of experiments using Verde to assess usefulness of Interactive Runtime Verification and performance of our implementation. Our results show that though debugger-based instrumentation incurs non-trivial performance costs, i-RV is appliable performance-wise in a variety of cases and helps studying bugs.

This work has been submitted at the ACM Transactions on Software Engineering and Methodology (TOSEM).

7.3.2. A Taxonomy for Classifying Runtime Verification Tools

Over the last 15 years Runtime Verification (RV) has grown into a diverse and active field, which has stimulated the development of numerous theoretical frameworks and tools. Many of the tools are at first sight very different and challenging to compare. Yet, there are similarities. In this work, we classify RV tools within a high-level taxonomy of concepts. We first present this taxonomy and discuss the different dimensions. Then, we survey RV tools and classify them according to the taxonomy. This work constitutes a snapshot of the current state of the art and enables a comparison of existing tools.

This work has been published in the proceedings of the 18th International Conference on Runtime Verification.

7.3.3. Bringing Runtime Verification Home

We use runtime verification (RV) to check various specifications in a smart apartment. The specifications can be broken down into three types: be- havioral correctness of the apartment sensors, detection of specific user activities (known as activities of daily living), and composition of specifications of the previous types. The context of the smart apartment provides us with a complex system with a large number of components with two different hierarchies to group specifications and sensors: geographically within the same room, floor or globally in the apartment, and logically following the different types of specifications. We leverage a recent approach to decentralized RV of decentralized specifications, where monitors have their own specifications and communicate together to verify more general specifications. This allows us to re-use specifications, and combine them to: (1) scale beyond existing centralized RV techniques, and (2) greatly reduce computation and communication costs.

This work has been published in the proceedings of the 18th International Conference on Runtime Verification.

7.3.4. Tracing Distributed Component-Based Systems, a Brief Overview

We overview a framework for tracing asynchronous distributed component-based systems with multiparty interactions managed by distributed schedulers. Neither the global state nor the total ordering of the system events is available at runtime. We instrument the system to retrieve local events from the local traces of the schedulers. Local events are sent to a global observer which reconstructs on-the-fly the global traces that are compatible with the local traces, in a concurrency-preserving and communication-delay insensitive fashion. The global traces are represented as an original lattice over partial states, such that any path of the lattice projected on a scheduler represents the corresponding lo- cal partial trace according to that scheduler (soundness), and all possible global traces of the system are recorded (completeness).

This work has been published in the proceedings of the 18th International Conference on Runtime Verification.

7.3.5. Can We Monitor All Multithreaded Programs?

Runtime Verification(RV)is a lightweight formal method which consists in verifying that an execution of a program is correct wrt a specification. The specification formalizes with properties the expected correct behavior of the system. Programs are instrumented to extract necessary information from the execution and feed it to monitors tasked with checking the properties. From the perspective of a monitor, the system is a black box; the trace is the only system information provided. Parallel programs generally introduce an added level of complexity on the program execution due to concurrency. A concurrent execution of a parallel program is best represented as a partial order. A large number of RV approaches generate monitors using formalisms that rely on total order, while more recent approaches utilize formalisms that consider multiple traces.

We made a tutorial where we review some of the main RV approaches and tools that handle multithreaded Java programs. We discuss their assumptions, limitations, expressiveness, and suitability when tackling parallel programs such as producer- consumer and readers-writers. By analyzing the interplay between specification formalisms and concurrent executions of programs, we identify four questions RV practitioners may ask themselves to classify and determine the situations in which it is sound to use the existing tools and approaches.

This work has been published in the proceedings of the 18th International Conference on Runtime Verification.

7.3.6. Facilitating the Implementation of Distributed Systems with Heterogeneous Interactions

We introduce HDBIP an extension of the Behavior Interaction Priority (BIP) framework. BIP is a component-based framework with a rigorous operational semantics and high-level and expressive interaction model. HDBIP extends BIP interaction model by allowing heterogeneous interactions targeting distributed systems. HDBIP allows both multiparty and direct send/receive interactions that can be directly mapped to an underlying communication library. Then, we present a correct and efficient code generation from HDBIP to C++ implementation using Message Passing Interface (MPI). We present a non-trivial case study showing the effectiveness of HDBIP.

This work has been published in the proceedings of the 14th International Conference on integrated Formal Methods.

7.3.7. Modularizing Behavioral and Architectural Crosscutting Concerns in Formal Component-Based Systems

We define a method to modularize crosscutting concerns in Component-Based Systems (CBSs) expressed using the Behavior Interaction Priority (BIP) framework. Our method is inspired from the Aspect Oriented Programming (AOP) paradigm which was initially conceived to support the separation of concerns during the development of monolithic systems. BIP has a formal operational semantics and makes a clear separation between architecture and behavior to allow for compositional and incremental design and analysis of systems. We distinguish local from global aspects. Local aspects model concerns at the component level and are used to refine the behavior of components. Global aspects model concerns at the architecture level, and hence refine communications (synchronization and data transfer) between components. We formalize local and global aspects as well as their composition and integration into a BIP system through rigorous transformation primitives. We present AOP-BIP, a tool for Aspect-Oriented Programming of BIP systems, demonstrate its use to modularize logging, security, and fault tolerance in a network protocol, and discuss its possible use in runtime verification of CBSs.

This work has been published in the Journal of Logical and Algebraic Methods in Programming.

7.4. Numa MeMory Analyzer

Participants: François Trahay [Télécom SudParis], Manuel Selva, Lionel Morel [CEA], Kevin Marquet [INSA Lyon].

Non Uniform Memory Access (NUMA) architectures are nowadays common for running High-Performance Computing (HPC) applications. In such architectures, several distinct physical memories are assembled to create a single shared memory. Nevertheless, because there are several physical memories, access times to these memories are not uniform depending on the location of the core performing the memory request and on the location of the target memory. Hence, threads and data placement are crucial to efficiently exploit such architectures. To help in taking decision about this placement, profiling tools are needed. NUMA MeMory Analyzer (NumaMMA) is a new profiling tool for understanding the memory access patterns of HPC applications. NumaMMA combines efficient collection of memory traces using hardware mechanisms with original visualization means allowing to see how memory access patterns evolve over time. The information reported by NumaMMA allows to understand the nature of these access patterns inside each object allocated by the application. We show how NumaMMA can help understanding the memory patterns of several HPC applications in order to optimize them and get speedups up to 28% over the standard non optimized version.

This work has been published in the 47th International Conference on Parallel Processing - ICPP 2018.

7.5. Towards an Easier Way to Program FPGAs in an HPC Context

Participants: Georgios Christodoulis, Manuel Selva, Francois Broquedis, Frederic Desprez, Olivier Muller [TIMA].

Heterogeneity in HPC nodes appears as a promising solution to improve the execution of a wide range of scientific applications, regarding both performance and energy consumption. Unlike CPUs and GPUs, FPGAs can be configured to fit the application needs, making them an appealing target to extend traditional heterogeneous HPC architectures. However, exploiting them requires an in-depth knowledge of low-level hardware and high expertise on vendor-provided tools, which should not be the primary concern of HPC application programmers. In the context of the Persyval HEAVEN project, we proposed a framework enabling a more straightforward development of scientific applications over FPGA enhanced platforms. Our solution requires the minimum knowledge of the underlying architecture, as well as fewer changes to the existing code. To fulfill these requirements, we extended the StarPU task programming library that initially targets heterogeneous architectures to support FPGAs. We used Vivado HLS, a high-level synthesis tool to deliver efficient hardware implementations of the tasks from high-level languages like C/C++. Our solution, validated on a blocking version of the matrix multiplication algorithm, offers an easier way to exploit FPGAs from an HPC application. We also conducted some preliminary experiments to validate our proof-of-concept implementation regarding performance.

This work has been published in the 13th International Symposium on Reconfigurable Communication-centric Systems-on-Chip and obtained the best paper award.

7.6. Automatic IPC Profile Analysis to Detect Phases in HPC Application

Participants: Mathieu Stoffel, François Broquedis, Frederic Desprez, Abdelhafid Mazouz [Atos/Bull], Philippe Rols [Atos/Bull].

Mathieu Stoffel started his PhD in February 2018 on a CIFRE contract with Atos/Bull. The purpose of this work is to enhance the energy consumption of HPC applications on large-scale platforms. The first phase of the thesis project consists in an in-depth study of the evolution of the metrics characterizing the state of the supercomputer during the execution of a highly parallel application. Indeed, the utilization rates of the different components of the HPC system may demonstrate extreme variations during the execution of the aforementioned application. These variations are sometimes subject to repeat themselves on a regular basis during the application execution. We refer to this phenomena as application "phases". In this context, we already generated precise IPC profiles out of many benchmarks and real-life applications and we worked on a methodology to adapt the CPU frequency based on these profiles. This part of the thesis has been published in an IEEE Cluster workshop (HPCMASPA). Currently, we are working on a detection tool for the application phases. It will implement an automated reconfiguration of the parameters of the HPC system during the execution of the application, in relation with the type of phase being executed. By doing so, the tool will aim at optimizing the energy consumption associated with the execution of the application, by adapting the state of the HPC systems all along the aforesaid execution.

7.7. Teaching of Algorithms, Programming, and Debugging

Participants: Florent Bouchez-Tichadou, Theo Barollet, Aurelien Flori, Thomas Herve.

7.7.1. Teaching Algorithms using Problem and Challenge Based Learning

Teaching algorithms is always a challenge at any level of the CS curriculum, as it is often viewed as a theoretical field. While many exercises revolve around classical examples that illustrate interesting algorithmic points, they are often disconnected from reality, which is a major drawback for students trying to learn. During the last four years, we have been trying to reconnect the teaching of algorithms with their applicability in the real world to M1 and L2 students, by giving them actual problems that could arise in their life of future software engineers, challenging enough to force them to use particular algorithmic techniques or data structures—e.g., linked lists, binary trees, dynamic programming or approximation algorithms.

By assigning students in groups of 5 to 6 members, we wanted to create an environment where they function as a team trying to work together to solve a problem. This allowed them to help each other in their respective comprehension, and made them more autonomous in their learning. The effective materials was provided as online pdf files so they had to read and learn from them by themselves, while the class sessions with a tutor (teacher) where used for the problem-solving part, with guidance from the tutor (who is there to make sure the learning takes place).

After four years of experimentation with M1 students, we found that the student's grades were stable, in particular there was no decrease in exams performances compared to the classical course that was taught in the previous years. However, the students progressed in trans-disciplinary skills such a communication and the writing of essays. More importantly, students show a strong adhesion to the teaching method, 50% of them rating it as “excellent” (6) and 25% as “good” (resp. 6 and 5 on a scale from 1 (terrible) to 6 (excellent)). No student rated the course below average.

This work has been published in the 23rd International Conference on Innovation and Technology in Computer Science Education, ITiCSE 2018.

7.7.2. Data Structures Visualization at Runtime

Debuggers are powerful tools to observe a program behaviour and find bugs but they are not often used by developers and especially beginners because of the hard learning curve of such tools. They provide information on low level data but are not able to analyze higher level elements such as data structures. This work tries to provide a more intuitive representation of the program execution to ease debugging and algorithms understanding.

We have a basic prototype, Moly, which is a GDB extension (GNU Project Debugger) to explore a program runtime memory and analyze its data structures. It also provides an interface with an external visualizer, Lotos, through a formatted output. Running Moly along with a dedicated visualizer should allow a programmer to spot bugs easier by seeing the subsequent whole memory states of the program and some data structures information.

The current status of Moly allows a programmer to explore all attainable memory at any point during the debug process, and already provides minimum information about the possible properties of the data structures, such as recognizing graphs, trees, or linked lists. Future work includes recognizing access patterns to the structures to extract for instance visit patterns and higher-level properties (such as the breaking of data structure properties between break-points).

The external visualizer, Lotos, is still in its early stages of development and was enough to make a proof-of-concept that it is possible to display via a web browser the information gathered by Moly. Our plans is to redesign this part from scratch using the knowledge gaining during the writing of this prototype.

7.7.3. AppoLab: an Online Platform to Engage Students in Their Learning

Classical teaching of algorithms and low-level data structures at the L2 european level is often tedious and unappealing to students, with much of the time being spent on analysing and devising algorithms for textbook cases, such as sorting lists of integers, visiting linked lists or trees, etc.

Using Problem-Based Learning helps to alleviate this problem, by presenting more complex problems to handle, hence engaging more students in their learning. This work revolves around the design of a learning platform that includes gamification in PBL. AppoLab is in its core a server that has scripted "exercices". Students can communicate with the server either manually, using telnet; but ultimately, they will need to script the communication also from their side, since the server will gradually impose constraints on the problems such as timeouts or large input sizes.

This preliminary work was used this year in some parts of an Algorithm course at the L2 level, and has received positive feedback from the students. This encourages us to continue this development and study more precisely the impact it has on students' engagement in their learning.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- CORSE is involved in a contract with Atos/Bull which objective is the objective is to optimize the energy consumption of HPC applications on large scale plateforms.

8.2. Bilateral Grants with Industry

- ES3CAP is a bilateral grant with Kalray. CORSE is involved in the optimisation of machine learning algorithms for many-core architectures.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. HEAVEN Persyval Project

- Title: HEterogenous Architectures: Versatile Exploitation and programiNg
- HEAVEN leaders: François Broquedis, Olivier Muller [TIMA lab]
- CORSE participants: François Broquedis, Frédéric Desprez, Georgios Christodoulis, Manuel Selva
- Computer architectures are getting more and more complex, exposing massive parallelism, hierarchically-organized memories and heterogeneous processing units. Such architectures are extremely difficult to program as they most of the time make application programmers choose between portability and performance.

  While standard programming environments like OpenMP are currently evolving to support the execution of applications on different kinds of processing units, such approaches suffer from two main issues. First, to exploit heterogeneous processing units from the application level, programmers need to explicitly deal with hardware-specific low-level mechanisms, such as the memory transfers between the host memory and private memories of a co-processor for example. Second, as the evolution of programming environments towards heterogeneous programming mainly focuses on CPU/GPU platforms, some hardware accelerators are still difficult to exploit from a general-purpose parallel application.

  FPGA is one of them. Unlike CPUs and GPUs, this hardware accelerator can be configured to fit the application needs. It contains arrays of programmable logic blocks that can be wired together to build a circuit specialized for the targeted application. For example, FPGAs can be configured to accelerate portions of code that are known to perform badly on CPUs or GPUs. The energy efficiency of FPGAs is also one of the main assets of this kind of accelerators compared to GPUs, which encourages the scientific community to consider FPGAs as one of the building blocks of large scale low-power heterogeneous multicore platforms.

  However, only a fraction of the community considers programming FPGAs for now, as configurations must be designed using low-level description languages such as VHDL that application programmers are not experienced with.

  The main objective of this project is to improve the accessibility of heterogeneous architectures containing FPGA accelerators to parallel application programmers. The proposed project focuses on three main aspects:

  – Portability: we don't want application programmers to redesign their applications completely to benefit from FPGA devices. This means extending standard parallel programming environments like OpenMP to support FPGA. Improving application portability also means leveraging most of the hardware-specific low-level mechanisms at the run-time system level;

– Performance: we want our solution to be flexible enough to get the most out of any heterogeneous platforms containing FPGA devices depending on specific performance needs, like computation throughput or energy consumption for example;
– Experiments: Experimenting with FPGA accelerators on real-life scientific applications is also a key element of our project proposal. In particular, the solutions developed in this project will allow comparisons between architectures on real-life applications from different domains like signal processing and computational finance.

Efficient programming and exploitation of heterogeneous architectures implies the development of methods and tools for system design, embedded or not. The HEAVEN project proposal fits in the PCS research action of the PERSYVAL-lab. The PhD of Georgios Christodoulis and the PostDoc of Manuel Selva are funded by this project.

9.2. National Initiatives

9.2.1. PIA ELCI

- Title: Software environment for computation-intensive applications
- Coordinator: Corinne Marchand (BULL SAS)
- CORSE participants: François Broquedis, Philippe Virouleau
- INRIA Partners: Avalon, Cardamon, Myriads; Realopt, Roma, Storm, Tadaam
- Other Partners: Algo'Tech, CEA, Cenaero, CERFACS, CORIA, Kitware, Onera, SAFRAN
- Duration: from Sept. 2014 to March 2018
- Abstract: The ELCI project main goal is to develop a highly-scalable new software stack to tackle high-end supercomputers, from numerical solvers to programming environments and run-time systems. In particular, the CORSE team is studying the scalability of OpenMP run-time systems on large scale shared memory machines through the PhD of Philippe Virouleau, co-advised by researchers from the CORSE and AVALON Inria teams. This work intends to propose new approaches based on a compiler/run-time cooperation to improve the execution of scientific task-based programs on NUMA platforms. The PhD of Philippe Virouleau is funded by this project.

9.2.2. IPL ZEP

- Title: Zero-Power computing systems
- Coordinator: Kevin Marquet (INRIA Socrate)
- CORSE participants: Fabrice Rastello
- Other INRIA Partners: Cairn, Pacap
- Duration: from Apr. 2017 to Sept. 2019
- Abstract: The ZEP project addresses the issue of designing tiny computing objects with no battery by combining non-volatile memory (NVRAM), energy harvesting, micro-architecture innovations, compiler optimizations, and static analysis. The main application target is Internet of Things (IoT) where small communicating objects will be composed of this computing part associated to a low-power wake-up radio system. The ZEP project gathers four Inria teams that have a scientific background in architecture, compilation, operating system and low power together with the CEA Lialp and Lisan laboratories of CEA LETI & LIST. The major outcomes of the project will be a prototype harvesting board including NVRAM and the design of a new microprocessor associated with its optimizing compiler and operating system.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. EoCoE

Title: Energy oriented Centre of Excellence for computer applications

Programm: H2020

Duration: October 2015 - October 2018

Coordinator: CEA

Partners:

Barcelona Supercomputing Center - Centro Nacional de Supercomputacion (Spain)

Commissariat A L Energie Atomique et Aux Energies Alternatives (France)

Centre Europeen de Recherche et de Formation Avancee en Calcul Scientifique (France)

Consiglio Nazionale Delle Ricerche (Italy)

The Cyprus Institute (Cyprus)

Agenzia Nazionale Per le Nuove Tecnologie, l'energia E Lo Sviluppo Economico Sostenibile (Italy)

Fraunhofer Gesellschaft Zur Forderung Der Angewandten Forschung Ev (Germany)

Instytut Chemii Bioorganicznej Polskiej Akademii Nauk (Poland)

Forschungszentrum Julich (Germany)

Max Planck Gesellschaft Zur Foerderung Der Wissenschaften E.V. (Germany)

University of Bath (United Kingdom)

Universite Libre de Bruxelles (Belgium)

Universita Degli Studi di Trento (Italy)

INRIA contact: Michel Kern

CORSE contact: Jean Francois Méhaut

CORSE participants: Jean Francois Méhaut, Frédéric Desprez and Francieli Zanon Boito

The aim of the present proposal is to establish an Energy Oriented Centre of Excellence for computing applications, (EoCoE). EoCoE (pronounce "Echo") will use the prodigious potential offered by the ever-growing computing infrastructure to foster and accelerate the European transition to a reliable and low carbon energy supply. To achieve this goal, we believe that the present revolution in hardware technology calls for a similar paradigm change in the way application codes are designed. EoCoE will assist the energy transition via targeted support to four renewable energy pillars: Meteo, Materials, Water and Fusion, each with a heavy reliance on numerical modeling. These four pillars will be anchored within a strong transverse multidisciplinary basis providing high-end expertise in applied mathematics and HPC. EoCoE is structured around a central Franco-German hub coordinating a pan-European network, gathering a total of 8 countries and 23 teams. Its partners are strongly engaged in both the HPC and energy fields; a prerequisite for the long-term sustainability of EoCoE and also ensuring that it is deeply integrated in the overall European strategy for HPC. The primary goal of EoCoE is to create a new, long lasting and sustainable community around computational energy science. At the same time, EoCoE is committed to deliver high-impact results within the first three years. It will resolve current bottlenecks in application codes, leading to new modeling capabilities and scientific advances among the four user communities; it will develop cutting-edge mathematical and numerical methods, and tools to foster the usage of Exascale computing. Dedicated services for laboratories and industries will be established to leverage this expertise and to foster an ecosystem around HPC for energy. EoCoE will give birth to new collaborations and working methods and will encourage widely spread best practices.

Francieli Zanon Boito started in November 2017 as post-doc for the EoCoe project. She is working with Frédéric Desprez, Thierry Deutsch (CEA INAC) and Jean Francois Méhaut. Francieli is investigating the data storage issues for the scientific workflows on the nano-scale characterization center (PFNC@Minatec http://inac.cea.fr/en/Phocea/Vie_des_labos/Ast/ast_technique.php?id_ast=217).

9.3.1.2. PRACE-5IP

Title: PRACE-5IP (PRACE Fifht Implementation Phase)

Program H2020

Duration: 01/01/2013 - 30/04/2019

Inria partners: Hiepacs team (Inria Bordeaux Sud-Ouest), Storm team (Inria Bordeaux Sud-Ouest), Nachos team (Inria Sophia Antipolis Méditerranée), CORSE team (Inria Grenoble Rhône Alpes) INRIA contact: Stéphane Lanteri (Nachos, Sophia Antipolis)

CORSE contact: Jean Francois Méhaut

CORSE participants: François Broquedis, Jean Francois Méhaut

The objectives of PRACE-5IP are to build on and seamlessly continue the successes of PRACE and start new innovative and collaborative activities proposed by the consortium. These include:

- assisting the transition to PRACE2 including analysis of TransNational Access;
- strengthening the internationally recognized PRACE brand;
- continuing and extend advanced training which so far provided more than 18800 person-training days;
- preparing strategies and best practices towards Exascale computing;
- coordinating and enhancing the operation of the multi-tier HPC systems and services;
- supporting users to exploit massively parallel systems and novel architectures.

The INRIA contribution is in the prolongation of involvement (jointly with CINES) in PRACE 4IP – WP7. The participation of Inria's researchers has been enlarged to include project-teams that were all involved in the C2S@Exa Inria Project Lab. The Inria teams will contribute to the WP7 and the following sub-tasks:

- Task 7.1: Applications Enabling Services for PRACE systems
- Task 7.4 Provision of Numerical Libraries for Heterogeneous/Hybrid Architectures

The activities are organized along two complementary lines

- Generic (or transverse) technologies for simulation software
- Specific (or vertical) technologies i.e. simulation software

The CORSE activities for PRACE-5IP will start with the hiring of one year postdoc in 2018. We will work on the DIOGENEs (DisOntinous GalErkin Nanoscale Solvers) software suite developed in the Nachos team. The post-doc will investigate the new vectorization features of processors.

9.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: COST

Project acronym: ArVI

Project title: Run-Time Verification beyond Monitoring

Duration: December 2014 - Dec 2018

Coordinator: Martin Leucker, University of Lubeck

Abstract: Run-Time verification (RV) is a computing analysis paradigm based on observing a system at run-time to check its expected behavior. RV has emerged in recent years as a practical application of formal verification, and a less ad-hoc approach to conventional testing by building monitors from formal specifications.

There is a great potential applicability of RV beyond software reliability, if one allows monitors to interact back with the observed system, and generalizes to new domains beyond computers programs (like hardware, devices, cloud computing and even human centric systems). Given the European leadership in computer based industries, novel applications of RV to these areas can have an enormous impact in terms of the new class of designs enabled and their reliability and cost effectiveness.

This Action aims to build expertise by putting together active researchers in different aspects of run-time verification, and meeting with experts from potential application disciplines. The main goal is to overcome the fragmentation of RV research by (1) the design of common input formats for tool cooperation and comparison; (2) the evaluation of different tools, building a growing sets benchmarks and running tool competitions; and (3) by designing a road-map and grand challenges extracted from application domains.

9.4. International Initiatives

9.4.1. Inria Associate Teams Not Involved in an Inria International Labs

9.4.1.1. IOComplexity

Title: Automatic characterization of data movement complexity

International Partner (Institution - Laboratory - Researcher):

Ohio State University (United States) - Computer Science and Artificial Intelligence Laboratory - P. Sadayappan

Start year: 2018

See also: https://team.inria.fr/corse/iocomplexity/

The goal of this project is to extend techniques for automatic characterisation of data movement of an application to the design of performance estimation.

The EA as three main objectives: 1. broader applicability of IO complexity analysis; 2. Hardware characterisation; 3. Performance model.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Mohamad Jaber visited the Inria Corse team in January 2018.

9.5.2. Visits to International Teams

9.5.2.1. Sabbatical programme

- Fabrice Rastello was on sabbatical at Colorado State University (USA) from July 2017 till July 2018.
- Yliès Falcone visited American University of Beirut (Lebanon) in May 2018 through an Erasmus exchange programme.

9.5.2.2. Research Stays Abroad

- Fabian Gruber visited the Colorado State University to work with Louis-Noël Pouchet from 18.03.2018 to 17.04.2018.
- Fabian Gruber visited the Ohio State University to work with P. Sadayappan , Changwan Hong, and Aravind Sukumaran-Rajam from 18.11.2018 to 01.12.2018.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Yliès Falcone chaired the programme committee of the Software Verification and Testing track of the 2018 ACM Symposium on Applied Computing.
- Yliès Falcone chaired the scientific organization of the 2nd international school on Runtime Verification.

10.1.1.2. Member of the Organizing Committees

- Fabrice Rastello: Steering Committee ACM/IEEE CGO; Steering Committee “Journées française de la compilation”

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

- Frédéric Desprez: Closer 2018, HPC ’18, SC18 (posters), CEBDA-2018 (with IPDPS’18), CLOUDCOM-2018.
- François Broquedis: IEEE IPDPS 2019, COMPAS 2019
- Fabrice Rastello: ACM SIGPLAN/SIGBED LCTES 2018
- Yliès Falcone: RUME’18, VORTEX’18, 4PAD’18, RV’18, TASE’18, DATE’18

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Frédéric Desprez: IEEE Transaction on Cloud Computing (associate editor)

10.1.4. Invited Talks

- Fabrice Rastello: Saarbruck University: "Automated Derivation of Roofline Performance Limits for Affine Programs"
- Fabrice Rastello: UC Denver: "Automated Derivation of Roofline Performance Limits for Affine Programs"
- Fabrice Rastello: CSU: "Data-Flow/Dependence Profiling for Structured Transformations"
- Frédéric Desprez: CCDSC workshop: SILECS: Super Infrastructure for Large-scale Experimental Computer Science

10.1.5. Leadership within the Scientific Community

- Frédéric Desprez: co-présidence du prix de thèse annuel du GDR Réseaux et Systèmes Distribués (RSD) en collaboration avec l’association ACM SIGOPS France (ASF)
- Frédéric Desprez: Scientific committee of ORAP
- Frédéric Desprez: Technical Committee of GENCI

10.1.6. Scientific Expertise

- Frédéric Desprez: Genci: attribution heures de calcul CT6
- Frédéric Desprez: Groupe de travail “Cloud recherche” du ministère
- Frédéric Desprez: comité des sages IRIT
- Frédéric Desprez: Netherlands Organisation for Scientific Research (NWO), TOP Grants for senior researchers

10.1.7. Research Administration

- Frédéric Desprez: Deputy Scientific Director at INRIA
- Frédéric Desprez: Director of the GIS GRID5000
- Frédéric Desprez: Conseil Scientifique ESIEE Paris

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master 1: Frédéric Desprez, Parallel Algorithms and Programming, 30 hours, M1 MoSIG and CS, UGA, France

License 3: François Broquedis, Imperative programming using python, 40 hours, Grenoble Institute of Technology (Ensimag)

License 3: François Broquedis, Computer architecture, 40 hours, Grenoble Institute of Technology (Ensimag)

License 3: François Broquedis, C programming, 80 hours, Grenoble Institute of Technology (Ensimag)

Master 1: François Broquedis, Operating systems and concurrent programming, 40 hours, Grenoble Institute of Technology (Ensimag)

Master 1: François Broquedis, Operating Systems Development Project - Fundamentals, 20 hours, Grenoble Institute of Technology (Ensimag)

Master 1: François Broquedis, Operating Systems Project, 20 hours, Grenoble Institute of Technology (Ensimag)

Master: Florent Bouchez Tichadou, Algorithmic Problem Solving, 41 hours, M1 MoSIG

Licence: Florent Bouchez Tichadou, Algorithms languages and programming, 113 hours, L2 UGA

Licence: Florent Bouchez Tichadou is responsible of the second year of INF (informatique) and MIN (mathématiques et informatique) students at UGA, eq. 85 hours

Master 1: Yliès Falcone, Proof Techniques and Logic Reminders, MoSIG, 3 hours

Master 1: Yliès Falcone, Programming Language Semantics and Compiler Design, MoSIG and Master informatique, 96 hours

License: Yliès Falcone, Languages and Automata, Univ. Grenoble Alpes, 105 hours

Master: Yliès Falcone, is co-responsible of the first year of the International Master of Computer Science (Univ. Grenoble Alpes and INP ENSIMAG)

10.2.2. Supervision

PhD in progress: Georgios Christodoulis, Adaptation of a heterogeneous run-time system to efficiently exploit FPGA, October 2015, advised by Frederic Desprez, Olivier Muller (TIMA/SLS), and François Broquedis

PhD in progress: Mathieu Stoffel, Static and dynamic approaches for the optimization of the energy consumption associated with applications of the High Performance Computing (HPC) field, February 2018, advised by François Broquedis, Frédéric Desprez, Abdelhafid Mazouz (Atos/Bull) and Philippe Rols (Atos/Bull)

PhD: Ye Xia, Scaling and placement for autonomic management of elasticity of applications in a widely distributed cloud, defended on December 17th 2018, Combining Heuristics for Optimizing and Scaling the Placement of IoT Applications in the Fog, advised by Thierry Coupaye (Orange), Frédéric Desprez, and Xavier Etchevers (Orange)

PhD in progress: Fabian Grüber, Interactive & iterative performance debugging, September 2016, advised by Fabrice Rastello and Yliès Falcone.

PhD: François Gindraud, Semantics and compilation for a data-flow model with a global address space and software cache coherency. Defended on January 11 2018, advised by Fabrice Rastello and Albert Cohen.

PhD: Thomas Messi Nguelé, Domain Specific Languages for Social Networks Analysis on Multi-Core Architectures, defended on September 15 2018, advised by Maurice Tchuenté (Yaoundé I, LIRIMA) and Jean Francois Méhaut

PhD: Philippe Virouleau, Improving the performance of task-based run-time systems on large scale NUMA machines, defended on June 5 2018, advised by Thierry Gautier (INRIA/AVALON), Fabrice Rastello, and François Broquedis

PhD: Antoine El-Hokayem, Decentralised and Distributed Monitoring of Cyber-Physical Systems, defended on December 18 2018, advised by Yliès Falcone.

PhD in progress: Pedro Henrique Penna, Towards an Operating System for Manycore Platforms, October 2017, advised by Marcio Castro (UFSC), François Broquedis, Henrique Cota de Freitas (PUC Minas) and Jean Francois Méhaut.

PhD in progress: Raphaël Jakse, Interactive Runtime Verification, to be defended in Fall 2019, advised by Jean-François Méhaut and Yliès Falcone.

10.2.3. Juries

10.2.3.1. Frédéric Desprez

- François Gindraud, examiner, *Semantics and compilation for a data-flow model with a global address space and software cache coherency*, PhD, Université Grenoble Alpes, January 11, 2018
- Guillaume Latu, reviewer, *Contribution à la simulation haute-performance et aux méthodes de calcul très extensibles*, HDR, Université de Strasbourg, April 18, 2018
- Bastien Confais, reviewer, *Conception d'un système de partage de données adapté à un environnement de Fog Computing*, PhD, Université de Nantes, July 10, 2018
- Hadrien Croubois, examiner/chair, *Toward an autonomic engine for scientific workflows and elastic Cloud infrastructure*, PhD, ENS Lyon, October 16 2018
- Estelle Dirand, examiner, *Développement d'un système in situ à base de tâches pour un code de dynamique moléculaire classique adapté aux machines exaflopiques*, PhD, Université Grenoble Alpes, November 6, 2018
- Ovidiu Marcu, reviewer, *KerA : A Unified Ingestion and Storage System for Scalable Big Data Processing*, PhD, Insa de Rennes, December 18, 2018
- Mohamed Abderrahim, reviewer, *Conception d'un système de supervision programmable et reconfigurable pour une infrastructure informatique et réseau répartie*, IMT Atlantique, December 19, 2018

10.2.3.2. Fabrice Rastello

- François Gindraud, advisor, *Système distribué à adressage global et cohérence logicielle pourl'exécution d'un modèle de tâche à flot de données*, Université Grenoble Alpes, January 11, 2018
- Johannes Doerfert, reviewer, *Applicable and Sound Polyhedral Optimization of Low-Level Programs*, Universität des Saarlandes, December 19, 2018
- Philippe Virouleau, advisor, *Etude et amélioration de l'exploitation des architectures NUMA à travers des supports exécutifs*, Université Grenoble Alpes, June 5, 2018

10.2.3.3. François Broquedis

- Philippe Virouleau, advisor, *Etude et amélioration de l'exploitation des architectures NUMA à travers des supports exécutifs*, Université Grenoble Alpes, June 5, 2018

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- Yliès Falcone: Elected member of the Research Council of Univ. Grenoble Alpes.
- Yliès Falcone: Elected member of the Academic Council of Univ. Grenoble Alpes.

- Yliès Falcone: Elected member of the Laboratory Council of the Laboratoire d'Informatique de Grenoble
- Yliès Falcone: Mission Valorisation for the Laboratoire d'Informatique de Grenoble.

11. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] J. BIGOT, V. GRANDGIRARD, G. LATU, J.-F. MÉHAUT, L. F. MILLANI, C. PASSERON, S. Q. MASNADA, J. RICHARD, B. VIDEAU.*Building and Auto-Tuning Computing Kernels: Experimenting with BOAST and StarPU in the GYSELA Code*, in "ESAIM: Proceedings and Surveys", October 2018, vol. 63 (2018), p. 152 - 178 [*DOI :* 10.1051/PROC/201863152], https://hal.inria.fr/hal-01909325

[2] A. EL-HOKAYEM, Y. FALCONE, M. JABER.*Modularizing Behavioral and Architectural Crosscutting Concerns in Formal Component-Based Systems - Application to the Behavior Interaction Priority Framework*, in "Journal of Logical and Algebraic Methods in Programming", 2018, vol. 99, p. 143–177 [*DOI :* 10.1016/J.JLAMP.2018.05.005], https://hal.inria.fr/hal-01796786

[3] P. J. PAVAN, R. K. LORENZONI, V. MACHADO, J. BEZ, E. PADOIN, F. ZANON BOITO, P. NAVAUX, J.-F. MÉHAUT.*Energy Efficiency and I/O Performance of Low-Power Architectures*, in "Concurrency and Computation: Practice and Experience", 2018 [*DOI :* 10.1002/CPE.4948], https://hal.inria.fr/hal-01784497

[4] P. H. PENNA, A. T. A. GOMES, M. CASTRO, P. PLENTZ, H. C. D. FREITAS, F. BROQUEDIS, J.-F. MEHAUT.*A Comprehensive Performance Evaluation of the BinLPT Workload-Aware Loop Scheduler*, in "Concurrency and Computation: Practice and Experience", 2019, https://hal.archives-ouvertes.fr/hal-01986361

[5] M. RENARD, Y. FALCONE, A. ROLLET, T. JÉRON, H. MARCHAND.*Optimal Enforcement of (Timed) Properties with Uncontrollable Events*, in "Mathematical Structures in Computer Science", 2019, vol. 29, n^{o} 1, p. 169-214 [*DOI :* 10.1017/S0960129517000123], https://hal.archives-ouvertes.fr/hal-01262444

[6] B. VIDEAU, K. POUGET, L. GENOVESE, T. DEUTSCH, D. KOMATITSCH, F. DESPREZ, J.-F. MÉHAUT.*BOAST: A metaprogramming framework to produce portable and efficient computing kernels for HPC applications*, in "International Journal of High Performance Computing Applications", January 2018, vol. 32, n^{o} 1, p. 28-44 [*DOI :* 10.1177/1094342017718068], https://hal.archives-ouvertes.fr/hal-01620778

[7] N. ZHOU, G. DELAVAL, B. ROBU, E. RUTTEN, J.-F. MÉHAUT.*An Autonomic-Computing Approach on Mapping Threads to Multi-cores for Software Transactional Memory*, in "Concurrency and Computation: Practice and Experience", September 2018, vol. 30, n^{o} 18, e4506 [*DOI :* 10.1002/CPE.4506], https://hal.archives-ouvertes.fr/hal-01742690

Invited Conferences

[8] Y. FALCONE.*Second School on Runtime Verification, as part of the ArVi COST Action 1402 Overview and Reflections*, in "RV 2018 - 18th International Conference on Runtime Verification", Limassol, Cyprus, November 2018, p. 1-5, https://hal.inria.fr/hal-01882413

International Conferences with Proceedings

[9] C. COLOMBO, Y. FALCONE, M. LEUCKER, G. REGER, C. SANCHEZ, G. SCHNEIDER, V. STOLZ.*COST Action IC1402 Runtime Verification beyond Monitoring*, in "RV 2018 - 18th International Conference on Runtime Verification", Limassol, Cyprus, November 2018, p. 1-8, https://hal.inria.fr/hal-01900195

[10] A. EL-HOKAYEM, Y. FALCONE.*Bringing Runtime Verification Home*, in "RV 2018 - 18th International Conference on Runtime Verification", Limassol, Cyprus, November 2018, p. 1-17, https://hal.inria.fr/hal-01882411

[11] A. EL-HOKAYEM, Y. FALCONE.*Can We Monitor All Multithreaded Programs?*, in "RV 2018 - 18th International Conference on Runtime Verification", Limassol, Cyprus, November 2018, p. 1-24, https://hal.inria.fr/hal-01882414

[12] S. FABRE, J. LUÍS GÜNTZEL, L. LIMA PILLA, R. NETTO, T. FONTANA, V. LIVRAMENTO.*Enhancing Multi-Threaded Legalization Through k-d Tree Circuit Partitioning*, in "SBCCI 2018 - 31st Symposium on Integrated Circuits and Systems Design", Bento Gonçalves, Brazil, August 2018, p. 1-9, https://hal.inria.fr/hal-01872451

[13] Y. FALCONE, S. KRSTIĆ, G. REGER, D. TRAYTEL.*A Taxonomy for Classifying Runtime Verification Tools*, in "RV 2018 - 18th International Conference on Runtime Verification", Limassol, Cyprus, November 2018, p. 1-18, https://hal.inria.fr/hal-01882410

[14] Y. FALCONE, H. NAZARPOUR, M. JABER, M. BOZGA, S. BENSALEM.*Tracing Distributed Component-Based Systems, a Brief Overview*, in "Proceedings of the 18th International Conference on Runtime Verification", Limassol, Cyprus, November 2018, https://hal.inria.fr/hal-01882412

[15] V. FREITAS, A. SANTANA, M. CASTRO, L. LIMA PILLA.*A Batch Task Migration Approach for Decentralized Global Rescheduling*, in "SBAC-PAD 2018 - International Symposium on Computer Architecture and High Performance Computing", Lyon, France, September 2018, p. 1-12, https://hal.inria.fr/hal-01860626

[16] C. HONG, A. SUKUMARAN-RAJAM, J. KIM, P. S. RAWAT, S. KRISHNAMOORTHY, L.-N. POUCHET, F. RASTELLO, P. SADAYAPPAN.*GPU Code Optimization using Abstract Kernel Emulation and Sensitivity Analysis*, in "PLDI 2018 - 39th ACM SIGPLAN Conference on Programming Language Design and Implementation", Philadelphia, United States, June 2018, p. 736-751 [*DOI :* 10.1145/3192366.3192397], https://hal.inria.fr/hal-01955475

[17] S. KOBEISSI, A. UTAYIM, M. JABER, Y. FALCONE.*Facilitating the Implementation of Distributed Systems with Heterogeneous Interactions*, in "IFM 2018 - 14th International Conference on integrated Formal Methods", Maynooth, Ireland, September 2018, p. 1-19, https://hal.inria.fr/hal-01868748

[18] A. RAMOS CARNEIRO, J. LUCA BEZ, F. ZANON BOITO, B. A. FAGUNDES, C. OSTHOFF, P. NAVAUX.*Collective I/O Performance on the Santos Dumont Supercomputer*, in "PDP 2018 - 26th Euromicro International Conference on Parallel, Distributed and Network-based Processing", Cambridge, United Kingdom, IEEE, March 2018, p. 45-52 [*DOI :* 10.1109/PDP2018.2018.00015], https://hal.inria.fr/hal-01711359

[19] P. SINGH, A. SUKUMARAN-RAJAM, A. ROUNTEV, F. RASTELLO, L.-N. POUCHET, P. SADAYAPPAN.*Register Optimizations for Stencils on GPUs*, in "PPoPP 2018 - 23rd ACM SIGPLAN Symposium on

Principles and Practice of Parallel Programming", Vienna, Austria, February 2018, p. 1-15, https://hal.inria.fr/hal-01955542

[20] P. SINGH RAWAT, A. SUKUMARAN-RAJAM, A. ROUNTEV, F. RASTELLO, L.-N. POUCHET, P. SADAYAPPAN.*Associative Instruction Reordering to Alleviate Register Pressure*, in "SC 2018 - International Conference for High Performance Computing, Networking, Storage, and Analysis", Dallas, United States, November 2018, p. 1-13, https://hal.inria.fr/hal-01956260

[21] Y. XIA, X. ETCHEVERS, L. LETONDEUR, T. COUPAYE, F. DESPREZ.*Combining hardware nodes and software components ordering-based heuristics for optimizing the placement of distributed IoT applications in the fog*, in "SAC 2018 - 33rd Annual ACM/SIGAPP Symposium on Applied Computing", Pau, France, ACM Press, April 2018, p. 751-760 [*DOI :* 10.1145/3167132.3167215], https://hal.inria.fr/hal-01908928

[22] Y. XIA, X. ETCHEVERS, L. LETONDEUR, A. LEBRE, T. COUPAYE, F. DESPREZ.*Combining Heuristics to Optimize and Scale the Placement of IoT Applications in the Fog*, in "UCC 2018 - 11th IEEE/ACM Conference on Utility and Cloud Computing", Zurich, Switzerland, December 2018, p. 1-11, https://hal.inria.fr/hal-01942097

National Conferences with Proceeding

[23] A. SANTANA, V. FREITAS, M. CASTRO, L. LIMA PILLA, J.-F. MÉHAUT.*Reducing Global Schedulers' Complexity Through Runtime System Decoupling*, in "WSCAD 2018 - XIX Simpósio de Sistemas Computacionais de Alto Desempenho", São Paulo, Brazil, October 2018, p. 1-12, https://hal.inria.fr/hal-01873526

Conferences without Proceedings

[24] F. BOUCHEZ-TICHADOU.*Problem solving to teach advanced algorithms in heterogeneous groups*, in "ITiCSE 2018 - 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education", Larnaca, Cyprus, ACM Press, July 2018, p. 200-205 [*DOI :* 10.1145/3197091.3197147], https://hal.archives-ouvertes.fr/hal-01929650

[25] *Best Paper*
G. CHRISTODOULIS, M. SELVA, F. BROQUEDIS, F. DESPREZ, O. MULLER.*An FPGA target for the StarPU heterogeneous runtime system*, in "13th International Symposium on Reconfigurable Communication-centric Systems-on-Chip (RECOSOC 2018)", Lille, France, IEEE, July 2018, p. 1-8, http://hal.univ-grenoble-alpes.fr/hal-01858951.

[26] F. TRAHAY, M. SELVA, L. MOREL, K. MARQUET.*NumaMMA: NUMA MeMory Analyzer*, in "ICPP 2018 - 47th International Conference on Parallel Processing", Eugene, United States, August 2018, p. 1-10 [*DOI :* 10.1145/3225058.3225094], https://hal-cea.archives-ouvertes.fr/cea-01854072

Scientific Books (or Scientific Book chapters)

[27] E. BARTOCCI, Y. FALCONE.*Lectures on Runtime Verification. Introductory and Advanced Topics*, LNCS, Springer, February 2018, vol. 10457, p. 1-240 [*DOI :* 10.1007/978-3-319-75632-5], https://hal.inria.fr/hal-01762298

[28] E. BARTOCCI, Y. FALCONE, A. FRANCALANZA, G. REGER.*Introduction to Runtime Verification*, in "Lectures on Runtime Verification. Introductory and Advanced Topics", Lecture Notes in Computer Science,

Springer, February 2018, vol. 10457, p. 1-33 [*DOI :* 10.1007/978-3-319-75632-5_1], https://hal.inria.fr/hal-01762297

[29] Y. FALCONE, L. MARIANI, A. ROLLET, S. SAHA.*Runtime Failure Prevention and Reaction*, in "Lectures on Runtime Verification", Lecture Notes in Computer Science, Springer, February 2018, vol. 10457, p. 103-134 [*DOI :* 10.1007/978-3-319-75632-5_4], https://hal.archives-ouvertes.fr/hal-01723606

Research Reports

[30] F. GRUBER, M. SELVA, D. SAMPAIO, C. GUILLON, L.-N. POUCHET, F. RASTELLO.*Building of a Polyhedral Representation from an Instrumented Execution: Making Dynamic Analyses of non-Affine Programs Scalable*, CORSE - Compiler Optimization and Run-time Systems, January 2019, n° RR-9244, https://hal.inria.fr/hal-01967828

[31] L. LIMA PILLA.*Basics of Vectorization for Fortran Applications*, Inria Grenoble Rhône-Alpes, January 2018, n° RR-9147, p. 1-9, https://hal.inria.fr/hal-01688488

Other Publications

[32] G. BERTHOU, A. CARER, H.-P. CHARLES, S. DERRIEN, K. MARQUET, I. MIRO-PANADES, D. PALA, I. PUAUT, F. RASTELLO, T. RISSET, E. ROHOU, G. SALAGNAC, O. SENTIEYS, B. YARAHMADI.*The Inria ZEP project: NVRAM and Harvesting for Zero Power Computations*, March 2018, 1, NVMW 2018 - 10th Annual Non-Volatile Memories Workshop, Poster, https://hal.inria.fr/hal-01941766

[33] P. H. PENNA, M. SOUZA, E. PODESTÁ JÚNIOR, B. NASCIMENTO, M. CASTRO, F. BROQUEDIS, H. FREITAS, J.-F. MÉHAUT.*An OS Service for Transparent Remote Memory Accesses in NoC-Based Lightweight Manycores*, October 2018, 1, NOCS 2018 - 12th IEEE/ACM International Symposium on Networks-on-Chip, Poster [*DOI :* 10.13140/RG.2.2.13022.08000], https://hal.archives-ouvertes.fr/hal-01907003

Project-Team CTRL-A

Control for safe Autonomic computing systems

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

IN PARTNERSHIP WITH:
Institut polytechnique de Grenoble

Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Distributed Systems and middleware

Table of contents

Project-Team CTRL-A

Creation of the Team: 2014 January 01, updated into Project-Team: 2017 June 01

Keywords:

Computer Science and Digital Science:

A1.1.2. - Hardware accelerators (GPGPU, FPGA, etc.)
A1.1.4. - High performance computing
A1.1.9. - Fault tolerant systems
A1.1.10. - Reconfigurable architectures
A1.3. - Distributed Systems
A1.3.5. - Cloud
A1.4. - Ubiquitous Systems
A2.1.9. - Synchronous languages
A2.1.10. - Domain-specific languages
A2.2. - Compilation
A2.3.1. - Embedded systems
A2.5.1. - Software Architecture & Design
A2.5.2. - Component-based Design
A2.5.4. - Software Maintenance & Evolution
A2.6.2. - Middleware
A4.9. - Security supervision
A4.9.1. - Intrusion detection
A4.9.3. - Reaction to attacks
A6.4.2. - Stochastic control

Other Research Topics and Application Domains:

B4.5. - Energy consumption
B5.1. - Factory of the future
B6.1. - Software industry
B6.1.1. - Software engineering
B6.1.2. - Software evolution, maintenance
B6.4. - Internet of things
B6.5. - Information systems
B6.6. - Embedded systems
B8.1. - Smart building/home

1. Team, Visitors, External Collaborators

Research Scientist

Eric Rutten [Team leader, Inria, Researcher, HDR]

Faculty Members

Gwenaël Delaval [Univ Grenoble Alpes, Associate Professor]
Stéphane Mocanu [Grenoble-INP, Associate Professor]

External Collaborator
Bogdan Robu [Univ Grenoble Alpes, Gipsa-lab]

Technical Staff
Soguy Mak Kare Gueye [Inria, Aug. 2018]

PhD Students
Neïl Ayeb [Orange Labs]
Adja Sylla [CEA, until Jan. 2018]

Post-Doctoral Fellow
Quang Pham Tran Anh [Inria/Nokia, co-advised EPI Dyonisos, Rennes, from May 2018]

Administrative Assistant
Maria Immaculada Presseguer [Inria]

2. Overall Objectives

2.1. Objective: control support for autonomic computing

CTRL-A is motivated by today's context where computing systems, large (data centers) or small (embedded), are more and more required to be adaptive to the dynamical fluctuations of their environments and workloads, evolutions of their computing infrastructures (shared, or subject to faults), or changes in application functionalities. Their administration, traditionally managed by human system administrators, needs to be automated in order to be efficient, safe and responsive. Autonomic Computing is the approach that emerged in the early 2000's in distributed systems to answer that challenge, in the form of self-administration control loops. They address objectives like self-configuration (e.g. in service-oriented systems), self-optimization (resource consumption management e.g., energy), self-healing (fault-tolerance, resilience), self-protection (security and privacy).

Therefore, there is a pressing and increasing demand for methods and tools to design controllers for self-adaptive computing systems, that ensure quality and safety of the behavior of the controlled system. The critical importance of the quality of control on performance and safety in automated systems, in computing as elsewhere, calls for a departure from traditional approaches relying on *ad hoc* techniques, often empirical, unsafe and application-specific solutions.

The main objective of the CTRL-A project-team is to develop a novel framework for model-based design of controllers in Autonomic Computing. We want to contribute generic Software Engineering methods and tools for developers to design appropriate controllers for their particular reconfigurable architectures, software or hardware, and integrate them at middleware level. We want to improve concrete usability of techniques from Control Theory, particularly Discrete Event Systems, by specialists of concrete systems (rather than formal models) [9], and to provide tool support for our methods in the form of specification languages and compilers. We address policies for self-configuration, self-optimization (resource management, low power), self-healing (fault tolerance) and self-protection (security).

3. Research Program

3.1. Modeling and control techniques for autonomic computing

The main objective of CTRL-A translates into a number of scientific challenges, the most important of these are:

- (i) programming language support, on the two facets of model-oriented languages, based on automata [6], and of domain specific languages, following e.g., a component-based approach [5], [1] or related to rule-based or HMI languages ;
- (ii) design methods for reconfiguration controller design in computing systems, proposing generic systems architectures and models based on transition systems [3], [8], classical continuous control or controlled stochastic systems.

We adopt a strategy of constant experimental identification of needs and validation of proposals, in application domains like middleware platforms for Cloud systems [7], multi-core HPC architectures [11], Dynamic Partial Reconfiguration in FPGA-based hardware [2] and the IoT and smart environments [4].

Achieving the goals of CTRL-A requires multidisciplinarity and expertise from several domains. The expertise in Autonomic Computing and programming languages is covered internally by members of the Ctrl-A team. On the side of theoretical aspects of control, we have active external collaborations with researchers specialized in Control Theory, in the domain of Discrete Event Systems as well as in classical, continuous control. Additionally, an important requirement for our research to have impact is to have access to concrete, real-world computing systems requiring reconfiguration control. We target autonomic computing at different scales, in embedded systems or in cloud infrastructures, which are traditionally different domains. This is addressed by external collaborations, with experts in either hardware or software platforms, who are generally missing our competences on model-based control of reconfigurations.

4. Application Domains

4.1. Self-adaptive and reconfigurable computing systems in HPC and the IoT

We are attacking the problem of designing well-regulated and efficient self-adaptive computing systems by the development of novel strategies for systems management.

The kind of systems we typically target involve relatively coarse grained computation tasks (e.g. image processing or HPC tasks, components or services), assembled in workflows, application dependency graphs, or composites. At that level, there can be parallel and conditional branches, as well as choices that can be made between alternative branches, corresponding to different ways to perform that part of the application. Such tasks can be achieved following a choice of implementations or versions, such as in service oriented approaches. Each implementation has its own characteristics and requirements, e.g., w.r.t. resources consumed and QoS offered. The systems execution infrastructures present heterogeneity, with different computing processors, a variety of peripheral devices (e.g., I/O, video port, accelerators), and different means of communications. This hardware or middleware level also presents adaptation potential e.g., in varying quantities of resources or sleep and stand-by modes.

The kinds of control problems encountered in these adaptive systems concern the navigation in the configurations space defined by choice points at the levels of applications, tasks, and architecture. Upon events or conditions triggering reconfiguration and adaptation, the controller has to choose a next configuration where, on the one hand, all consistency constraints are satisfied w.r.t. dependencies and resources requirements. On the other hand, it has to apply a policy or strategy deciding between correct configurations which one to chose e.g. by optimizing one or more criteria, or by insuring reachability of some later configuration (goal or fallback). This targeted class of computing systems we consider is mid-sized, in the sense that the combinatorial complexity is large enough for manual solving to be impossible, while remaining within the range where supervisory control techniques are tractable. The pace of control is more sporadic, and slower than the instruction-level computation performance within the coarse-grained tasks.

The objectives of CTRL-A will be achieved and evaluated in both of our main application domains, thereby exhibiting their similarities from the point of view of reconfiguration control. A first application domain is High Performance Computing. In this area, we currently focus especially on the management of Dynamic Partial Reconfiguration in FPGA based hardware, at the level of middleware. Here the particular application we consider is, as in our ANR project HPeC starting end of 2015, video image flow processing for smart cameras implemented on DPR FPGASs themselves embedded in drones.

A second general application domain to confront our models is the Internet of Things (IoT), more specifically self-adaptive middleware platforms for Smart Environments, or Industry 4.0 related topics, like SCADA. We focus on providing coordination components and controllers of software components and services, or rule-based middleware platforms. The adaptation problems concern both the functional aspects of the applications in a smart building, and the middleware support deployment and reconfiguration issues. We are considering perspectives concerning self-protection and security.

5. New Software and Platforms

5.1. Heptagon

KEYWORDS: Compilers - Synchronous Language - Controller synthesis
FUNCTIONAL DESCRIPTION: Heptagon is an experimental language for the implementation of embedded real-time reactive systems. It is developed inside the Synchronics large-scale initiative, in collaboration with Inria Rhones-Alpes. It is essentially a subset of Lucid Synchrone, without type inference, type polymorphism and higher-order. It is thus a Lustre-like language extended with hierchical automata in a form very close to SCADE 6. The intention for making this new language and compiler is to develop new aggressive optimization techniques for sequential C code and compilation methods for generating parallel code for different platforms. This explains much of the simplifications we have made in order to ease the development of compilation techniques.

The current version of the compiler includes the following features: - Inclusion of discrete controller synthesis within the compilation: the language is equipped with a behavioral contract mechanisms, where assumptions can be described, as well as an "enforce" property part. The semantics of this latter is that the property should be enforced by controlling the behaviour of the node equipped with the contract. This property will be enforced by an automatically built controller, which will act on free controllable variables given by the programmer. This extension has been named BZR in previous works. - Expression and compilation of array values with modular memory optimization. The language allows the expression and operations on arrays (access, modification, iterators). With the use of location annotations, the programmer can avoid unnecessary array copies.

- Participants: Adrien Guatto, Brice Gelineau, Cédric Pasteur, Eric Rutten, Gwenaël Delaval, Léonard Gérard and Marc Pouzet
- Partners: UGA - ENS Paris - Inria - LIG
- Contact: Gwenaël Delaval
- URL: http://heptagon.gforge.inria.fr

6. New Results

6.1. Programming support for Autonomic Computing

6.1.1. Reactive languages

Participants: Gwenaël Delaval, Eric Rutten.

Our work in reactive programming for autonomic computing systems is focused on the specification and compilation of declarative control objectives, under the form of contracts, enforced upon classical mode automata as defined in synchronous languages. The compilation involves a phase of Discrete Controller Synthesis, integrating the tool ReaX, in order to obtain an imperative executable code. The programming language Heptagon / BZR (see Section Software and Platforms) integrates our research results [6].

Recent work concerns exploring new possibilities offered by logico-numeric control. We consider Symbolic Limited Lookahead Control for Best-effort Dynamic Computing Resource Management. We put forward a new modeling technique for Dynamic Resource Management (DRM) based on discrete events control for symbolic logico-numerical systems, especially Discrete Controller Synthesis (DCS). The resulting models involve state and input variables defined on an infinite domain (Integers), thereby no exact DCS algorithm exists for safety control. We thus formally define the notion of limited lookahead, and associated best-effort control objectives targeting safety and optimization on a sliding window for a number of steps ahead. We give symbolic algorithms, illustrate our approach on an example model for DRM, and report on performance results based on an implementation in the tool ReaX. This work is in cooperation with the Sumo team at Inria Rennes (Hervé Marchand) and University of Liverpool (Nicolas Berthier), and is published in the WODES 2018 conference [14].

We also have ongoing activities on abstraction methods for compilation using discrete controller synthesis (needed for example, in order to program the controllers for systems where the useful data for control can be of arbitrary types (integer, real, ...) , or also for systems which are naturally distributed, and require a decentralized controller) and on compilation and diagnosis for discrete controller synthesis (which is made special by the declarative nature of the compilation, where it is not easy to precisely diagnose cases where no solution can be found).

On the applicative side, we also consider such modular and logico-numeric approaches for the control of different targets in self-adaptive and reconfigurable systems (see below in Section 6.2.2.1 and 6.2.1.3 [20], [15].

6.1.2. Domain-specific languages

Participants: Gwenaël Delaval, Soguy Mak Kare Gueye, Eric Rutten.

Our work in Domain-specific languages (DSLs) is founded on our work in component-based programming for autonomic computing systems as examplified by e.g., FRACTAL. We consider essentially the problem of specifying the control of components assembly reconfiguration, with an approach based on the integration within such a component-based framework of a reactive language as in Section 6.1.1 [5].

In recent work, we proposed an extension of a classical Software Architecture Description Languages (ADL) with Ctrl-F, DSL for the specification of dynamic reconfiguration behavior in a [1].

Based on this experience, we are working on a proposal for a DSL called Ctrl-DPR, allowing designers to easily generate Autonomic Managers for DPR FPGA systems (see Section 6.2.1.3). Users can describe their system and their management strategies, in terms of the entities composing the system : tasks, versions, applications, ressources, policies. The DSL relies on a behavioral modeling of these entities, targeted at the design of autonomic managers to control the reconfigurations in such a way as to enforce given policies and strategies. These model-based control techniques are embedded in a compiler, connected to the reactive language and discrete controller synthesis tool of Section 6.1.1, which enables to generate a C implementation of the controller enforcing the management strategies. We apply our DSL for the management of a video application on a UAV. This work is in cooperation with LabSticc in Lorient (Jean-Philippe Diguet), and is published in the ICAC 2018 conference [16].

Ongoing work involves a generalization from our experiences in software components, DPR FPGA, as well as Rule-based autonomic manager as in Section 6.1.3. As we observed a similarity in objects and structures (e.g., tasks, implmentation versions, resources, and upper-level application layer), we are considering a more general DSL, which could be specialized towards such different target domains, and where the compilation towards reactive models could be studied and improved, especially considering the features of Section 6.1.1. This direction will also lead us to study the definition of architectural patterns for multiple loop Autonomic Managers, particularly hierarchical, with lower layers autonomy alleviating management burden from the upper layers.

6.1.3. Rule-based systems

Participants: Adja Sylla, Gwenaël Delaval, Eric Rutten.

This work concerns a high-level language for safe rule-based programming in the LINC transactional rule-based platform developed at CEA [10]. Rule based middlewares such as LINC enable high level programming of distributed adaptive systems behaviours. LINC also provides the systems with transactional guarantees and hence ensures their reliability at runtime. However, the set of rules may contain design errors (e.g. conflicts, violations of constraints) that can bring the system in unsafe safe or undesirables states, despite the guarantees provided by LINC. On the other hand, automata based languages such as Heptagon/BZR enable formal verification and especially synthesis of discrete controllers to deal with design errors. Our work studies these two languages and combines their execution mechanisms, from a technical perspective. We target applications to the domain of Internet of Things and more particularly smart building, office or home (see Section 6.2.2.1).

This work is in cooperation with CEA LETI/DACLE (Maxime Louvel), it was the topic of the PhD of Adja Sylla at CEA, co-advised with M. Louvel, and aspects on applications of logico-numeric control are published in the CCTA 2018 conference [20].

6.2. Design methods for reconfiguration controller design in computing systems

We apply the results of the previous axes of the team's activity, as well as other control techniques, to a range of infrastructures of different natures, but sharing a transversal problem of reconfiguration control design. From this very diversity of validations and experiences, we draw a synthesis of the whole approach, towards a general view of Feedback Control as MAPE-K loop in Autonomic Computing [23] [9].

6.2.1. High-Performance Computing

Participants: Agustin Yabo, Soguy Mak Kare Gueye, Gwenaël Delaval, Stéphane Mocanu, Bogdan Robu, Eric Rutten.

6.2.1.1. Automated regulation and software transactional memory

A parallel program needs to manage the trade-off between the time spent in synchronisation and computation. This trade-off is significantly affected by its parallelism degree. A high parallelism degree may decrease computing time while increasing synchronisation cost. We performed work on dynamic control of thread parallelism and mapping. We address concurrency issues via Software Transactional Memory (STM). We implement feedback control loops to automate management of threads and diminish program execution time.

This work was performed in the framework on the PhD of Naweiluo Zhou, and published in the journal on Concurrency and Computation: Practice and Experience [13].

6.2.1.2. A Control-Theory based approach to minimize cluster underuse

HPC systems are facing more and more variability in their behavior, related to e.g., performance and power consumption, and the fact that they are less predictable requires more runtime management. One such problem is found in the context of CiGri, a simple, lightweight, scalable and fault tolerant grid system which exploits the unused resources of a set of computing clusters. This work resulted in first results addressing the problem of automated resource management in an HPC infrastructure, using techniques from Control Theory to design a controller that maximizes cluster utilization while avoiding overload. We put in place a mechanism for feedback (Proportional Integral, PI) control system software, through a maximum number of jobs to be sent to the cluster, in response to system information about the current number of jobs processed. Additionally, we developed a Model-Predictive Controller to improve the performance of the system.

This work is done in cooperation with the Datamove team of Inria/LIG, and Gipsa-lab. It was the topic of the Master's thesis of Agustin Yabo [25]. Preliminary results were published in the AIScience workshop (Autonomous Infrastructure for Science) of the HPDC conference [19].

6.2.1.3. Reconfiguration control in DPR FPGA

6.2.1.3.1. DPR FPGA and discrete control for reconfiguration

Implementing self-adaptive embedded systems, such as UAV drones, involves an offline provisioning of the several implementations of the embedded functionalities with different characteristics in resource usage and performance in order for the system to dynamically adapt itself under uncertainties. We propose an autonomic control architecture for self-adaptive and self-reconfigurable FPGA-based embedded systems. The control architecture is structured in three layers: a mission manager, a reconfiguration manager and a scheduling manager. In this work we focus on the design of the reconfiguration manager. We propose a design approach using automata-based discrete control. It involves reactive programming that provides formal semantics, and discrete controller synthesis from declarative objectives.

This work is in the framework of the ANR project HPeC (see Section 8.2.1), and is published in the International Workshop on High Performance and Dynamic Reconfigurable Systems and Networks (DRSN 2018), part of the HPCS 2018 conference [17] ; for the evaluation of the apllication of logico-numeric control, in the CCTA 18 conference [15] ; for the proposal of a Domain Specific Language, in the ICAC 2018 conference [16].

6.2.1.3.2. Mission management and stochastic control

In the Mission Management workpackage of the ANR project HPeC, a concurrent control methodology is constructed for the optimal mission planning of a U.A.V. in stochastic environnement. The control approach is based on parallel ressource sharing Partially Observable Markov Decision Processes modeling of the mission. The parallel POMDP are reduced to discrete Markov Decision Models using Bayesian Networks evidence for state identification. The control synthesis is an iterative two step procedure : first MDP are solved for the optimisation of a finite horizon cost problem ; then the possible ressource conflicts between parallel actions are solved either by a priority policy or by a QoS degradation of actions, e.g., like using a lower resolution version of the image processing task if the ressource availability is critical.

This work was performed in the framework on the PhD of Chabha Hireche, and published in the journal on Sensors [24], [12].

6.2.2. IoT

Participants: Neïl Ayeb, Adja Sylla, Gwenaël Delaval, Stéphane Mocanu, Eric Rutten.

6.2.2.1. Control of smart buildings

A smart environment is equipped with numerous devices (i.e., sensors, actuators) that are possibly distributed over different locations (e.g., rooms of a smart building). These devices are automatically controlled to achieve different objectives related, for instance, to comfort, security and energy savings. Our work proposes a design framework based on the combination of the rule based middleware LINC and the automata based language Heptagon/BZR (H/BZR). It consists of: an abstraction layer for the heterogeneity of devices, a transactional execution mechanism to avoid inconsistencies and a controller that, based on a generic model of the environment, makes appropriate decisions and avoids conflicts. A case study with concrete devices, in the field of building automation, is presented to illustrate the framework.

This work is in the framework of the cooperation with CEA (see Section 7.1), and is published in the CCTA 2018 conference [20].

6.2.2.2. Device management

The research topic is targeting an adaptative and decentralized management for the IoT. It will contribute design methods for processes in virtualized gateways in order to enhance IoT infrastructures. More precisely, it concerns Device Management in the case of large numbers of connected sensors and actuators, as can be found in Smart Home and Building, Smart Electricity grids, and industrial frameworks as in Industry 4.0.

This work is in the framework of the Inria/Orange labs joint laboratory (see Section 7.2.1), and supported by the CIFRE PhD thesis grant of Neïl Ayeb, starting dec. 2017.

6.2.2.3. Security in SCADA industrial systems

We focus mainly on vulnerability search, automatic attack vectors synthesis and intrusion detection. Model checking techniques are used for vulnerability search and automatic attack vectors construction. Intrusion detection is mainly based on process-oriented detection with a technical approach from run-time monitoring. The LTL formalism is used to express safety properties which are mined on an attack-free dataset. The resulting monitors are used for fast intrusion detections.

A demonstrator of attack/defense scenario in SCADA systems will be built on the existing G-ICS lab (hosted by ENSE3/Grenoble-INP).

This work is in the framework of the ANR project Sacade on cybersecurity of industrial systems (see Section 8.2.2) [18] [22] [21].

The work is also supported by Grenoble Alpes Cybersecurity Institute (see Section 8.1.1).

Ongoing work concerns the complementary topic of analysis and identification of reaction mechanisms for self-protection in cybersecurity, where, beyond classical defense mechanisms that detect intrusions and attacks or assess the kind of danger that is caused by them, we explore models and control techniques for the automated reaction to attacks, in order to use detection information to take the appropriate defense and repair actions.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

Our cooperation with CEA (an EPIC, industrial and commercial public institution) concerns the LETI/LIST DACLE laboratory at Grenoble Minatec; it is bilateral, involving the CEA PhD grant of Adja Sylla (finished end of january 2018), to work with F. Pacull and M. Louvel on high-level programming on top of a rule-based middleware (See Sections 6.1.3 and 6.2.2.1).

7.2. Bilateral Grants with Industry

7.2.1. Orange

We have a cooperation with Orange labs, around a CIFRE PhD grant, on the topic of autonomic device management (see Section 6.2.2.2). This activity is part of the Inria/Orange joint laboratory.

7.2.2. Nokia / Bell labs

We are starting a research action with Nokia / Bell labs, around a post-doctorate, co-advised with project-team Dyonisos at Inria Rennes, on the topic of the integration of FPGA-based accelerators in network nodes, and their reconfiguration management in coordination with higher level Software Defined Networks management. This activity is part of the Inria/ Nokia / Bell labs joint laboratory, and is in cooperation with the Dyonisos EPI at Inria Rennes Bretagne Atlantique (Yassine Hadjhadj), and the post-doctorate topic of Quang Pham Tran Anh.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. Grenoble Alpes Cybersecurity Institute Cross-Disciplinary Project of the Idex

The Grenoble Alpes Cybersecurity Institute aims at undertaking ground-breaking interdisciplinary research in order to address cybersecurity and privacy challenges. Our main technical focus is on low-cost secure elements, critical infrastructures, vulnerability analysis and validation of large systems, including practical resilience across the industry and the society.

In Ctrl-A, it is funding an internship position followed by a PhD position to be provided in September 2019 and supervised by Stephane Mocanu.

8.2. National Initiatives

8.2.1. ANR HPeC

HPeC is an ANR project on Self-Adaptive, Energy Efficient High Performance Embedded Computing, with a UAV case study (http://hpec.fr/). The Coordinator is Lab-STICC / MOCS (Lorient / Brest), and the duration: 42 month from october 2015. Others Partners are: UBO, U. Clermont-Ferrand, InPixal.

In Ctrl-A, it is funding a post-doc position, hired in Grenoble and co-advised with Lorient : Soguy Gueye. The work will be continued with a post-doc hired in Lorient : Erwan Moreac. A PhD based in Brest, Chabha Hireche, is co-advised by Stéphane Mocanu.

8.2.2. ANR Sacade

The ANR ASTRID Sacade project is funded by DGA. Stéphane Mocanu is in charge of several workpackages including a demonstrator. An expert engineer position is funded for the implementation of attack/defense scenarios in SCADA.

8.2.3. Informal National Partners

We have contacts with colleagues in France, in addition to the cooperation mentioned before, and with whom we are submitting collaboration projects, co-organizing events and workshops, etc. They feature : Avalon Inria team in Lyon (Ch. Perez, L. Lefevre, E. Caron), LIP6 (J. Malenfant), Scales Inria team in Sophia-Antipolis (L. Henrio), LIRRM in Montpellier (A. Gamatié, K. Godary, D. Simon), IRISA/Inria Rennes (J. Buisson, J.L. Pazat, ...), Telecom Paris-Tech (A. Diaconescu, E. Najm), LAAS (Thierry Monteil), LURPA ENS Cachan (J.M. Faure, J.J. Lesage).

8.2.4. Informal National Industrial Partners

We have ongoing discussions with several industrial actors in our application domains, some of them in the framework of cooperation contracts, other more informal: Eolas/Business decision (G. Dulac, I. Saffiedine), ST Microelectronics (V. Bertin), Schneider Electric (C. El-Kaed, P. Nappey, M. Pitel).

8.3. International Initiatives

8.3.1. Inria International Labs

We participate in the jLESC, Joint Laboratory for Extreme Scale Computing, with partners Inria, the University of Illinois, Argonne National Laboratory, Barcelona Supercomputing Center, Jülich Supercomputing Centre and RIKEN AICS.

We participated to the 7th Workshop of the JLESC at Urbana-Champain in July 2017.

We started a cooperation with Argonne National Labs, on Improving the performance and energy efficiency of HPC applications using autonomic computing techniques.

https://jlesc.github.io/projects/energy_autonomic/

We are also exploring possibilities on the topic of integrating FPGAs in HPC grids, with a participation in a workshop at FPT 18.

https://collab.cels.anl.gov/display/HPCFPGA/HPC-FPGA

8.3.2. Inria International Partners

8.3.2.1. Informal International Partners

We have ongoing relations with international colleagues in the emerging community on our topic of control for computing e.g., in Sweden at Lund (K.E. Arzen, M. Maggio), Mälardalen (A. Papadopoulos) and Linnaeus Universities (D. Weyns, N. Khakpour), in the Netherlands at CWI/leiden University (F. Arbab), in the U.K. at Liverpool U. (N. Berthier), in China at Heifei University (Xin An), in Italy at University Milano (C. Ghezzi, A. Leva), in the USA at Ann Arbor University (S. Lafortune) and UMass (P. Shenoy, E. Cecchet).

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

Eric Rutten is co-chairing, with Bogdan Robu (Gipsa-lab), the 40th Summer School of Automatic Control, Grenoble, September 2019, on the special topic of Control for Computing Systems.

9.1.2. Scientific Events Selection

9.1.2.1. Chair of Conference Program Committees

Eric Rutten is co-chair, with A. Filieri (Imp. Coll. UK), of the International Workshop on Autonomic High Performance Computing,(AHPC 2018) (http://hpcs2018.cisedu.info/2-conference/workshops—hpcs2018/workshop09-ahpc) part of The International Conference on High Performance Computing & Simulation (HPCS 2018), July 16 - 20, 2018,Orléans, France (http://hpcs2018.cisedu.info/)

9.1.2.2. Member of the Conference Program Committees

Eric Rutten is PC member for :

- international conferences
  - ICAC 2018 (15th IEEE International Conference on Autonomic Computing), Sept 3-7, 2018, Trento, Italy (http://icac2018.informatik.uni-wuerzburg.de/)
  - (Associate Editor) 2nd IEEE Conference on Control Technology and Applications, CCTA 18, Copenhagen, Denmark, August 21-24, 2018 (http://ccta2018.ieeecss.org/)
  - 14th Workshop on Discrete Event Systems, WODES'18, Sorrento Coast, Italy, May 30 - June 1, 2018. (http://wodes2018.unisa.it/)
  - AI-Science'18, workshop on autonomic cyberinfrastructure for science, in conjunction with the ACM HPDC 2018, Tempe, AZ, United States, June 11, 2018 (http://www.hpdc.org/2018/)
  - 16th High Performance Computing & Simulation Conference (HPCS 2018). July 16 ? 20, 2018,Orléans, France (http://hpcs2018.cisedu.info/)
  - UCC Cloud Challenge event in conjunction with IEEE/ACM UCC/BDCAT 2018 in Zurich (http://events.cbmi.htw-berlin.de/ucc18-cloudChallenge/index.html)
  - ICAC 2019 (16th IEEE International Conference on Autonomic Computing), Umeøa, Sweden, June 16 ? 20, 2 019 (http://icac2019.cs.umu.se)

9.1.2.3. Reviewer

Eric Rutten is reviewer for SafeProcess 18 and CDC 18.

Stéphane Mocanu is reviewer for WODES'18.

9.1.3. Journal

9.1.3.1. Reviewer - Reviewing Activities

Eric Rutten is reviewer for

- Science of Computer Programming
- IEEE TSC Transactions on Services Computing
- IEEE Transactions on Industrial Informatics

9.1.4. Invited Talks

Eric Rutten was invited to give a talk at :

- FETCH 2018, Ecole d'Hiver francophone sur la technologie de conception des systèmes embarqués hétérogènes, 24-26 Janvier 2018, Saint Malo http://fetch-conference.org
- with J.Ph. Diguet, : WRC: 12th Workshop on Reconfigurable Computing at HiPeac 2018, https://www.hipeac.net/2018/manchester/#/schedule/ Jan. 24, Manchester U.K.
- International Workshop on High Performance and Dynamic Reconfigurable Systems and Networks (DRSN 2018) part of The International Conference on High Performance Computing & Simulation (HPCS 2018), July 16 ? 20, 2018, Orléans, France http://hpcs2018.cisedu.info/
- FETCH 2019, Ecole d'Hiver francophone sur la technologie de conception des systèmes embarqués hétérogènes, Louvain-la-Neuve, Belgique, 28-30 janvier 2019 http://fetch-conference.org

9.1.5. Leadership within the Scientific Community

Eric Rutten is co-chair of the Technical Committee on Discrete Event Systems (DESTC), a part of the IEEE Control Systems Society (CSS) Technical Activities Board (http://discrete-event-systems.ieeecss.org/tc-discrete/home) ; and member of the IFAC Technical Committee 1.3 on Discrete Event and Hybrid Systems, for the 2017-2020 triennum (http://tc.ifac-control.org/1/3).

9.1.6. Research Administration

Eric Rutten is member of the LIG laboratory council, and in charge of scientific relations between Inria Grenoble Rhône-Alpes and CEA Tech.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence : G. Delaval, Algorithmics and imperative programming, 18h class, 18h lab, L2, Université Grenoble Alpes

Licence : G. Delaval, Basis of software development : modularity, tests, 15h class, 15h lab, L2, Université Grenoble Alpes

Master : G. Delaval, Programming languages and compiler design, 33h, M1, Université Grenoble Alpes

Master : S. Mocanu, Computer Networks and Cybersecurity, 16h class, 34h lab, M1, Grenoble-INP/ENSE3

Master : S. Mocanu, Industriel Computer Networks, 8h class, 8h lab, niveau (M1, M2), M2, Grenoble-INP/ENSE3

Master : S. Mocanu, Reliability, 10h class, 8h lab, M2, Grenoble-INP/ENSE3

Master : S. Mocanu, Intrusion Detection and Defense in Depth labs, niveau M2, Grenoble-ENSE3/ENSIMAG

9.2.2. Supervision

- PhD : Oualid Koucham ; Détection d'intrusions dans les systèmes de contrôle industriels ; 12 nov. 2018 ; co-advised by S. Mocanu with J-M Thiriet (Gipsa-lab).
- PhD in progress : Chabha Hireche ; Etude et implémentation d'une approche probabiliste de contrôle de mission de drone autonome ; oct. 2015 ; co-advised by S. Mocanu with Catherine Dezan (U. Bretagne Occidentale), and Jean-Philippe Diguet (U. Bretagne Sud).
- PhD in progress : Neïl Ayeb ; Vers un management des objets de l'IoT décentralisé et adaptable ; dec. 2017 ; co-advised by Eric Rutten with S. Bolle, T. Coupaye (Orange labs).

9.3. Popularization

9.3.1. Articles and contents

The Ctrl-A team is featured in a special issue on the numeric world in the "*Dauhiné Libéré des enfants*" (nov.-dec. 2018).

9.3.2. Internal action

The Ctrl-A team participated in the event "mon équipe en 180 secondes" at Inria Montbonnot.

10. Bibliography

Major publications by the team in recent years

[1] F. ALVARES, E. RUTTEN, L. SEINTURIER.*A Domain-specific Language for The Control of Self-adaptive Component-based Architecture*, in "Journal of Systems and Software", January 2017, https://hal.archives-ouvertes.fr/hal-01450517

[2] X. AN, E. RUTTEN, J.-P. DIGUET, A. GAMATIÉ.*Model-based design of correct controllers for dynamically reconfigurable architectures*, in "ACM Transactions on Embedded Computing Systems (TECS)", February 2016, vol. 15, n^{o} 3, https://hal.inria.fr/hal-01272077

[3] N. BERTHIER, E. RUTTEN, N. DE PALMA, S. M.-K. GUEYE.*Designing Autonomic Management Systems by using Reactive Control Techniques*, in "IEEE Transactions on Software Engineering", July 2016, vol. 42, n^{o} 7, 18, https://hal.inria.fr/hal-01242853

[4] J. CANO, G. DELAVAL, E. RUTTEN.*Coordination of ECA rules by verification and control*, in "16th International Conference on Coordination Models and Languages", Berlin, Germany, June 2014, 16 p., https://hal.archives-ouvertes.fr/hal-01006186

[5] G. DELAVAL, E. RUTTEN.*Reactive model-based control of reconfiguration in the Fractal component-based model*, in "Proceedings of the 13th International Symposium on Component Based Software Engineering (CBSE), Prague, Czech Republic, 23-25 June", 2010, p. 93–112, best paper award, http://dx.doi.org/10.1007/978-3-642-13238-4_6

[6] G. DELAVAL, E. RUTTEN, H. MARCHAND.*Integrating Discrete Controller Synthesis in a Reactive Programming Language Compiler*, in "journal of Discrete Event Dynamic System, jDEDS, special issue on Modeling of Reactive Systems", 2013, vol. 23, n^{o} 4, p. 385-418, http://dx.doi.org/10.1007/s10626-013-0163-5

[7] S. M.-K. GUEYE, N. DE PALMA, E. RUTTEN, A. TCHANA, N. BERTHIER.*Coordinating self-sizing and self-repair managers for multi-tier systems*, in "Future Generation Computer Systems", June 2014, vol. 35, p. 14 - 26 [*DOI :* 10.1016/J.FUTURE.2013.12.037], https://hal.inria.fr/hal-00949556

[8] N. KHAKPOUR, F. ARBAB, E. RUTTEN.*Synthesizing structural and behavioral control for reconfigurations in component-based systems*, in "Formal Aspects of Computing", December 2015 [*DOI :* 10.1007/S00165-015-0346-Y], https://hal.inria.fr/hal-01247524

[9] M. LITOIU, M. SHAW, G. TAMURA, N. M. VILLEGAS, H. MÜLLER, H. GIESE, R. ROUVOY, E. RUTTEN.*What Can Control Theory Teach Us About Assurances in Self-Adaptive Software Systems?*, in "Software Engineering for Self-Adaptive Systems 3: Assurances", R. DE LEMOS, D. GARLAN, C. GHEZZI, H. GIESE (editors), LNCS, Springer, May 2017, vol. 9640, https://hal.inria.fr/hal-01281063

[10] M. LOUVEL, F. PACULL, E. RUTTEN, A. N. SYLLA.*Development Tools for Rule-Based Coordination Programming in LINC*, in "19th International Conference on Coordination Languages and Models (COORDINATION)", Neuchâtel, Switzerland, J.-M. JACQUET, M. MASSINK (editors), Coordination Models and Languages, Springer International Publishing, June 2017, vol. LNCS-10319, p. 78-96, Part 2: Languages and Tools [*DOI :* 10.1007/978-3-319-59746-1_5], https://hal-cea.archives-ouvertes.fr/cea-01531019

[11] N. ZHOU, G. DELAVAL, B. ROBU, E. RUTTEN, J.-F. MÉHAUT.*Autonomic Parallelism and Thread Mapping Control on Software Transactional Memory*, in "13th IEEE International Conference on Autonomic Computing (ICAC 2016)", Würzburg, Germany, July 2016, https://hal.archives-ouvertes.fr/hal-01309681

Publications of the year

Articles in International Peer-Reviewed Journal

[12] C. HIRECHE, C. DEZAN, S. MOCANU, D. HELLER, J.-P. DIGUET.*Context/Resource-Aware Mission Planning Based on BNs and Concurrent MDPs for Autonomous UAVs*, in "Sensors", December 2018, vol. 18, n^o^ 4266, p. 1-28 [*DOI :* 10.3390/S18124266], https://hal.univ-brest.fr/hal-01944680

[13] N. ZHOU, G. DELAVAL, B. ROBU, E. RUTTEN, J.-F. MÉHAUT.*An Autonomic-Computing Approach on Mapping Threads to Multi-cores for Software Transactional Memory*, in "Concurrency and Computation: Practice and Experience", September 2018, vol. 30, n^o^ 18, e4506 [*DOI :* 10.1002/CPE.4506], https://hal.archives-ouvertes.fr/hal-01742690

International Conferences with Proceedings

[14] N. BERTHIER, H. MARCHAND, E. RUTTEN.*Symbolic Limited Lookahead Control for Best-effort Dynamic Computing Resource Management*, in "WODES 2018 - 14th Workshop on Discrete Event Systems", Sorrento Coast, Italy, Elsevier, May 2018, p. 1-8 [*DOI :* 10.1016/J.IFACOL.2018.06.288], https://hal.inria.fr/hal-01807284

[15] S. M.-K. GUEYE, G. DELAVAL, E. RUTTEN, J.-P. DIGUET.*Discrete and Logico-numerical Control for Dynamic Partial Reconfigurable FPGA-based Embedded Systems : a Case Study*, in "CCTA 2018 - 2nd IEEE Conference on Control Technology and Applications", Copenhaguen, Denmark, 2nd IEEE Conference on Control Technology and Applications CCTA, August 2018, p. 1480-1487, https://hal.archives-ouvertes.fr/hal-01862619

[16] S. M.-K. GUEYE, G. DELAVAL, E. RUTTEN, D. HELLER, J.-P. DIGUET.*A Domain-specific Language for Autonomic Managers in FPGA Reconfigurable Architectures*, in "ICAC 2018 - 15th IEEE International Conference on Autonomic Computing", Trento, Italy, IEEE, September 2018, p. 1-10, https://hal.archives-ouvertes.fr/hal-01868675

[17] S. M. K. GUEYE, E. RUTTEN, J.-P. DIGUET.*Autonomic Management of Reconfigurations in DPR FPGA-based Embedded System: INVITED TALK EXTENDED ABSTRACT*, in "International Workshop on High Performance and Dynamic Reconfigurable Systems and Networks (DRSN 2018), part of HPCS 2018 - International Conference on High Performance Computing & Simulation", Orléans, France, International Workshop on High Performance and Dynamic Reconfigurable Systems and Networks (DRSN 2018), part of the 2018 International Conference on High Performance Computing & Simulation, HPCS, July 2018, p. 557 - 558 [*DOI :* 10.1109/HPCS.2018.00093], https://hal.archives-ouvertes.fr/hal-01862631

[18] O. KOUCHAM, S. MOCANU, G. HIET, J.-M. THIRIET, F. MAJORCZYK.*Efficient Mining of Temporal Safety Properties for Intrusion Detection in Industrial Control Systems*, in "SAFEPROCESS 2018 - 10th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes", Warsaw, Poland, August 2018, p. 1-8, https://hal.archives-ouvertes.fr/hal-01877109

[19] E. STAHL, A. G. YABO, O. RICHARD, B. BZEZNIK, B. ROBU, E. RUTTEN.*Towards a control-theory approach for minimizing unused grid resources*, in "AI-Science'18 - workshop on Autonomous Infrastructure for Science, in conjunction with the ACM HPDC 2018", Tempe, AZ, United States, June 2018, p. 1-8 [*DOI :* 10.1145/3217197.3217201], https://hal.archives-ouvertes.fr/hal-01823787

[20] A. N. SYLLA, M. LOUVEL, E. RUTTEN, G. DELAVAL.*Modular and Hierarchical Discrete Control for Applications and Middleware Deployment in IoT and Smart Buildings*, in "CCTA 2018 - 2nd IEEE Conference on Control Technology and Applications", Copenhagen, Denmark, 2nd IEEE Conference on Control Technology and Applications CCTA, August 2018, p. 1472–1479, https://hal.archives-ouvertes.fr/hal-01862608

[21] J.-M. THIRIET, S. MOCANU.*Some Considerations on Dependability Issues and Cyber-Security of Cyber-Physical Systems*, in "The 7 th IEEE International Conference on Smart Communications in Network Technologies (SACONET'18)", El Oued, Algeria, France, October 2018, https://hal.archives-ouvertes.fr/hal-01909025

National Conferences with Proceeding

[22] S. MOCANU.*Formation cybersécurité des systèmes industriels pour les ingénieurs non-informaticiens*, in "RESSI 2018 - Rendez-Vous de la Recherche et de l'Enseignement de la Sécurité des Systèmes d'Information", Nancy / La Bresse, France, May 2018, p. 1-3, https://hal.archives-ouvertes.fr/hal-01908938

Scientific Books (or Scientific Book chapters)

[23] E. RUTTEN, N. MARCHAND, D. SIMON.*Feedback Control as MAPE-K loop in Autonomic Computing*, in "Software Engineering for Self-Adaptive Systems III. Assurances", Lecture Notes in Computer Science, Springer, January 2018, vol. 9640, p. 349-373 [*DOI :* 10.1007/978-3-319-74183-3_12], https://hal.inria.fr/hal-01285014

Other Publications

[24] C. HIRECHE, C. DEZAN, J.-P. DIGUET, S. MOCANU.*Planification de Mission de Drone: Implémentation Logicielle/Matérielle*, June 2018, GDR SoC2, Poster, https://hal.univ-brest.fr/hal-01844331

References in notes

[25] A. YABO.*A control-theory approach for cluster autonomic management: maximizing usage while avoiding overload* , Masters Thesis, M2R MISCIT, Université Grenoble Alpes, France, 2018

Project-Team DANTE

Dynamic Networks : Temporal and Structural Capture Approach

IN COLLABORATION WITH: Laboratoire de l'Informatique du Parallélisme (LIP)

IN PARTNERSHIP WITH:
Ecole normale supérieure de Lyon

Université Claude Bernard (Lyon 1)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Networks and Telecommunications

Table of contents

Project-Team DANTE

Creation of the Team: 2012 November 01, updated into Project-Team: 2015 January 01

Keywords:

Computer Science and Digital Science:

A1.2. - Networks
A1.2.4. - QoS, performance evaluation
A1.2.5. - Internet of things
A1.2.6. - Sensor networks
A1.2.9. - Social Networks
A3.4.1. - Supervised learning
A3.5. - Social networks
A3.5.1. - Analysis of large graphs
A5.9. - Signal processing
A5.9.4. - Signal processing over graphs
A8.1. - Discrete mathematics, combinatorics
A8.7. - Graph theory
A8.8. - Network science
A8.9. - Performance evaluation

Other Research Topics and Application Domains:

B2.3. - Epidemiology
B6. - IT and telecom
B6.3.4. - Social Networks
B6.4. - Internet of things
B9.5.1. - Computer science
B9.5.6. - Data science
B9.6.5. - Sociology
B9.6.8. - Linguistics
B9.6.10. - Digital humanities

1. Team, Visitors, External Collaborators

Research Scientists

Paulo Gonçalves [Team leader, Inria, Senior Researcher, HDR]
Philippe Nain [Inria, Senior Researcher, HDR]
Tommaso Venturini [Inria, Advanced Research Position, until Aug 2018, HDR]

Faculty Members

Thomas Begin [Univ de Claude Bernard, Associate Professor, HDR]
Anthony Busson [Univ de Claude Bernard, Professor, HDR]
Christophe Crespelle [Univ de Claude Bernard, Associate Professor, HDR]
Éric Fleury [Team leader, Ecole Normale Supérieure Lyon, Chair, until Mar 2018, HDR]
Marion Foare [Associate Professor, Ecole supérieure de chimie physique électronique de Lyon, since Sept. 2018]

Isabelle Guérin Lassous [Univ de Claude Bernard, Professor, HDR]
Márton Karsai [Ecole Normale Supérieure Lyon, Associate Professor]
Jean-Pierre Chevrot [Professor Univ Grenoble Alpes, ext. collaborator until Aug 2018]
Éric Philippe Guichard [Associate Professor Enssib, Lyon, ext. collaborator until march 2020]
Jean-Philippe Magué [Associate Professor Ecole Normale Supérieure Lyon, ext. collaborator until Aug 2018]

PhD Students

Lafdal Abdelwedoud [Bourse du gouvernement Français pour étudiants étrangers]
Mohammed Amer [Ministère de défense Algérien, until Sep 2018]
Dominique Barbe [Univ de Lyon, from Sep 2018]
Esteban Bautista Ruiz [Bourse CONACyT, Méxique]
Nour El Houda Bouzouita [Ecole Normale Supérieure Lyon, from Nov 2018]
Sicheng Dai [Univ de Lyon]
Gaetan Frusque [Ecole Normale Supérieure Lyon]
Remy Grunblatt [Inria]
Jacob Levy Abitbol [Inria]
Marija Stojanova [Ecole Normale Supérieure Lyon]
Samuel Unicomb [Inria]

Post-Doctoral Fellows

Sarah de Nigris [CNRS, until Jun 2018]
Sebastien Lerique [Inria]
Mikhail Tsitsvero [Univ de Lyon, Labex MILyon]

Visiting Scientists

Dorsaf Ghozlani [Innov'Com, Tunise, from Apr 2018 until Jul 2018]
Donald Towsley [U. of Massachusetts Amherst, US, from Mar 2018 until Jun 2018]
Alexandre Brandwajn [U. of California, Santa Cruz, US, from Feb 2018 until Mar 2018]

Administrative Assistant

Laetitia Gauthe [Inria, until Nov 2018]

2. Overall Objectives

2.1. Overall Objectives

The goal of DANTE is to develop **novel models, algorithms and methods to analyse the dynamics of large-scale networks**, (*e.g. social networks, technological networks such as the Web and hyperlinks, Articles and co-citation, email exchanges, economic relations, bacteria/virus propagation in human networks...*). Large datasets describing such networks are nowadays more "accessible" due to the emergence of online activities and new techniques of data collection. These advantages provide us an unprecedented avalanche of large data sets, recording the digital footprints of millions of entities (*e.g.* individuals, computers, documents, stocks, etc.) and their temporal interactions [0]. Such large amount of information allows for easier and more precise traceability of social activities, better observation of the structural and temporal evolution of social/technological/economical networks, the emergence of their localized and cascading failures, and provides information about the general roles of self-organization in an interdisciplinary sense. All these questions represent a major scientific, economic, and social challenge, which has the potential to revolutionize our understanding of the arising socio-technical world of our age.

[0] YouTube claims to receive 48 hours of video every minute, Google and Facebook represent major world companies that generate millions of traces on our activities every second. Every day, hundreds of millions of posts are added to the blogosphere, from which information on citizen opinions and their evolutions can be collected.

Our main challenge is to propose **generic methodologies and concepts to develop relevant formal tools to model, analyse the dynamics and evolution of such networks, that is, to formalise the dynamic properties of both structural and temporal interactions of network entities/relations**:

- **Ask** application domains relevant questions, to learn something new about such domains instead of merely playing with powerful computers on huge data sets.
- **Access** and collect data with adapted and efficient tools. This includes a reflexive step on the biases of the data collected and their relations to real activities/application domain.
- **Model** the dynamics of networks by analyzing their structural and temporal properties jointly, inventing original approaches combining graph theory with signal processing. A key point is to capture temporal features in the data, which may reveal meaningful insights on the evolution of the networks.
- **Interpret** the results, make the knowledge robust and useful in order to be able to control, optimise and (re)-act on the network structure itself and on the protocols exchange/interactions in order to obtain a better performance of the global system.

The challenge is to solve a major scientific puzzle, common to several application domains (*e.g.*, sociology, information technology, epidemiology) and central in network science: how to understand the causality between the evolution of macro-structures and individuals, at local and global scales?

3. Research Program

3.1. Graph-based signal processing

Participants: Paulo Gonçalves, Éric Fleury, Márton Karsai, Marion Foare, Thomas Begin.

Evolving networks can be regarded as "*out of equilibrium*" systems. Indeed, their dynamics are typically characterized by non standard and intricate statistical properties, such as non-stationarity, long range memory effects, intricate space and time correlations.

Analyzing, modeling, and even defining adapted concepts for dynamic graphs is at the heart of DANTE. This is a largely open question that has to be answered by keeping a balance between specificity (solutions triggered by specific data sets) and generality (universal approaches disconnected from social realities). We will tackle this challenge from a graph-based signal processing perspective involving signal analysts and computer scientists, together with experts of the data domain application. One can distinguish two different issues in this challenge, one related to the graph-based organization of the data and the other to the time dependency that naturally exits in the dynamic graph object. In both cases, a number of contributions can be found in the literature, albeit in different contexts. In our application domain, high-dimensional data "naturally reside" on the vertices of weighted graphs. The emerging field of signal processing on graphs merges algebraic and spectral graph theoretic concepts with computational harmonic analysis to process such signals on graphs [76].

As for the first point, adapting well-founded signal processing techniques to data represented as graphs is an emerging, yet quickly developing field which has already received key contributions. Some of them are very general and delineate ambitious programs aimed at defining universal, generally unsupervised methods for exploring high-dimensional data sets and processing them. This is the case for instance of the “diffusion wavelets” and “diffusion maps” pushed forward at Yale and Duke [58]. Others are more traditionally connected with standard signal processing concepts, in the spirit of elaborating new methodologies via some bridging between networks and time series, see for instance [71] and references therein. Other viewpoints can be found as well, including multi-resolution Markov models [79], Bayesian networks or distributed processing over sensor networks [70]. Such approaches can be particularly successful for handling static graphs and unveiling aspects of their organization in terms of dependencies between nodes, grouping, etc. Incorporating possible time dependencies within the whole picture calls however for the addition of an extra dimension to the problem "as it would be the case when switching from one image to a video sequence", a situation for which one can imagine to take advantage of the whole body of knowledge attached to non-stationary signal processing [59].

3.2. Theory and Structure of dynamic Networks

Participants: Christophe Crespelle, Éric Fleury, Anthony Busson, Márton Karsai, Jean-Philippe Magué, Éric Guichard, Jean-Pierre Chevrot, Tommaso Venturini.

Characterization of the dynamics of complex networks. We need to focus on intrinsic properties of evolving/dynamic complex networks. New notions (as opposed to classical static graph properties) have to be introduced: rate of vertices or links appearances or disappearances, the duration of link presences or absences. Moreover, more specific properties related to the dynamics have to be defined and are somehow related to the way to model a dynamic graph.

Through the systematic analysis and characterization of static network representations of many different systems, researchers of several disciplines have unveiled complex topologies and heterogeneous structures, with connectivity patterns statistically characterized by heavy-tails and large fluctuations, scale-free properties and non trivial correlations such as high clustering and hierarchical ordering [73]. A large amount of work has been devoted to the development of new tools for statistical characterisation and modelling of networks, in order to identify their most relevant properties, and to understand which growth mechanisms could lead to these properties. Most of those contributions have focused on static graphs or on dynamic process (*e.g.* diffusion) occurring on static graphs. This has called forth a major effort in developing the methodology to characterize the topology and temporal behaviour of complex networks [73], [63], [80], [69], to describe the observed structural and temporal heterogeneities [56], [63], [57], to detect and measure emerging community structures [60], [77], [78], to see how the functionality of networks determines their evolving structure [68], and to determine what kinds of correlations play a role in their dynamics [64], [67], [72].

The challenge is now to extend this kind of statistical characterization to dynamical graphs. In other words, links in dynamic networks are temporal events, called contacts, which can be either punctual or last for some period of time. Because of the complexity of this analysis, the temporal dimension of the network is often ignored or only roughly considered. Therefore, fully taking into account the dynamics of the links into a network is a crucial and highly challenging issue.

Another powerful approach to model time-varying graphs is via activity driven network models. In this case, the only assumption relates to the distribution of activity rates of interacting entities. The activity rate is realistically broadly distributed and refers to the probability that an entity becomes active and creates a connection with another entity within a unit time step [75]. Even the generic model is already capable to recover some realistic features of the emerging graph, its main advantage is to provide a general framework to study various types of correlations present in real temporal networks. By synthesising such correlations (*e.g.* memory effects, preferential attachment, triangular closing mechanisms, ...) from the real data, we are able to extend the general mechanism and build a temporal network model, which shows certain realistic feature in a controlled way. This can be used to study the effect of selected correlations on the evolution of the emerging structure [66] and its co-evolution with ongoing processes like spreading phenomena, synchronisation, evolution of consensus, random walk etc. [66], [74]. This approach allows also to develop control and immunisation strategies by fully considering the temporal nature of the backgrounding network.

3.3. Distributed Algorithms for dynamic networks: regulation, adaptation and interaction

Participants: Thomas Begin, Anthony Busson, Isabelle Guérin Lassous, Philippe Nain.

Dedicated algorithms for dynamic networks. First, the dynamic network object itself trigger original algorithmic questions. It mainly concerns distributed algorithms that should be designed and deployed to efficiently measure the object itself and get an accurate view of its dynamic behavior. Such distributed measure should be “transparent”, that is, it should introduce no bias or at least a bias that is controllable and corrigible. Such problem is encountered in all distributed metrology measures / distributed probes: P2P, sensor network, wireless network, QoS routing... This question raises naturally the intrinsic notion of adaptation and control of the dynamic network itself since it appears that autonomous networks and traffic aware routing are becoming crucial.

Communication networks are dynamic networks that potentially undergo high dynamicity. The dynamicity exhibited by these networks results from several factors including, for instance, changes in the topology and varying workload conditions. Although most implemented protocols and existing solutions in the literature can cope with a dynamic behavior, the evolution of their behavior operates identically whatever the actual properties of the dynamicity. For instance, parameters of the routing protocols (*e.g.* hello packets transmission frequency) or routing methods (*e.g.* reactive / proactive) are commonly hold constant regardless of the nodes mobility. Similarly, the algorithms ruling CSMA/CA (*e.g.* size of the contention window) are tuned identically and they do not change according to the actual workload and observed topology.

Dynamicity in computer networks tends to affect a large number of performance parameters (if not all) coming from various layers (viz. physical, link, routing and transport). To find out which ones matter the most for our intended purpose, we expect to rely on the tools developed by the two former axes. These quantities should capture and characterize the actual network dynamicity. Our goal is to take advantage of this latter information in order to refine existing protocols, or even to propose new solutions. More precisely, we will attempt to associate "fundamental" changes occurring in the underlying graph of a network (reported through graph-based signal tools) to quantitative performance that are matter of interests for networking applications and the end-users. We expect to rely on available testbeds such as SensLab and FIT to experiment our solutions and ultimately validate our approach.

4. Application Domains

4.1. Life Science & Health

In parallel to the advances in modern medicine, health sciences and public health policy, epidemic models aided by computer simulations and information technologies offer an increasingly important tool for the understanding of transmission dynamics and of epidemic patterns. The increased computational power and use of Information and Communication Technologies make feasible sophisticated modelling approaches augmented by detailed in vivo data sets, and allow to study a variety of possible scenarios and control strategies, helping and supporting the decision process at the scientific, medical and public health level. The research conducted in the DANTE project finds direct applications in the domain of LSH since modelling approaches crucially depend on our ability to describe the interactions of individuals in the population. In the TUBEXPO and ARIBO projects, we are collaborating with Pr. Jean-Christopge Lucet (Professeur des université Paris VII, Praticien hospitalier APHP). Within PhD work of G. Frusque, we collaborate with Dr. Julien Jung from Hôpital de Neurologie de Bron (HCL) and with Nadine Ravel, DR CNRS (CRNL, INSERM).

4.2. Network Science / Complex networks

In the last ten years the science of complex networks has been assigned an increasingly relevant role in defining a conceptual framework for the analysis of complex systems. Network science is concerned with graphs that map entities and their interactions to nodes and links. For a long time, this mathematical abstraction has contributed to the understanding of real-world systems in physics, computer science, biology, chemistry, social sciences, and economics. Recently, however, enormous amounts of detailed data, electronically collected and meticulously catalogued, have finally become available for scientific analysis and study. This has led to the discovery that most networks describing real world systems show the presence of complex properties and heterogeneities, which cannot be neglected in their topological and dynamical description. This has called forth a major effort in developing the methodology to characterise the topology and temporal behaviour of complex networks, to describe the observed structural and temporal heterogeneities, to detect and measure emerging community structure, to see how the functionality of networks determines their evolving structure, and to determine what kinds of correlations play a role in their dynamics. All these efforts have brought us to a point where the science of complex networks has become advanced enough to help us to disclose the deeper roles of complexity and gain understanding about the behaviour of very complicated systems.

In this endeavour the DANTE project targets the study of dynamically evolving networks, concentrating on questions about the evolving structure and dynamical processes taking place on them. During the last year we developed developed several projects along these lines concerning three major datasets:

- Mobile telephony data: In projects with academic partners and Grandata we performed projects based on two large independent datasets collecting the telephone call and SMS event records for million of anonymised individuals. The datasets record the time and duration of mobile phone interactions and some coarse grained location and demographic data for some users. In addition one of the dataset is coupled with anonymised bank credit information allowing us to study directly the socioeconomic structure of a society and how it determines the communication dynamics and structure of individuals.
- Skype data: Together with Skype Labs/STACC and other academic groups we were leading projects in the subject of social spreading phenomena. These projects were based on observations taken from a temporally detailed description of the evolving social network of (anonymised) Skype users registered between 2003 and 2011. This data contains dates of registration and link creation together with gradual information about their location and service usage dynamics.
- Twitter data: In collaboration with ICAR-ENS Lyon we collected a large dataset about the microblogs and communications of millions of Twitter users in the French Twitter space. This data allows us to follow the spreading of fads/opinions/hashtags/ideas and more importantly linguistic features in online communities. The aim of this collaboration is to set the ground for a quantitative framework studying the evolution of linguistic features and dialects in an social-communication space mediated by online social interactions.

4.3. Social Sciences / Epistemology

Political impacts of the internet and of networks begin to be well known (Cambridge Analytica, Russian trolls, etc.). Hence the public at large begins to be aware of the abuses of the leaders of the internet (privacy by firms and advertising, surveillance by states, fake news by activists, etc.). In the same time, on-line exchanges now give scientific estimations of the political life [61] and policital sciences begin to consider the internet as a relevant subject of study. As the internet is a *technology*, philosophy is the best approach to understand what socially happens (or can be made) with the internet. We develop it by two ways:

- political philosophy. Some Dante members are working with Triangle laboratory (social sciences, philosophy and politics; Ens de Lyon and CNRS).
- epistemology, because computer sciences discoveries are related with the evolution of science(s): we discovered that, in the case of the internet, political philosophy can do few if not strongly related to epistemology [62]. Epistemological approach is developped in collaboration with Jean Dhombres (who holds a seminar at Enssib: http://barthes.enssib.fr/cours/Dhombres2018-2019.html and with Hcéres (new relations between social and exact sciences, cf. point 9.2.6).

This approach should help computer scientists to understand how their research may depend on foreign initiatives and to create new links between social sciences and Inria.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Machine Learning & Data Science for Complex and Dynamical Models

The Dante team is part of a consortium (including the LIP, the Physics Lab from ENS de Lyon, the LabHC from U. Jean Monnet and LIRIS from U. Lyon 1) that got funded a 4 years project within the call “Scientific Breakthrough” of **IDEX de Lyon**.

With a total envelope of 1.2M euros, the project "mAChine LeArning & Data sciEnce for coMplex and dynamICal modelS" (ACADEMICS) combines **Machine Learning (ML) and Data Science (DS)** for the purpose of scientific research into two challenging directions:

1. **Computing and information processing** – develop new theoretical frameworks and learning algorithms adapted to difficult scientific contexts involving heterogeneous, irregular, error-prone, dynamic and complex data, while taking into account prior knowledge whenever it is relevant.
2. **Complex and dynamic models learning** – leverage the synergy between ML and DS to devise data-driven models in two scientific domains: **climate modeling**, and **quantitative understanding of social systems**. Focusing on these two case studies, the project will tackle the key issue of how to learn intricate models from numerous, heterogeneous and dynamic data.

The **research program** is elaborated along specific scientific issues that can reasonably lead to significant results within the 3-year lifetime of the project. The two case studies are instrumental to frame the way ML and DS can combine to yield relevant models. The methodological axes are:

- **Representation and model learning for complex data:** How to find sparse latent spaces for complex data or graphs, and how to learn compressed models? How to identify exceptional phenomena?
- **Estimation and learning from multi-source and/or dynamic data:** How to transfer a model learned from *source* data to related but different *target* data? How to learn from multi-source complex data?
- **Distributed and adaptive machine learning for graphs and complex models:** How to design distributed optimization-based learning? How to develop adaptive and distributed model inference in high dimension?

In close connexion with these methodological questions, the climate modeling use-case raises the central interrogation of **how to learn effective dynamic models**, firstly in a nonparametric way by means of ML tools and secondly, by mixing several data sources (from observations and simulations). As for computational social science, the challenge is **to embed together in ML approaches, individual features, global structures and dynamics of social networks**. The goal here, is to benefit from their complementarity to infer latent correlations, to identify behavioral mechanisms and to better model emergent social phenomena.

5.1.2. Books on Dynamic Networks by Márton Karsai

After a book chapter on *Control Strategies of Contagion Processes in Time-varying Networks* in Temporal Network Epidemiology in collaboration with Nicola Perra [65], a full book on *Bursty Human Dynamics* was just released at the end of the year in collaboration with Hang-Hyun Jo and Kimmo Kaski [40].

5.1.3. Awards

- Márton Karsai received the Junior Scientific Award of the Complex System Society, Sept. 2018.
- Márton Karsai, awarded Fellow of the ISI Foundation (Torino, Italy), 2018.
- Samuel Unicomb (PhD of Márton Karsai) obtained the best poster award at the NetSci'18 conference in Paris in June 2018.

6. New Software and Platforms

6.1. GraSP

Graph Signal Processing
KEYWORDS: Matlab - LaTeX - Graph - Graph visualization - Signal processing - GNU Octave
FUNCTIONAL DESCRIPTION: Matlab / GNU Octave toolbox to manipulate and visualize signals on graphs. LaTeX package to draw signals.

- Contact: Benjamin Girault

6.2. IoT-LAB aggregation-tools

KEYWORD: Internet of things
FUNCTIONAL DESCRIPTION: IoT-LAB aggregation-tools allow aggregating data results from many nodes at a time. It connects to several tcp connections and handle the received data.

- Participant: Gaetan Harter
- Contact: Eric Fleury
- URL: https://github.com/iot-lab/aggregation-tools

6.3. IoT-LAB cli-tools

KEYWORD: Internet of things
FUNCTIONAL DESCRIPTION: IoT-LAB cli-tools provide a basic set of operations for managing IoT-LAB experiments from the command-line.

- Participants: Frédéric Saint-Marcel and Gaetan Harter
- Contact: Eric Fleury
- URL: https://github.com/iot-lab/cli-tools

6.4. IoT-LAB gateway

KEYWORD: Internet of things
FUNCTIONAL DESCRIPTION: IoT-LAB software embedded on a IoT-LAB gateway node new generation provides the local management of the experiment on that node. It is a software bridge between the IoT-LAB server, the user open node and the control node.

- Contact: Frédéric Saint-Marcel
- URL: https://github.com/iot-lab/iot-lab-gateway

6.5. IoT-LAB robots

KEYWORDS: Internet of things - Robotics
FUNCTIONAL DESCRIPTION: IoT-LAB robots is an embedded robot controler on a Turtlebot2 providing the IoT-LAB node mobility functionnality

- Partner: Université de Strasbourg
- Contact: Julien Vandaële
- URL: https://github.com/iot-lab/

6.6. Queueing Systems

FUNCTIONAL DESCRIPTION: This tool aims at providing a simple web interface to promote the use of our proposed solutions to numerically solve classical queueing systems.

- Participants: Alexandre Brandwajn and Thomas Begin
- Contact: Thomas Begin
- URL: http://queueing-systems.ens-lyon.fr/

6.7. WSNet

KEYWORD: Network simulator

FUNCTIONAL DESCRIPTION: WSNet is a modular event-driven simulator targeted to Wireless Sensor Networks. Its main goals are to offer scalabiliy, extensibility and modularity for the integration of new protocols/hardware models and a precise radio medium simulation. We still hope to find the proper resource to make WSNet evolve into a wireless capillary network simulator suitable for conducting simulations at the urban scale.

- Participants: Rodrigue Domga Komguem and Fabrice Valois
- Partner: CEA-LETI
- Contact: Guillaume Chelius
- URL: https://gforge.inria.fr/projects/wsnet-3/

7. New Results

7.1. Graph Signal Processing and Machine Learning

Participants: Paulo Gonçalves, Esteban Bautista Ruiz, Mikhail Tsitsvero, Sarah de Nigris.

7.1.1. Analytic signal in many dimensions

In a series of two articles [30] and [54] (in collaboration with P. Borgnat), we extended analytic signal to the multidimensional case. First we showed how to obtain separate phase-shifted components and how to combine them into instantaneous amplitude and phase. Secondly we defined the proper hypercomplex analytic signal as a holomorphic hypercomplex function on the boundary of polydisk in the hypercomplex space. Next it was shown that the correct phase-shifted components can be obtained by positive frequency restriction of the Scheffers-Fourier transform based on the commutative and associative algebra generated by the set of elliptic hypercomplex numbers. Moreover we demonstrated that for d > 2 there is no corresponding Clifford-Fourier transform that allows to recover phase-shifted components correctly. Finally the euclidean-domain construction of instantaneous amplitude was extended to manifold and manifold-like graphs and point clouds.

7.1.2. BGP Zombies: an Analysis of Beacons Stuck Routes

Joint work with Romain Fontugne (IIJ Research Lab, Japan) and Patrice Abry (CNRS, Physics Lab of ENS de Lyon) [25].

Network operators use the Border Gateway Protocol (BGP) to control the global visibility of their networks. When withdrawing an IP prefix from the Internet, an origin network sends BGP withdraw messages, which are expected to propagate to all BGP routers that hold an entry for that address space in their routing table. Yet network operators occasionally report issues where routers maintain routes to IP prefixes withdrawn by their origin network. We refer to this problem as BGP zombies and characterize their appearance using RIS BGP beacons, a set of prefixes withdrawn every four hours at predetermined times. Across the 27 monitored beacon prefixes, we observe usually more than one zombie outbreak per day. But their presence is highly volatile, on average a monitored peer misses 1.8% withdraws for an IPv4 beacon (2.7% for IPv6). We also discovered that BGP zombies can propagate to other ASes, for example, zombies in a transit network are inevitably affect- ing its customer networks. **We employ a graph-based semi-supervised machine learning technique to estimate the scope of zombies propagation**, and found that most of the observed zombie outbreaks are small (i.e. on average 10% of monitored ASes for IPv4 and 17% for IPv6). We also report some large zombie outbreaks with almost all monitored ASes affected.

7.1.3. Design of graph filters and filterbanks

Book chapter [43], co-authored with Nicolas Tremblay (CNRS, UGA Gipsa-Lab) and Pierre Borgnat (CNRS, Physics Lab, ENS de Lyon).

Basic operations in graph signal processing consist in processing signals indexed on graphs either by filtering them or by changing their domain of representation, in order to better extract or analyze the important information they contain. The aim of this chapter is to review general concepts underlying such filters and representations of graph signals. We first recall the different Graph Fourier Transforms that have been developed in the literature, and show how to introduce a notion of frequency analysis for graph signals by looking at their variations. Then, we move to the introduction of graph filters, that are defined like the classical equivalent for 1D signals or 2D images, as linear systems which operate on each frequency band of a signal. Some examples of filters and of their implementations are given. Finally, as alternate representations of graph signals, we focus on multiscale transforms that are defined from filters. Continuous multiscale transforms such as spectral wavelets on graphs are reviewed, as well as the versatile approaches of filterbanks on graphs. Several variants of graph filterbanks are discussed, for structured as well as arbitrary graphs, with a focus on the central point of the choice of the decimation or aggregation operators.

7.2. Optimization

Participant: Marion Foare.

7.2.1. A new proximal method for joint image restoration and edge detection with the Mumford-Shah model

Joint work with Nelly Pustelnik (CNRS, Physics Lab of ENS de Lyon) and Laurent Condat (CNRS, GIPSA Lab) [24].

In this paper, we propose an adaptation of the PAM algorithm to the minimization of a nonconvex functional designed for joint image denoising and contour detection. This new functional is based on the Ambrosio–Tortorelli approximation of the well-known Mumford–Shah functional. We motivate the proposed approximation, offering flexibility in the choice of the possibly non-smooth penalization, and we derive closed form expression for the proximal steps involved in the algorithm. We focus our attention on two types of penalization: 1-norm and a proposed quadratic-1 function. Numerical experiments show that the proposed method is able to detect sharp contours and to reconstruct piecewise smooth approximations with low computational cost and convergence guarantees. We also compare the results with state-of-the-art re-laxations of the Mumford–Shah functional and a recent discrete formulation of the Ambrosio–Tortorelli functional.

7.2.2. Semi-Linearized Proximal Alternating Minimization for a Discrete Mumford–Shah Model

Joint work with Nelly Pustelnik (CNRS, Physics Lab of ENS de Lyon) and Laurent Condat (CNRS, GIPSA Lab) [51].

The Mumford–Shah model is a standard model in image segmentation and many approximations have been proposed in order to approximate it. The major interest of this functional is to be able to perform jointly image restoration and contour detection. In this work, we propose a general formulation of the discrete counterpart of the Mumford–Shah functional, adapted to nonsmooth penalizations, fitting the assumptions required by the Proximal Alternating Linearized Minimization (PALM), with convergence guarantees. A second contribution aims to relax some assumptions on the involved functionals and derive a novel Semi-Linearized Proximal Alternated Minimization (SL-PAM) algorithm, with proved convergence. We compare the performances of the algorithm with several nonsmooth penalizations, for Gaussian and Poisson denoising, image restoration and RGB-color denoising. We compare the results with state-of-the-art convex relaxations of the Mumford–Shah functional, and a discrete version of the Ambrosio–Tortorelli functional. We show that the SL-PAM algorithm is faster than the original PALM algorithm, and leads to competitive denoising, restoration and segmentation results.

7.2.3. Discrete Mumford-Shah on graph for mixing matrix estimation

Joint work with Yacouba Kaloga (Physics Lab of ENS de Lyon), Nelly Pustelnik (CNRS, Physics Lab of ENS de Lyon) and Pablo Jensen (CNRS, Physics Lab of ENS de Lyon) [53].

The discrete Mumford-Shah formalism has been introduced for the image denoising problem, allowing to capture both smooth behavior inside an object and sharp transitions on the boundary. In the present work, we propose first to extend this formalism to graphs and to the problem of mixing matrix estimation. New algorithmic schemes with convergence guarantees relying on proximal alternating minimization strategies are derived and their efficiency (good estimation and robustness to initialization) are evaluated on simulated data, in the context of vote transfer matrix estimation.

7.3. Wireless & Wired Networks

Participants: Thomas Begin, Anthony Busson, Isabelle Guérin Lassous.

7.3.1. Conflict graph-based model for IEEE 802.11 networks: A Divide-and-Conquer approach

WLANs (Wireless Local Area Networks) based on the IEEE 802.11 standard have become ubiquitous in our daily lives. We typically augment the number of APs (Access Points) within a WLAN to extend its coverage and transmission capacity. This leads to network densification, which in turn demands some form of coordination between APs so as to avoid potential misconfigurations. In our article [20], we describe a performance modeling method that can provide guidance for configuring WLANs and be used as a decision-support tool by a network architect or as an algorithm embedded within a WLAN controller. The proposed approach estimates the attained throughput of each AP, as a function of the WLAN's conflict graph, the AP loads, the frame sizes, and the link transmission rates. Our modeling approach employs a Divide-and-Conquer strategy which breaks down the original problem into multiple sub-problems, whose solutions are then combined to provide the solution to the original problem. We conducted extensive simulation experiments using the ns-3 simulator that show the model's accuracy is generally good with relative errors typically less than 10%. We then explore two issues of WLAN configuration: choosing a channel allocation for the APs and enabling frame aggregation on APs.

7.3.2. Video on Demand in IEEE 802.11p-based Vehicular Networks: Analysis and Dimensioning

This is a joint work with A. Boukerche. In [31], we consider a VoD (Video on-Demand) platform designed for vehicles traveling on a highway or other major roadway. Typically, cars or buses would subscribe to this delivery service so that their passengers get access to a catalog of movies and series stored on a back-end server. Videos are delivered through IEEE 802.11p Road Side Units deployed along the highway. In this paper, we propose a simple analytical and yet accurate solution to estimate (at the speed of a click) two key performance parameters for a VoD platform: (i) the total amount of data down-loaded by a vehicle over its journey and (ii) the total "interruption time" , which corresponds to the time a vehicle spends with the playback of its video interrupted because of an empty buffer. After validating its accuracy against a set of simulations run with ns-3, we show an example of application of our analytical solution for the sizing of an IEEE 802.11p-based VoD platform.

7.3.3. An accurate and efficient modeling framework for the performance evaluation of DPDK-based virtual switches

This is a joint work with B. Baynat, G. Artero Gallardo and V. Jardin [4]. Data plane development kit (DPDK) works as a specialized library that enables virtual switches to accelerate the processing of incoming packets by, among other things, balancing the incoming flow of packets over all the CPU cores and processing packets by batches to make a better use of the CPU cache. Although DPDK has become a de facto standard, the performance modeling of a DPDK-based vSwitch remains a challenging problem. In this paper, we present an analytical queueing model to evaluate the performance of a DPDK-based vSwitch. Such a virtual equipment is represented by a complex polling system in which packets are processed by batches, i.e., a given CPU core processes several packets of one of its attached input queues before switching to the next one. To reduce the complexity of the associated model, we develop a general framework that consists in decoupling the polling system into several queueing subsystems, each one corresponding to a given CPU core. We resort to servers with vacation to capture the interactions between subsystems. Our proposed solution is conceptually simple, easy to implement and computationally efficient. Tens of comparisons against a discrete-event simulator show that our models typically deliver accurate estimates of the performance parameters of interest (e.g., attained throughput, packet latency or loss rate). We illustrate how our models can help in determining an adequate setting of the vSwitch parameters using several real-life case studies.

7.3.4. Association optimization in Wi-Fi networks

Densification of Wi-Fi networks has led to the possibility for a wireless station to choose between several access points (APs), improving coverage, wireless link quality and mobility. But densification of APs may generate interference, contention and decrease the global throughput as these APs have to share a limited number of channels. The recent trend in which Wi-Fi networks are managed in a centralized way offers the opportunity to alleviate this problem through a global optimization of the resource usage. In particular, optimizing the association step between APs and stations can increase the overall throughput and fairness between stations. In this work, we propose an original solution to this optimization problem. First, we propose a mathematical model to evaluate and forecast the throughput achieved for each station for a given association. The best association is then defined as the one that maximizes a logarithmic utility function using the stations' throughputs predicted by the model. The use of a logarithmic utility function allows to achieve a good trade-off between overall throughput and fairness. A heuristic based on a local search algorithm is used to propose approximate solutions to this optimization problem. This approach has the benefit to be tuned according to the CPU and time constraints of the WLAN controller. A comparison between different heuristic versions and the optimum solution shows that the proposed heuristic offers solutions very close to the optimum with a significant gain of time.

In the first place, we consider a saturated network. Even if such traffic conditions are rare, the optimization of the association step under this assumption has the benefit to fairly share the bandwidth between stations. Nevertheless, traffic demands may be very different from one station to another and it may be more useful to optimize associations according to the stations' demands. In a second step, we propose an optimization of the association step based on the stations' throughputs and the channel busy time fraction (BTF). The latter is defined as the proportion of time the channel is sensed busy by an AP. We propose an analytical model that predicts BTF for any configuration. Associations are optimized in order to minimize the greatest BTF in the network. This original approach allows the Wi-Fi manager to unload the most congested AP, increase the throughput for most of the stations, and offer more bandwidth to stations that need it. We present a local search technique that finds local optima to this optimization problem. This heuristic relies on an analytical model that predicts BTF for any configuration. The model is based on a Markov network and a Wi-Fi conflict graph. NS-3 simulations including a large set of scenarios highlight the benefits of our approach and its ability to improve the performance in congested and non-congested Wi-Fi networks.

Lastly, we consider the latest amendments of the IEEE 802.11 standard. The main challenges are to propose models that take into account recent enhancements such as spatial multiplexing (MIMO) at the physical layer and frame aggregation mechanism at the MAC layer. To assess these new features, we derive an association optimization approach based on a new metric, named Hypothetical Busy Time Fraction (H-BTF), that combines the classical Busy Time Fraction (BTF) and the frame aggregation mechanism [3].

7.3.5. Transient analysis of idle time in VANETs using Markov-reward models

The development of analytical models to analyze the behavior of vehicular ad hoc networks (VANETs) is a challenging aim. Adaptive methods are suitable for many algorithms (*e.g.* choice of forwarding paths, dynamic resource allocation, channel control congestion) and services (*e.g.* provision of multimedia services, message dissemination). These adaptive algorithms help the network to maintain a desired performance level. However, this is a difficult goal to achieve, especially in VANETs due to fast position changes of the VANET nodes. Adaptive decisions should be taken according to the current conditions of the VANET. Therefore, evaluation of transient measures is required for the characterization of VANETs. In the literature, different works address the characterization and measurement of the idle (or busy) time to be used in different proposals to attain a more efficient usage of wireless network. We focus on the idle time of the link between two VANET nodes. Specifically, we have developed an analytical model based on a straightforward Markov reward chain (MRC) to obtain transient measurements of this idle time. Numerical results from the analytical model fit well with simulation results [12].

7.4. Performance Evaluation of Communication Networks

Participants: Thomas Begin, Philippe Nain, Isabelle Guérin Lassous.

7.4.1. First-Come-First-Served Queues with Multiple Servers and Customer Classes

This a joint work with A. Brandwajn [5]. We present a simple approach to the solution of a multi-server FCFS queueing system with several classes of customers and phase-type service time distributions. The proposed solution relies on solving a single two-class model in which we distinguish one of the classes and we aggregate the remaining customer classes. We use a reduced state approximation to solve this two-class model. We propose two types of aggregation: exact, in which we merge the phase-type service time distributions exactly, and approximate, in which we simplify the phase-type distribution for the aggregated class by matching only its first two moments. The proposed approach uses simple mathematics and is highly scalable in terms of the number of servers, the number of classes, as well as the number of phases per class. Our approach applies both to queues with finite and infinite buffer space.

7.4.2. A study of systems with multiple operating levels, probabilistic thresholds and hysteresis

This a joint work with A. Brandwajn [6]. Current architecture of many computer systems relies on dynamic allocation of a pool of resources according to workload conditions to meet specific performance objectives while minimizing cost (e.g., energy or billing). In such systems, different levels of operation may be defined, and switching between operating levels occurs at certain thresholds of system congestion. To avoid rapid oscillations between levels of service, "hysteresis" is introduced by using different thresholds for increasing and decreasing workload levels, respectively. We propose a model of such systems with general arrivals, arbitrary number of servers and operating levels where each higher operating level may correspond to an arbitrary number of additional servers and soft (i.e. non-deterministic) thresholds to account for "inertia" in switching between operating levels. In our model, request service times are assumed to be memoryless and server processing rates may be a function of the current operating level and of the number of requests (users) in the system. Additionally, we allow for delays in the activation of additional operating levels. We use simple mathematics to obtain a semi-numerical solution of our model. We illustrate the versatility of our model using several case study examples inspired by features of real systems. In particular, we explore optimal thresholds as a tradeoff between performance and energy consumption.

7.4.3. Covert cycle stealing in an M/G/1 queue

Consider an M/G/1 queue where arriving jobs are under control of a party (Willie). There exists a second party, Alice who may or may not want to introduce a sequence of jobs to be serviced. Her goal is to prevent Willie from being able to distinguish between these two cases. The question that we address is: can Alice introduce her stream of jobs covertly, i.e., prevent Willie from distinguishing between the two possibilities, either her introducing the stream or not, and if so, at what rate can she introduce her jobs? We present a square-root law on the amount of service Alice can receive covertly. The covertness criterion is that the probabilities of false alarm and missed detection is arbitrarily close to one. One result we have established is the following: consider exponential service times for Alice's jobs and Willies' jobs with rate μ_1 and μ_2, respectively. During n Willie's job busy periods, Alice can submit covertly $O(\sqrt{n})$ jobs if $\mu_1 < 2\mu_2$, $O(\sqrt{n/\log n})$ jobs if $\mu_1 = 2\mu_2$, and $O(n^{\mu_1/\mu_2})$ jobs if $\mu_1 > 2\mu_2$. This is the first time that such a phase transition has been observed in this context. This ongoing research, carried out by P. Nain in collaboration with D. Towsley (Univ. Massachusetts) and B. Jiang (Shanghai Jiao Tong Univ.), has various applications in the context of service level agreement.

7.4.4. LRU caches

The work on network caches operating under the standard Least-Recently-Used (LRU) management policy, initiated in 2017 (see 2017 Dante Activity Report), has been completed and published [13]. Under weak statistical assumptions on the content request process, this work establishes the validity of the so-called "Che's approximation" as the cache size and the number of content go to infinity.

7.4.5. Stochastic Multilayer Networks

A stochastic multilayer network is the aggregation of M networks (one per layer) where each is a subgraph of a foundational network G. Each layer network is the result of probabilistically removing links and nodes from G. The resulting network includes any link that appears in at least K layers. This model is an instance of a non-standard site-bond percolation model. Two sets of results are obtained in [28]: first, we derive the probability distribution that the M-layer network is in a given configuration for some particular graph structures (explicit results are provided for a line and an algorithm is provided for a tree), where a configuration is the collective state of all links (each either active or inactive). Next, we show that for appropriate scalings of the node and link selection processes in a layer, links are asymptotically independent as the number of layers goes to infinity, and follow Poisson distributions. Numerical results are provided to highlight the impact of having several layers on some metrics of interest (including expected size of the cluster a node belongs to in the case of the line). This model finds applications in wireless communication networks with multichannel radios, multiple social networks with overlapping memberships, transportation networks, and, more generally, in any scenario where a common set of nodes can be linked via co-existing means of connectivity.

7.5. Computational Human Dynamics and Temporal Networks

Participants: Márton Karsai, Éric Fleury, Jean-Philippe Magué, Philippe Nain, Jean-Pierre Chevrot.

7.5.1. Correlations and dynamics of consumption patterns in social-economic networks

In [16], we analyse a coupled dataset collecting the mobile phone communications and bank transactions history of a large number of individuals living in a Latin American country [16]. After mapping the social structure and introducing indicators of socioeconomic status, demographic features, and purchasing habits of individuals, we show that typical consumption patterns are strongly correlated with identified socioeconomic classes leading to patterns of stratification in the social structure. In addition, we measure correlations between merchant categories and introduce a correlation network, which emerges with a meaningful community structure. We detect multivariate relations between merchant categories and show correlations in purchasing habits of individuals. Finally, by analysing individual consumption histories, we detect dynamical patterns in purchase behaviour and their correlations with the socioeconomic status, demographic characters and the egocentric social network of individuals. Our work provides novel and detailed insight into the relations between social and consuming behaviour with potential applications in resource allocation, marketing, and recommendation system design.

7.5.2. Mapping temporal-network percolation to weighted, static event graphs

The dynamics of diffusion-like processes on temporal networks are influenced by correlations in the times of contacts. This influence is particularly strong for processes where the spreading agent has a limited lifetime at nodes: disease spreading (recovery time), diffusion of rumors (lifetime of information), and passenger routing (maximum acceptable time between transfers). In [14], we introduce weighted event graphs as a powerful and fast framework for studying connectivity determined by time-respecting paths where the allowed waiting times between contacts have an upper limit. We study percolation on the weighted event graphs and in the underlying temporal networks, with simulated and real-world networks. We show that this type of temporal-network percolation is analogous to directed percolation, and that it can be characterized by multiple order parameters.

7.5.3. Randomized reference models for temporal networks

Many real-world dynamical systems can successfully be analyzed using the temporal network formalism. Empirical temporal networks and dynamic processes that take place in these situations show heterogeneous, non-Markovian, and intrinsically correlated dynamics, making their analysis particularly challenging. Randomized reference models (RRMs) for temporal networks constitute a versatile toolbox for studying such systems. Defined as ensembles of random networks with given features constrained to match those of an input (empirical) network, they may be used to identify statistically significant motifs in empirical temporal networks (i.e. over-represented w.r.t. the null random networks) and to infer the effects of such motifs on dynamical processes unfolding in the network. However, the effects of most randomization procedures on temporal network characteristics remain poorly understood, rendering their use non-trivial and susceptible to misinterpretation. In the work presented in [52], we propose a unified framework for classifying and understanding microcanonical RRMs (MRRMs). We use this framework to propose a canonical naming convention for existing randomization procedures, classify them, and deduce their effects on a range of important temporal network features. We furthermore show that certain classes of compatible MRRMs may be applied in sequential composition to generate more than a hundred new MRRMs from existing ones surveyed in this article. We provide a tutorial for the use of MRRMs to analyze an empirical temporal network and we review applications of MRRMs found in literature. The taxonomy of MRRMs we have developed provides a reference to ease the use of MRRMs, and the theoretical foundations laid here may further serve as a base for the development of a principled and systematic way to generate and apply randomized reference null models for the study of temporal networks.

7.5.4. Socioeconomic dependencies of linguistic patterns in Twitter: a multivariate analysis

Our usage of language is not solely reliant on cognition but is arguably determined by myriad external factors leading to a global variability of linguistic patterns. This issue, which lies at the core of sociolinguistics and is backed by many small-scale studies on face-to-face communication, is addressed in [29], by constructing a dataset combining the largest French Twitter corpus to date with detailed socioeconomic maps obtained from national census in France. We show how key linguistic variables measured in individual Twitter streams depend on factors like socioeconomic status, location, time, and the social network of individuals. We found that (i) people of higher socioeconomic status, active to a greater degree during the daytime, use a more standard language; (ii) the southern part of the country is more prone to use more standard language than the northern one, while locally the used variety or dialect is determined by the spatial distribution of socioeconomic status; and (iii) individuals connected in the social network are closer linguistically than disconnected ones, even after the effects of status homophily have been removed. Our results inform sociolinguistic theory and may inspire novel learning methods for the inference of socioeconomic status of people from the way they tweet.

7.5.5. Threshold driven contagion on weighted networks

Weighted networks capture the structure of complex systems where interaction strength is meaningful. This information is essential to a large number of processes, such as threshold dynamics, where link weights reflect the amount of influence that neighbours have in determining a node's behaviour. Despite describing numerous cascading phenomena, such as neural firing or social contagion, the modelling of threshold dynamics on weighted networks has been largely overlooked. We fill this gap in [21], by studying a dynamical threshold model over synthetic and real weighted networks with numerical and analytical tools. We show that the time of cascade emergence depends non-monotonously on weight heterogeneities, which accelerate or decelerate the dynamics, and lead to non-trivial parameter spaces for various networks and weight distributions. Our methodology applies to arbitrary binary state processes and link properties, and may prove instrumental in understanding the role of edge heterogeneities in various natural and social phenomena.

7.5.6. Link transmission centrality in large-scale social networks

Understanding the importance of links in transmitting information in a network can provide ways to hinder or postpone ongoing dynamical phenomena like the spreading of epidemic or the diffusion of information. In our work [22], we propose a new measure based on stochastic diffusion processes, the*transmission centrality*, that captures the importance of links by estimating the average number of nodes to whom they transfer information during a global spreading diffusion process. We propose a simple algorithmic solution to compute transmission centrality and to approximate it in very large networks at low computational cost. Finally we apply transmission centrality in the identification of weak ties in three large empirical social networks, showing that this metric outperforms other centrality measures in identifying links that drive spreading processes in a social network.

7.5.7. Prepaid or Postpaid? That Is the Question: Novel Methods of Subscription Type Prediction in Mobile Phone Services

In the paper [41], we investigate the behavioural differences between mobile phone customers with prepaid and postpaid subscriptions. Our study reveals that (a) postpaid customers are more active in terms of service usage and (b) there are strong structural correlations in the mobile phone call network as connections between customers of the same subscription type are much more frequent than those between customers of different subscription types. Based on these observations, we provide methods to detect the subscription type of customers by using information about their personal call statistics, and also their egocentric networks simultaneously. The key of our first approach is to cast this classification problem as a problem of graph labelling, which can be solved by max-flow min-cut algorithms. Our experiments show that, by using both user attributes and relationships, the proposed graph labelling approach is able to achieve a classification accuracy of $\sim 87\%$, which outperforms by $\sim 7\%$ supervised learning methods using only user attributes. In our second problem, we aim to infer the subscription type of customers of external operators. We propose via approximate methods to solve this problem by using node attributes, and a two-way indirect inference method based on observed homophilic structural correlations. Our results have straightforward applications in behavioural prediction and personal marketing.

7.5.8. Service Adoption Spreading in Online Social Networks

The collective behaviour of people adopting an innovation, product or online service is commonly interpreted as a spreading phenomenon throughout the fabric of society. This process is arguably driven by social influence, social learning and by external effects like media. Observations of such processes date back to the seminal studies by Rogers and Bass, and their mathematical modelling has taken two directions: One paradigm, called simple contagion, identifies adoption spreading with an epidemic process. The other one, named complex contagion, is concerned with behavioural thresholds and successfully explains the emergence of large cascades of adoption resulting in a rapid spreading often seen in empirical data. The observation of real-world adoption processes has become easier lately due to the availability of large digital social network and behavioural datasets. This has allowed simultaneous study of network structures and dynamics of online service adoption, shedding light on the mechanisms and external effects that influence the temporal evolution of behavioural or innovation adoption. These advancements have induced the development of more realistic models of social spreading phenomena, which in turn have provided remarkably good predictions of various empirical adoption processes. In our chapter [39], we review recent data-driven studies addressing real-world service adoption processes. Our studies provide the first detailed empirical evidence of a heterogeneous threshold distribution in adoption. We also describe the modelling of such phenomena with formal methods and data-driven simulations. Our objective is to understand the effects of identified social mechanisms on service adoption spreading, and to provide potential new directions and open questions for future research.

7.5.9. Attention on Weak Ties in Social and Communication Networks

Granovetter's weak tie theory of social networks is built around two central hypotheses. The first states that strong social ties carry the large majority of interaction events; the second maintains that weak social ties, although less active, are often relevant for the exchange of especially important information (e.g., about potential new jobs in Granovetter's work). While several empirical studies have provided support for the first hypothesis, the second has been the object of far less scrutiny. A possible reason is that it involves notions relative to the nature and importance of the information that are hard to quantify and measure, especially in large scale studies. In our work [48], we search for empirical validation of both Granovetter's hypotheses. We find clear empirical support for the first. We also provide empirical evidence and a quantitative interpretation for the second. We show that attention, measured as the fraction of interactions devoted to a particular social connection, is high on weak ties—possibly reflecting the postulated informational purposes of such ties—but also on very strong ties. Data from online social media and mobile communication reveal network-dependent mixtures of these two effects on the basis of a platform's typical usage. Our results establish a clear relationships between attention, importance, and strength of social links, and could lead to improved algorithms to prioritize social media content.

8. Bilateral Contracts and Grants with Industry

8.1. GranData

Participants: Márton Karsai [correspondant], Éric Fleury.

Founded in 2012, Grandata is a Palo Alto-based company that leverages advanced research in Human Dynamics (the application of « big data » to social relationships and human behaviour) to identify market trends and predict customer actions. Leading telecom and financial services firms are using Grandata's Social Universe product to transform « big data » into impressive business results.

The DANTE team and Grandata started to collaborate in 2014 on the analysis of large datasets provided by the company. The aim of the collaboration is to gain better understanding about the dynamical patterns of human interactions, mobility, and the socio-economic structure of the society. This collaboration was very successful over the years, leading to several publications within the PhD thesis of Yannick Leo. Currently the collaboration is supported by the MOTIF Stic-AmSud project (2018-2020) (coordinated by Márton Karsai) which allows to meet frequently with the company. Recent projects within this collaboration are focusing on socioeconomic inference using remote sensing techniques.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Idex Lyon ACADEMICS

Participants: Paulo Gonçalves, Dominique Barbe, Gaetan Frusque.

See Section 5.1.1 for an executive summary of this project. The project brings together a **consortium of 4 teams from Laboratories of Université de Lyon** (UdL) and will form a working group with complementary expertise in machine learning (deep learning, statistical learning, data mining), in data science (complex data analysis, adaptive and/or data-driven methods, network science) and in the studies of climate modeling and of computational social science. It comprises:

- Laboratoire Informatique du Parallélisme (LIP): P. Gonçalves (PI), M. Karsai (PI for Comp. Social Sc.)
- Laboratoire de Physique (LP): P. Borgnat (Coordinator), F. Bouchet (PI for Climate)
- Laboratoire Hubert Curien (LabHC), Université Jean Monnet: M. Sebban (PI)
- Laboratoire d'InfoRmatique en Images et Systèmes d'information (LIRIS): C. Robardet (PI)

The **impacts** of the project will stem from the efficiency of our proposed methods to learn from complex and dynamic data, and if so, **future applications** will naturally follow in many areas: social science and study of social interactions, climate and environmental science but also in technological networks, neuroscience with the study of brain networks and more generally in any domain where effective dynamical models of complex situations are to be learned from data. All these situations go beyond the current classical applicative frameworks of ML (time measurements, 2D images, or texts) and compel us to work out a major scientific breakthrough.

9.1.2. ISI Torino / Dante

Participant: Márton Karsai [correspondant].

Duration of the project: **October 2016 - October 2020**.

This project involves M. Karsai and L. Gauvin (ISI Torino) and funded by the IXXI Complex System Institute. The purpose of this project is to investigate the presence and the importance of higher-order correlations in dynamical networks. As the first attempt to address this problem we applied autoencoder, a recent representation using deep neural networks, on modelled and small-scale real temporal networks. However, since the results were trivial on the modelled network and not convincing on the real one we decided to take a different approach during the second phase of the project. We involved an ISI PhD student Maddalena Toricelli, to work out a method for temporal network embedding. Our idea is to extend the node2vec representation of static networks for time-varying structures, by using a local random walk to explore the structural-temporal neighbourhood of a node. Based on such local information we can effectively propose an embedding, which captures the temporal and structural properties of nodes in a temporal network.

9.2. National Initiatives

9.2.1. Equipex FIT (Futur Internet of Things)

Participant: Éric Fleury [correspondant].

FIT was one of 52 winning projects in the Equipex research grant program. It will set up a competitive and innovative experimental facility that brings France to the forefront of Future Internet research. FIT benefits from 5.8 million euro grant from the French government Running from 22.02.11 – 31.12.2019. The main ambition is to create a first-class facility to promote experimentally driven research and to facilitate the emergence of the Internet of the future.

9.2.2. ANR GRAPHSIP (Graph Signal Processing)

Participants: Paulo Gonçalves [correspondant], Éric Fleury, Thomas Begin, Mikhail Tsitsvero.

Duration of the project: **October 2014 - October 2018**.

An increasing number of application areas require the processing of massive datasets. These data can often be represented by graphs in order to encode complex interactions. When data vectors are associated with graph vertices, a so-called graph signal is obtained. The processing of such graph signals includes several open challenges because of the nature of the involved information. Indeed graph theory and signal and image processing methodologies do not combine readily. In particular, such a combination requires new developments, allowing classical signal processing methods to work on irregular grids and non Euclidean spaces. Considering the significant success of classical signal processing tools, it appears essential to generalise their use to graph signals. The GRAPHSIP project aims at developing a set of advanced methods and algorithms for the processing of graph signals: multi-scale transforms and solutions of variational problems on graphs. The major outcomes of this project are expected to lead to significant breakthroughs for graph data processing. The project will also focus on two novel applications on instances of graph signals: brain networks and 3D colour point clouds. They will exemplify and illustrate the proposed methodological advances on emerging applications.

9.2.3. ANR SoSweet

Participants: Jean Pierre Chevrot, Éric Fleury, Márton Karsai [correspondant], Jean-Philippe Magué [PI].

Duration of the project: **November 2015 - November 2019**.

The SoSweet project focuses on the synchronic variation and the diachronic evolution of the variety of French used on Twitter. The recent rise of novel digital services opens up new areas of expression which support new linguistic behaviours. In particular, social medias such as Twitter provide channels of communication through which speakers/writers use their language in ways that differ from standard written and oral forms. The result is the emergence of new varieties of languages. The main goal of SoSweet is to provide a detailed account of the links between linguistic variation and social structure in Twitter, both synchronically and diachronically. Through this specific example, and aware of its bias, we aim at providing a more detailed understanding of the dynamic links between individuals, social structure and language variation and change.

9.2.4. ANR DylNet

Participants: Márton Karsai [correspondant], Jean Pierre Chevrot, Jean-Philippe Magué, Éric Fleury.

Duration of the project: **September 2016 - September 2020**.

The DylNet project aims to observe and to characterise the relationships between childhood sociability and oral-language learning at kindergarten. With a view to this, it takes an multidisciplinary approach combining work on language acquisition, sociolinguistics, and network science. It will be implemented by following all the children ($\approx$ 220) and teaching staff in one kindergarten over a 3-year period. The use of wireless proximity sensors will enable collection of social contacts throughout the study. The data on sociability will be linked to the results of language tests and recordings of verbal interactions used to follow the children's progress on both a psycholinguistic level (lexicon, syntax, pragmatics) and a sociolinguistic level (features showing belonging to a social group). The aim is to better understand the mechanisms of adaptation and integration at work when young children first come into contact with the school context.

9.2.5. Inria PRE LIAISON

Participants: Márton Karsai [correspondant], Éric Fleury.

Duration of the project: **November 2017 - December 2019**.

This project implements unsupervised deep learning approaches to infer correlations/patterns that exist between dynamic linguistic variables, the mesoscopic and dynamic structure of the social network, and their socio-economic attributes. This interdisciplinary project is positioned at the crossroads of Natural Language Processing (NLP), Network Science, Data Science and Machine Learning.

More precisely, we develop a joint feature-network embedding, named AN2VEC (Attributed Network to Vector), which ultimately aims at disentangling the information shared by the structure of a network and the features of its nodes. Building on the recent developments of Graph Convolutional Networks (GCN), we use a multitask GCN Variational Autoencoder where different dimensions of the generated embeddings can be dedicated to encoding feature information, network structure, or shared feature-network information separately. This method thus defines a range of models whose performance in embedding a given data set varies depending with the allocation of dimensions. By exploring the behaviour of these models on synthetic data sets having different levels of feature-network correlation, we show (i) that embeddings relying on shared information perform better than the corresponding reference with unshared information, and (ii) that this performance gap increases with the correlation between network and feature structure, thus confirming that our embedding is able to capture joint information of structure and features.

9.2.6. Inria & HCERES

Participants: Éric Guichard [correspondant], Éric Fleury.

Bilateral project on the evolution of the Multi/inter-disciplinary of SHS.

An increasing number of researchers in SHS has the desire to develop new researches with computer scientists or mathematicians because they want to apply new methodologies (according to various or numerous data) or to develop older ones, which can now be easily implemented online. Some also develop a reflexion on their discipline, with the idea that epistemological questions are revitalized by the internet. This reality invite them to discuss with philosophers or with other SHS scientists who have the same intuition (eg: cartography, visualisation).

The project is hence to measure these new forms or inter-multi-disciplinarity. The main source will be the publications of all academics of French SHS laboratories, to find out who writes a paper with somebody of a different discipline and/or laboratories. All data are anonymized,

9.2.7. Inria IPL BetterNet

Participant: Éric Guichard.

An Observatory to Measure and Improve Internet Service Access from User Experience [0].

[0] https://www.inria.fr/en/research/research-teams/inria-project-labs

BetterNet aims at building and delivering a scientific and technical collaborative observatory to measure and improve the Internet service access as perceived by users. In this Inria Project Lab, we will propose new original user-centered measurement methods, which will associate social sciences to better understand Internet usage and the quality of services and networks with a particular focus on geography and cartography.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. EMBERS

- Title: Enabling a Mobility Back-End as a Robust Service
- Programm: H2020
- Duration: December 2015 - November 2018
- Coordinator: UPMC
- Partners:
  - Fraunhofer Gesellschaft Zur Forderung Der Angewandten Forschung Ev (Germany)
  - Technische Universitat Berlin (Germany)
  - Universite Pierre et Marie Curie - Paris 6 (France)
  - Ubiwhere Lda (Portugal)
- Inria contact: Eric Fleury
- EMBERS will bring to market a back-end for smart city mobility that is developed by a European small enterprise based upon its smart parking and smart traffic management products that two municipalities in Portugal currently deploy. The Mobility Back-end as a Service (MBaaS) replaces such all-in-one systems, in which a municipality purchases the full set of components from a single vendor. Instead, the city manager can purchase best-of-breed devices and apps developed by third parties, with the only constraint being that they interoperate with the back-end via a free, open, smart city mobility API. This domain-specific API lowers barriers to entry for app and device developers, making it easier for innovative SMEs to enter the market. Furthermore, the API is offered via a variety of generic interfaces, including oneM2M, ETSI M2M, OMA LWM2M, and FIWARE NGSI. EMBERS thus clears the way for developers and to municipalities that have adopted any one of these potential emerging machine-to-machine (M2M) communication standards. Beyond its primary goal of bringing the MBaaS to market, EMBERS will stimulate development of an entire ecosystem around the MBaaS smart city mobility API. Separating out the back-end from the other components will, however, require rigorous testing. EMBERS will experiment with the system on two testbeds that are part of the FIRE OneLab facility: the FUSECO Playground, for M2M communications, and FIT IoT-LAB, for wireless sensor devices. EMBERS will host a hackathon and an app challenge to bring in third party developers. The project will also include three demonstrators by third parties via an open call. These activities will contribute back to FIRE by demonstrating successful experimentation by SMEs developing close-to-market products. The project will also conduct real world pilots in two or more cities as a final step in bringing the MBaaS to market.

9.3.1.2. ARMOUR

Title: Large-Scale Experiments of IoT Security & Trust (Project nÂ°688237)

Programm: H2020

Duration: 2015 Dec to 2018

Coordinator: UPMC

Partners:

Synelixis Lyseis Pliroforikis Automatismou & Tilepikoinonion Monoprosopi EPE (Greece)

Smartesting Solutions & Services (France)

Unparallel Innovation, Lda (Portugal)

Easy Global Market (France)

ODIN Solutions (Spain)

Universite Pierre et Marie Curie - Paris 6 (France)

Inria contact: Eric Fleury

ARMOUR will provide duly tested, benchmarked and certified Security & Trust solutions for large-scale IoT using upgraded FIRE large-scale IoT/Cloud testbeds properly-equipped for Security & Trust experimentations. ARMOUR takes the top large-scale FIT IoT-LAB testbed a FIRE OpenLAB / FIT IoT LAB facility â and enhances it as to enable experimentally-driven research on a key research dimension: large-scale IoT Security & Trust. Presently, no proper installations exist to experiment IoT Security & Trust on large-scale conditions; ARMOUR will develop and install such capability.

9.4. International Initiatives

9.4.1. Participation in International Programs

9.4.1.1. PHC Peridot

Participants: Mohammed Amer, Thomas Begin, Anthony Busson, Isabelle Guérin Lassous.

Framework for Control and Monitoring of Wireless Mesh Networks (WMN) using Software-Defined Networking (SDN). The main objective of this project is propose mechanisms and modifications in the SDN architecture, specifically in the OpenFlow, which allow SDN mechanisms to operate over WMN considering the dynamic network topology that WMN may experience and some other relevant characteristics. The project will involve devising mechanisms for controlling mesh switches through controllers in a wireless environment, which will require developing novel and WMN-specific rules, actions and commands. The project will involve proposing mechanism that consider dynamic environment of WMN along with providing redundancy in the network. Besides, there is a requirement to have an adaptive measurement API for WMN. This is the second objective of our research project. The proposed measurement API will enable the network operators to monitor network traffic over WMN which may be content-specific or host-specific. This is a joint project between DANTE and M. A. Jinnah University, Islamabad. It started in June 2015 and will end in June 2018.

9.4.2. International Initiatives

9.4.2.1. MOTif - Mobile phone sensing of human dynamics in techno-social environment

- Program: Stic AmSud
- Duration: January 2018 - December 2019
- Coordinator: Inria DANTE (Márton Karsai)
- Partners:
  - Universidad de Buenos Aires (Argentina)
  - Grandata (USA-Argentina)
  - Universidade Federal de Minas Gerais (Brazil)
  - LNCC (Brazil)

This project brings together two Inria teams - INFINE (Saclay) - DANTE (Lyon) and multiple Latin-American partners to work together with the IT company Grandata.

Scope of the project: Information and Communication Technology (ICT) is becoming increasingly social, as demonstrated by the multitude of emerging technologies and technology platforms that facilitate social interactions, taking place as communication via telephone, text message, email, online social networks etc. At the same time, our social activities are increasingly embedded in the ICT environments that enable and enhance our ability to transact, share experiences, and maintain social relationships. One of the best ways to explore these developments is through the mining and analysis of data, which are collected through mobile phones and allow us to investigate how individuals act when embedded in a technology-enabled environment. Unlimited access to a wide range of mobile applications and services may change our way to gain information, to communicate, or even to behave in different contextual places like home, work, or anywhere else. Thus understanding individual activity patterns and the source of decisions behind them is moreover important for the design of future services and to estimate the demand on the infrastructure. The MOTIf project builds on the analysis and modeling of geo-localized temporally detailed but fully anonymised mobile phone call networks. These datasets allow us to address the two scientific objectives about spatiotemporal patterns of service usage of anonymised individuals to learn when, where, and what people are doing; and about the fine-grained sociodemographic structure of society and its effect on the the individual social behaviour. In other words our goal in general is to understand how individuals behave in a dynamic techno-social environment.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Don Towsley, University of Massachusetts, Amherst, USA (march 15 - june 15, 2018).
- Alexandre Brandwajn, University of California Santa Cruz, USA (march 2018)

9.5.2. Visits to International Teams

9.5.2.1. Research Stays Abroad

- Philippe Nain visited the University of Massachusetts at Amherst (Sep. 1, 2018 - Jan. 31, 2019). He taught a graduate course on Performance Evaluation of Computer and Communication Systems and collaborated with Prof. D. Towsley and some of his students/collaborators on several research projects.
- Christophe Crespelle is on leave with a Marie Sklodowska-Curie Grant from EU. He is currently at the University of Bergen (Norway) until the beginning of 2020.
- Christophe Crespelle visited the Institute of Mathematics of the Vietnam Academy of Science and Technology, Hanoi, Vietnam (Nov. 6 - Dec. 6, 2018).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Márton Karsai was in the organizing team (general chair) of the Computational Social Science Satellite of the CCS'18 Conference in September 2018 in Thessaloniki
- Márton Karsai was in the organizing team (general chair) of the Machine Learning and Network Science Satellite of the NetSci'18 Conference in June 2018 in Paris

10.1.1.2. Member of the Organizing Committees

- Éric Guichard was the chair and organiser of the summer school *Cartography and visualisation 2018*[0].
- Márton Karsai was in the organising team (sponsor chair) of the NetSci'18 Conference held in June 2018 in Paris
- Márton Karsai was in the organising team (poster chair) of the Complex Networks Conference held in December 2018 in Cambridge

[0] http://barthes.enssib.fr/ECV-2018

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

- Thomas Begin was on the PC of the 43nd IEEE Conference on Local Computer Networks (LCN).
- Isabelle Guérin Lassous was, in 2018, a member of the program committee of the conferences ACM MSWiM, IEEE ICC and Globecom.
- Philippe Nain was a member of the program committee of the IFIP Performance 2018 conference (Toulouse, France, december 2018) and of the MAMA 2018 workshop (Irvine, CA, USA, June 2018).
- Márton Karsai was a member of the program committee of the following conferences in 2018: Complenet 2018, MLSN2018, LADAS2018, ICCS2018, Dyno 2018, COMPLEX NETWORKS 2018, NetSci 2018.
- Christophe Crespelle was a member of the program committee of the 10th International Colloquium on Graph Theory and Combinatorics (ICGT 2018).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Isabelle Guérin Lassous is member of the editorial boards of Computer Communications (Elsevier), Ad Hoc Networks (Elsevier) and Discrete Mathematics & Computer Science.
- Anthony Busson is member of the editorial boards of Computer Communications (Elsevier).
- Márton Karsai is member of the editorial boards of Advances in Complex Systems (World Scientific).
- Márton Karsai was the invited editor of the Special Issue on Complex Networks and Their Applications, Applied Network Science, Springer (2018)
- Philippe Nain is a member of the Advisor Board of Performance Evaluation (Elsevier).

10.1.3.2. Reviewer - Reviewing Activities

- Márton Karsai was acting as a reviewer for the journals of Nature Communications, PNAS, PRL, PRX, PRE, Scientific Reports, EPJ Data Science, SNAM, EPL, EPJ B, PLoS One, Journal of Statistical Mechanics, Physics Letter A, Advanced in Complex Systems, Journal of Physics: Condensed Matter, Complex Networks, New Journal of Physics, Physica Scripta, Network Science; and for project agencies as ANR, FET-H2020 RIA.
- Paulo Gonçalves was reviewer for the following journals: IEEE Transactions on Signal Processing, IEEE Signal Processing Letter, Signal Processing (Elsevier).
- Christophe Crespelle was a reviewer for Algorithmica (Springer) and Theoretical Computer Science (Elsevier).

10.1.4. Invited Talks

- Éric Guichard gave a talk at the one-day conference *Humanités et numérique: vers plus de confiance dans la production du discours scientifique*: Les humanités numériques, un slogan creux? (Paris, Dec 4, 2018).
- Éric Guichard gave a talk at the international conference in honor of Gérard Noiriel, called *Héritages et actualités de la socio-histoire*: Comment l'informatique et les statistiques ont aidé à penser l'histoire des catégories juridiques et sociales de l'immigration? (June 14, 2018, Ehess, Paris).
- Éric Guichard gave a talk at the seminar of the PhD students of LIRIS (Univ. Lyon-1): Numérique et fracture sociale (April 3rd, 2018.)
- Éric Guichard gave a talk at the conference « Quelle est l'utilité d'un produit digital connecté dans le monde du sport? », rencontres *maths et industrie* (institut Henri Poincaré, Paris, March 13, 2018).

- Éric Guichard gave a talk at the one-day conference *Imaginaires présents, numérique à venir: impensés et idéologies*: Culture de l'écrit spatial, imaginaires et subjectivités (ESAD St-Etienne, March 7, 2018).
- Éric Guichard gave a talk at the *Séminaire interdisciplinaire PHITECO (Philosophie, Technique et Cognition)*: Epistémologie et philosophie politique du numérique (UTC, Compiègne, Jan 19, 2018).
- Paulo Gonçalves gave a talk at the seminars of the EPC Inria Maracas (Nov. 2018).
- Isabelle Guérin Lassous gave a talk at CUST and NUST in Pakistan and a seminar at the LIG laboratory on Wi-Fi association.
- Thomas Begin gave a talk at DIVA lab of Univ. of Ottawa in Canada on the modeling of DPDK-based virtual switches.
- Thomas Begin gave a talk at CITI Lab of INSA Lyon on Modeling WiFi in a multihop wireless network : feedback on experience.
- Thomas Begin gave a talk at LIP Lab of ENS Lyon on A brief tour of Machine Learning techniques.
- Philippe Nain gave a seminar on A class of stochastic multilayer networks: percolation, exact and asymptotic results, Inria Grenoble (March 22, 2018).
- Márton Karsai gave an invited talk at the Higher Order Models in Network Science, NetSci'18 Satellite (12 June 2018, Paris, France)
- Márton Karsai gave an invited talk at the Art, Networks and Technology, NetSci'18 Satellite (12 June 2018, Paris, France)
- Márton Karsai gave an invited talk at the BrainTime workshop, Institut de Neurosciences de la Timone (18 September 2018, Marseille, France)
- Márton Karsai gave an invited talk at the Language Seminar Series, Laboratoire d'Informatique de Grenoble (20 September 2018, Grenoble, France)
- Márton Karsai gave an invited talk at the Op-La-Dyn workshop, CCS'18 Satellite (26 September 2018, Thessaloniki, Greece)
- Márton Karsai gave an invited talk at the Complex systems for the most vulnerable – UNICEF workshop, CCS'18 Satellite (27 September 2018, Thessaloniki, Greece)
- Márton Karsai gave an invited talk at the SpaceNet workshop, CCS'18 Satellite (27 September 2018, Thessaloniki, Greece)
- Christophe Crespelle gave an invited talk in the Workshop on Graph Theory and Applications at the Vietnam Institute for Advanced Studies in Mathematics (VIASM), Hanoi, Vietnam.
- Christophe Crespelle gave a talk at the ICTLab of the University of Science and Technology of Hanoi (USTH), Vietnam.

10.1.5. Leadership within the Scientific Community

- Éric Fleury is Co-chair of the Networking group ResCom of the CNRS GDR ASR. He is also a member of the scientific committee of the GDR ASR.
- Philippe Nain is the coordinator of the "Strategic Technology Monitoring & Prospective Studies Inria Unit".

10.1.6. Scientific Expertise

- Isabelle Guérin Lassous is a member of the research committee of the Milyon labex.
- Isabelle Guérin Lassous is the president of the HCERES evaluation committee of SAMOVAR.
- Éric Fleury is member of the Inria Advanced and starting research position jury and junior research position (CR2/CR1)
- Éric Fleury has been an expert for the Fund for Scientific Research - FNRS.

- Éric Fleury has been a member of evaluation panels as part of the French National Research Agencys (ANR) and member of the program committee of the ANR ROSE Challenge (RObotique et Capteurs au Service d'Ecophyto)
- Éric Fleury is member of the Inria Evaluation Committee.
- Éric Guichard is a member of the scientific committee of LASCO Idea Lab of the IMT (LAboratoire Sens et COmpréhension du monde contemporain de l'Institut Mines-Télécom).
- Éric Guichard is a member of the international evaluation board of the doctoral program *Filosofia da Ciencia, Tecnologia, Arte e Sociedade* of the University of Lisbon.
- Éric Guichard is the manager of the RAIL [0] (Réseau de l'Atelier Internet Lyonnais), founded in 2017 and supported by IXXI and Enssib.

10.1.7. Research Administration

- Paulo Gonçalves is scientific liaison officer for international relations in Inria Research Centre of Rhône-Alpes.
- Paulo Gonçalves is a member of the executive committee of the Milyon labex and referent for its valorisation committee.
- Paulo Gonçalves is correspondant for the theme "Big Data" of the *Fédération d'Informatique de Lyon*.
- Paulo Gonçalves is member of the Council of the LIP laboratory.
- Isabelle Guérin Lassous is member of the department council of the Computer Science department of Université Lyon 1.
- Isabelle Guérin Lassous is the managing director of the Foundation Blaise Pascal.
- Anthony Busson is member of the Thesis Commission at LIP.
- Anthony Busson is head of the computer science department at IUT (Institut Universitaire de Technologie) - Université Lyon Claude Bernard Lyon 1.
- Anthony Busson was a HCERES expert member (laboratory LMIA-MIPS - Université de Haute Alsace).
- Thomas Begin is an elected member of the Council of the LIP laboratory.
- Thomas Begin is an elected member of the department council of the Computer Science department of Université Lyon 1.
- Jean Pierre Chevrot is member of the steering committee of the IXXI - Rhône-Alpes Complex Systems Institute.
- Márton Karsai is the co-responsible for the M2 master program in Modelling of Complex Systems at ENS Lyon
- Márton Karsai is the elected council member of the Complex System Society (2015-)
- Márton Karsai is the elected member of executive committee of the Complex System Society (2018-)
- Márton Karsai is the elected member of the steering committee of the IXXI Complex System Institute (2017-)
- Márton Karsai is the member of the computational infrastructure board of LIP
- Márton Karsai is the member of the communication board of LIP
- Éric Guichard is a member of the steering committee of the IXXI Complex System Institute
- Jean-Philippe Magué is a member of the executive committee of the IXXI Complex System Institute
- Jean-Philippe Magué is a member of the executive committee of the Aslan Labex, in charge of the language complexity work package.

[0] http://barthes.enssib.fr/RAIL

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence :

Thomas Begin: Computer Networks, 18.5h, L3, Université Lyon 1.

Márton Karsai: Introduction to Complex Networks, 6h, L3, ENS Lyon

Master :

Thomas Begin: Distributed Algorithms, 18h, M1, Université Lyon 1.

Thomas Begin: Computer Networks, 44h, M1, Université Lyon 1.

Thomas Begin: System Administration & Security, 10.5h, M2, Université Lyon 1.

Thomas Begin: Advanced Networks, 58h, M2, Université Lyon 1.

Thomas Begin: Cloud Computing, 9h, M2, Université Lyon 1.

Isabelle Guérin-Lassous: Distributed Algorithms, 30h, M1, Université Lyon 1.

Isabelle Guérin-Lassous: Networking, 14h, M1, Université Lyon 1.

Isabelle Guérin-Lassous: Wireless networks, 9h, M2, Université Lyon 1.

Isabelle Guérin-Lassous: Quality of Service, 5h, M2, Université Lyon 1.

Isabelle Guérin-Lassous: ToIP and streaming, 12h, M1, Université Lyon 1.

Éric Guichard: Économie du web et du document, 36h, M2, Enssib & Univ. Lyon 1.

Éric Guichard: Programmation éditoriale, 18h, M2, Enssib & Univ. Lyon 1.

Márton Karsai: Complex Networks, 36h, M2, ENS Lyon

Márton Karsai: Data Bases Data Mining, 20h, M1, ENS Lyon

Philippe Nain: Performance Evaluation of Computing and Communication Systems (graduate course, Fall semester 2018), 32h, University of Massachusetts, Amherst, USA.

Cycle ingénieur (Bac+3 à Bac+5):

Paulo Gonçalves: Traitement du Signal (déterministe, aléatoire, numérique), Estimation statistique. 80 heures Eq. TD. CPE Lyon, France.

Marion Foare: Traitement du Signal (déterministe, aléatoire), Traitement d'images, Compression, Projets. 115 heures Eq. TD. CPE Lyon, France.

10.2.2. Supervision

HDR defense: Thomas BEGIN, Contributions to the Performance Modeling of Computer Networks. December 2018.

PhD defense: Mohammed AMER,WiFi network management: a SDN approach, Novembre 2018.

PhD in progress: Esteban BAUTISTA RUIZ, Statistical Graph Signal Processing. P. Gonçalves (P. Abry, co-advisor). Started Sept. 1st, 2016.

PhD in progress: Mohamed Adbelwedoud LAFDAL, Inference of conflict graph in IEEE 802.11 networks. September 2017, A. Busson and I. Guérin Lassous

PhD in progress: Samuel UNICOMB, Spreading processes on temporal networks, Oct 2016, E. Fleury and M. Karsai

PhD in progress: Jacobo Levy ABITBOL, Information diffusion and language evolution on dynamical social networks, Oct 2016, E. Fleury and M. Karsai

PhD in progress: Marija STOJANOVA, Performance Modelling of IEEE 802.11 networks , Oct 2016, T. Begin

PhD in progress: Sicheng DAI, Dynamic Multilayer Network Modelling, M. Karsai supervisor (E. Fleury director). Started October 1st, 2017.

PhD in progress: Gaetan FRUSQUE, *Modal Decompositions of Dynamic Graphs : Application in Neurosciences*, P. Gonçalves (P. Borgnat, co-advisor). Started October 1st, 2017.

PhD in progress: Rémy GRÜNBLATT, Controlled mobility for UAV networks, October 2017, I. Guérin Lassous and O. Simonin.

PhD in progress: Dominique BARBE, From local to global learning, P. Gonçalves (P. Borgnat, co-advisor). Started October 1st, 2018.

PhD in progress: Nour el Houda BOUZOUITA, supervised by A. Busson and Hervé Rivano. Wi-Fi network Optimization through crowd sensing applications. November 2018 - November 2021.

10.2.3. Juries

Paulo Gonçalves was a member of the Ph.D thesis examination board of Jean-Charles Vialatte, IMT Atlantique, Université Bretagne Loire, December 2018.

Isabelle Guérin Lassous was a reviewer of the Ph.D thesis examination board of Jonatan Krolikowski, Univ. Paris-Sud ; Raphaël Naves, Univ. de Toulouse ; Antoine Auger, Univ. de Toulouse

Isabelle Guérin Lassous was a president of the Ph.D thesis examination board of Elodie Morin, Univ. Grenoble Alpes ; Jonatan Krolikowski, Univ. Paris-Sud

Isabelle Guérin Lassous was a member of the Ph.D thesis examination board of Narcisse Kamtchoun, UPMC

Isabelle Guérin Lassous was a member of the HdR thesis examination board of Valeria Loscri, Univ. de Lille.

Éric Guichard was a member of the Ph.D thesis examination board of Amélie Turet, Univ. Bordeaux-Montaigne.

Christophe Crespelle was a reviewer of the Ph.D thesis examination board of Thibaud Arnoux, UPMC

Anthony Busson was a jury member of Cristhian IZA PAREDES's PhD thesis. Universitat Polytecnica de Catalunya. July 12 July 2018.

Anthony Busson was reviewer and jury member of Lam-Thanh TU's PhD thesis. Université Paris Saclay. 18 June 2018.

Anthony Busson was reviewer and jury member of Nesrine KHERNANE's PhD thesis. Université de Bourgogne Franche comté. 13 November 2018.

Anthony Busson reviewer and jury member of Lucas's RIVOIRARD PhD thesis. Université de Lille. 21 September 2018.

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

Isabelle Guérin Lassous is the managing director of the Foundation Blaise Pascal. The objectives of the foundation are to promote mathematics and computer science and to attract young people to scientific fields like computer science and mathematics. The actions of the FBP focus on:

- a support to actors that promote mathematics and computer science via allocated funding based on call of proposals;
- a structuring of actors to increase the impacts of their actions, to coordinate the efforts and to share experiences;
- a development of innovative experiences via summer camps and clubs of mathematics and computer science.

10.3.2. Articles and contents

- Éric Guichard wrote an article with Nicolas Schabanel in newspaper *Libération*, September 10, 2018, p. 24: « Copyright numérique: stériliser pour mieux tuer », https://www.liberation.fr/debats/2018/09/09/copyright-numerique-steriliser-pour-mieux-tuer_1677568.
- Éric Guichard gave an interview in newspaper *O Globo* (Brazil), Jan 1st , 2018.

10.3.3. Interventions

Paulo Gonçalves and Eric Guichard animated a citizen meeting on Artificial Intelligence and Social impacts. November 2018. Orliénas. https://www.vepm.net/2018/10/26/cafe-citoyen-orlienas-9-novembre-2018-lintelligence-artificielle/

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] T. BEGIN.*Contributions to the Performance Modeling of Computer Networks*, Université Claude Bernard Lyon 1, December 2018, Habilitation à diriger des recherches, https://hal.archives-ouvertes.fr/tel-01951250

Articles in International Peer-Reviewed Journal

[2] I. ACHOUR, T. BEJAOUI, A. BUSSON, S. TABBANE.*Performance Modeling of SEAD Protocol in Vehicular Environment*, in "Wireless Networks", 2018, p. 1-18 [*DOI :* 10.1007/s11276-018-1706-3], https://hal.inria.fr/hal-01782477

[3] M. AMER, I. GUÉRIN LASSOUS, A. BUSSON.*Association optimization based on access fairness for Wi-Fi networks*, in "Computer Networks", June 2018, vol. 137, p. 173 - 188 [*DOI :* 10.1016/J.COMNET.2018.03.004], https://hal.inria.fr/hal-01906319

[4] T. BEGIN, B. BAYNAT, G. ARTERO GALLARDO, V. JARDIN.*An accurate and efficient modeling framework for the performance evaluation of DPDK-based virtual switches*, in "IEEE Transactions on Network and Service Management", 2018, https://hal.archives-ouvertes.fr/hal-01901493

[5] A. BRANDWAJN, T. BEGIN.*First-Come-First-Served Queues with Multiple Servers and Customer Classes*, in "Performance Evaluation", 2018, https://hal.archives-ouvertes.fr/hal-01912975

[6] A. BRANDWAJN, T. BEGIN, H. CASTEL-TALEB, T. ATMACA.*A study of systems with multiple operating levels, probabilistic thresholds and hysteresis*, in "IEEE Transactions on Parallel and Distributed Systems", April 2018, vol. 29, no 4, p. 748-757 [*DOI :* 10.1109/TPDS.2017.2773496], https://hal.inria.fr/hal-01670974

[7] L. BUSON, A. NARDY, D. MULLER, J.-P. CHEVROT.*The sociolinguistic repetition task: A new paradigm for exploring the cognitive coherence of language varieties*, in "Topics in cognitive science", 2018, vol. 10, no 4, p. 803-817 [*DOI :* 10.1111/TOPS.12380], https://hal.inria.fr/hal-01674578

[8] A. BUSSON, I. LAHSEN-CHERIF.*Impact of Resource Blocks Allocation Strategies on Downlink Interference and SIR Distributions in LTE Networks: A Stochastic Geometry Approach*, in "Wireless Communications and Mobile Computing", June 2018, vol. 2018, p. 1 - 15 [*DOI :* 10.1155/2018/9163783], https://hal.inria.fr/hal-01906324

[9] C. CRESPELLE, A. PEREZ, I. TODINCA.*An $O(n^2)$ time algorithm for the minimal permutation completion problem*, in "Discrete Applied Mathematics", 2018, https://hal.inria.fr/hal-01969498

[10] M. DEHGHAN, W. CHU, P. NAIN, D. TOWSLEY, Z.-L. ZHANG.*Sharing Cache Resources among Content Providers: A Utility-Based Approach*, in "IEEE/ACM Transactions on Networking", 2019, vol. 40, n^o 8, https://hal.inria.fr/hal-01672961

[11] A. DUVAL, T. OBADIA, L. MARTINET, P.-Y. BOËLLE, É. FLEURY, D. GUILLEMOT, L. OPATOWSKI, L. TEMIME.*Measuring dynamic social contacts in a rehabilitation hospital: effect of wards, patient and staff characteristics*, in "Scientific Reports", January 2018, vol. 8, n^o 1, 1686 [*DOI :* 10.1038/s41598-018-20008-w], https://hal.inria.fr/hal-01696186

[12] I. GUÉRIN LASSOUS, I. FAUS MARTIN, L. URQUIZA AGUIAR, M. AGUILAR IGARTUA.*Transient analysis of idle time in VANETs using Markov-reward models*, in "IEEE Transactions on Vehicular Technology", 2018, vol. 67, n^o 4, p. 2833-2847 [*DOI :* 10.1109/TVT.2017.2766449], https://hal.inria.fr/hal-01671587

[13] B. JIANG, P. NAIN, D. TOWSLEY.*On the Convergence of the TTL Approximation for an LRU Cache under Independent Stationary Request Processes*, in "ACM Transactions on Modeling and Performance Evaluation of Computing Systems", 2018, vol. 3, n^o 4, https://hal.inria.fr/hal-01673272

[14] M. KIVELÄ, J. CAMBE, J. SARAMÄKI, M. KARSAI.*Mapping temporal-network percolation to weighted, static event graphs*, in "Scientific Reports", December 2018, vol. 8, n^o 1 [*DOI :* 10.1038/s41598-018-29577-2], https://hal.inria.fr/hal-01952142

[15] Y. LÉO, C. CRESPELLE, É. FLEURY.*Non-altering time scales for aggregation of dynamic networks into series of graphs*, in "Computer Networks", 2018, https://hal.inria.fr/hal-01969504

[16] Y. LÉO, M. KARSAI, C. SARRAUTE, É. FLEURY.*Correlations and dynamics of consumption patterns in social-economic networks*, in "Social Network Analysis and Mining", January 2018, vol. 8, p. 1-16, https://arxiv.org/abs/1801.08856 [*DOI :* 10.1007/s13278-018-0486-1], https://hal.inria.fr/hal-01697280

[17] L. MARTINET, C. CRESPELLE, É. FLEURY, P.-Y. BOËLLE, D. GUILLEMOT.*The link stream of contacts in a whole hospital*, in "Social Network Analysis and Mining", December 2018, vol. 8, n^o 59 [*DOI :* 10.1007/s13278-018-0535-9], https://hal.inria.fr/hal-01969491

[18] S. NAIMI, A. BUSSON, V. VÈQUE, R. BOUALLEGUE.*Metric Anticipation to manage Mobility in Mobile Mesh and ad-hoc Wireless Networks*, in "Annals of Telecommunications - annales des télécommunications", December 2018, vol. 73, n^o 11-12, p. 787–802 [*DOI :* 10.1007/s12243-018-0666-z], https://hal.inria.fr/hal-01940198

[19] K. REIN, T. VENTURINI.*Ploughing digital landscapes: How Facebook influences the evolution of live video streaming*, in "New Media and Society", 2018, vol. 20, n^o 9, p. 3359-3380 [*DOI :* 10.1177/1461444817748954], https://hal.archives-ouvertes.fr/hal-01672284

[20] M. STOJANOVA, T. BEGIN, A. BUSSON.*Conflict graph-based model for IEEE 802.11 networks: A Divide-and-Conquer approach*, in "Performance Evaluation", 2018, https://hal.archives-ouvertes.fr/hal-01955177

[21] S. UNICOMB, G. IÑIGUEZ, M. KARSAI.*Threshold driven contagion on weighted networks*, in "Scientific Reports", February 2018, vol. 8, n^{o} 1, p. 1-10 [*DOI :* 10.1038/S41598-018-21261-9], https://hal.inria.fr/hal-01831490

[22] Q. ZHANG, M. KARSAI, A. VESPIGNANI.*Link transmission centrality in large-scale social networks*, in "EPJ Data Science", December 2018, vol. 7, n^{o} 1 [*DOI :* 10.1140/EPJDS/S13688-018-0162-8], https://hal.inria.fr/hal-01952143

International Conferences with Proceedings

[23] A. BUSSON, I. G. LASSOUS, H. P. THE ANH.*Channel assignment in IEEE 802.11-based substitution networks*, in "BalkanCom 2018 - Second International Balkan Conference on Communications and Networking", Podgorica, Montenegro, June 2018, p. 1-5, https://hal.inria.fr/hal-01906335

[24] M. FOARE, N. PUSTELNIK, L. CONDAT.*A new proximal method for joint image restoration and edge detection with the Mumford-Shah model*, in "IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018)", Calgary, Alberta, Canada, April 2018, https://hal.archives-ouvertes.fr/hal-01736997

[25] R. FONTUGNE, E. BAUTISTA, C. PETRIE, Y. NOMURA, P. ABRY, P. GONÇALVES, K. FUKUDA, E. ABEN.*BGP Zombies: an Analysis of Beacons Stuck Routes*, in "20th Passive and Active Measurements Conference (PAM 2019)", Puerto Varas, Chile, March 2019, https://hal.inria.fr/hal-01970596

[26] I. GUÉRIN LASSOUS, A. BUSSON, M. AMER.*Association Optimization in Wi-Fi Networks based on the Channel Busy Time Estimation*, in "IFIP Networking", Zürich, Switzerland, 2018, https://hal.inria.fr/hal-01965313

[27] I. GUÉRIN LASSOUS, A. BUSSON, M. AMER.*Considering Frame Aggregation in Association Optimization for High Throughput Wi-Fi Networks*, in "ACM PE-WASUN", Montreal, Canada, 2018, https://hal.inria.fr/hal-01965315

[28] B. JIANG, P. NAIN, D. TOWSLEY, S. GUHA.*On a Class of Stochastic Multilayer Networks*, in "ACM Sigmetrics", Irvine, CA, United States, Proceedings of the ACM on Measurement and Analysis of Computing Systems (POMACS), June 2018, vol. 2, n^{o} 1, p. 1-24, https://hal.inria.fr/hal-01669368

[29] J. LEVY ABITBOL, M. KARSAI, J.-P. MAGUÉ, J.-P. CHEVROT, É. FLEURY.*Socioeconomic dependencies of linguistic patterns in Twitter: a multivariate analysis*, in "WWW '18 - World Wide Web Conference", Lyon, France, ACM, April 2018, p. 1125-1134 [*DOI :* 10.1145/3178876.3186011], https://hal.inria.fr/hal-01674620

[30] M. TSITSVERO, P. BORGNAT, P. GONÇALVES.*Multidimensional analytic signal with application on graphs*, in "SSP 2018 - IEEE Workshop on Statistical Signal Processing", Freiburg, Germany, IEEE, June 2018, p. 1-5, https://hal.inria.fr/hal-01878753

Conferences without Proceedings

[31] T. BEGIN, A. BUSSON, I. GUÉRIN LASSOUS, A. BOUKERCHE.*Video on Demand in IEEE 802.11p-based Vehicular Networks: Analysis and Dimensioning*, in "MSWiM'18 - 21st ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems", Montreal, Canada, October 2018 [*DOI :* 10.1145/3242102.3242109], https://hal.archives-ouvertes.fr/hal-01858019

[32] H. BOUCHET, M. KARSAI, S. BARBU, J.-P. CHEVROT, É. FLEURY, A. NARDY.*Recueil de données massives de contacts sociaux en école maternelle : validation éthologique de l'utilisation des capteurs de proximité*, in "SFECA 2018 - 48ème Colloque Annuel de la Socété Française pour l'Etude du Comportement Animal", Rennes, France, Université de Rennes 1 and Société Française pour l'Etude du Comportement Animal, June 2018, https://hal-univ-rennes1.archives-ouvertes.fr/hal-01826302

[33] J.-P. CHEVROT, J. LEVY ABITBOL, M. KARSAI, J.-P. MAGUÉ, É. FLEURY.*Variations du (ne) négatif du français dans Twitter. Que peut apporter l'étude des données massives aux questions de sociolinguistique ?*, in "XXIXe Congrès international de linguistique et de philologie romanes", Copenhague, Denmark, July 2019, https://hal.archives-ouvertes.fr/hal-01970511

[34] S. DE NIGRIS, E. BAUTISTA, P. ABRY, K. AVRACHENKOV, P. GONÇALVES.*Fractional graph-based semi-supervised learning*, in "International School and Conference on Network Science", Paris, France, June 2018, 1, https://hal.inria.fr/hal-01877500

[35] J. LEVY ABITBOL, M. KARSAI, J.-P. CHEVROT, J.-P. MAGUÉ, É. FLEURY.*Socioeconomic and network dependencies of linguistic patterns in Twitter*, in "4th Annual International Conference on Computational Social Science", Evanston, Illinois, United States, July 2018, https://hal.archives-ouvertes.fr/hal-01970536

Scientific Books (or Scientific Book chapters)

[36] J.-P. CHEVROT, A. GHIMENTON.*Bilingualism and bidialectalism*, in "The Cambridge Handbook of Bilingualism", L. ORTEGA, A. DE HOUWER (editors), Cambridge University Press, 2018, https://hal.inria.fr/hal-01674604

[37] É. GUICHARD.*L'histoire et l'écriture numérique : approche technique, politique, épistémologique*, in "Histoire et numérique", S. LAMASSÉ, G. BONNOT (editors), Presses de la Sorbonne, 2019, p. 1-18, https://hal.archives-ouvertes.fr/hal-01673099

[38] É. GUICHARD.*Proximités intellectuelles entre François Dagognet et Jack Goody à l'heure de l'internet*, in "Science, technique, société: hommage à François Dagognet", J.-C. BEAUNE, M. SALHAB (editors), Presses de l'Université Libano-française, 2019, p. 1-16, https://hal.archives-ouvertes.fr/hal-01672738

[39] G. IÑIGUEZ, Z. RUAN, K. KASKI, J. KERTÉSZ, M. KARSAI.*Service Adoption Spreading in Online Social Networks*, in "Complex Spreading Phenomena in Social Systems", June 2018, p. 151-175 [*DOI :* 10.1007/978-3-319-77332-2_9], https://hal.inria.fr/hal-01952149

[40] M. KARSAI, H.-H. JO, K. KASKI.*Bursty Human Dynamics*, Springer, January 2018, XIV, 121 [*DOI :* 10.1007/978-3-319-68540-3], https://hal.inria.fr/hal-01675082

[41] Y. LIAO, W. DU, M. KARSAI, C. SARRAUTE, M. MINNONI, É. FLEURY.*Prepaid or Postpaid? That Is the Question: Novel Methods of Subscription Type Prediction in Mobile Phone Services*, in "Social Network Based Big Data Analysis and Applications", Lecture Notes in Social Networks, Springer, May 2018, p. 165-181 [*DOI :* 10.1007/978-3-319-78196-9_8], https://hal.inria.fr/hal-01831494

[42] Y. LIAO, W. DU, M. KARSAI, C. SARRAUTE, M. MINNONI, É. FLEURY.*Prepaid or Postpaid? That Is the Question: Novel Methods of Subscription Type Prediction in Mobile Phone Services*, in "Social Network Based Big Data Analysis and Applications", May 2018, p. 165-181 [*DOI :* 10.1007/978-3-319-78196-9_8], https://hal.inria.fr/hal-01952151

[43] N. TREMBLAY, P. GONÇALVES, P. BORGNAT.*Design of graph filters and filterbanks*, in "Cooperative and Graph Signal Processing", 2018, https://arxiv.org/abs/1711.02046 , https://hal.inria.fr/hal-01675375

[44] T. VENTURINI, M. JACOMY, L. BOUNEGRU, J. GRAY.*Visual Network Exploration for Data Journalists*, in "The Routledge Handbook to Developments in Digital Journalism Studies (forthcoming)", Routledge, 2018, https://hal.archives-ouvertes.fr/hal-01672282

[45] T. VENTURINI, A. K. MUNK, M. JACOMY.*Actor-Network VS Network Analysis VS Digital Networks Are We Talking About the Same Networks?*, in "DigitalSTS: A Handbook and Fieldguide", 2018, https://hal.archives-ouvertes.fr/hal-01672289

[46] T. VENTURINI, A. MUNK, A. MEUNIER.*Data-Sprinting: a Public Approach to Digital Research*, in "Interdisciplinary Research Methods", 2018, https://hal.archives-ouvertes.fr/hal-01672288

[47] T. VENTURINI.*The Fish Tank Complex of Social Modelling on Space and Time in Understanding Collective Dynamics*, in "Frontiers of Social Science: A Philosophical Reflection", 2018, https://hal.archives-ouvertes.fr/hal-01672283

[48] L. WENG, M. KARSAI, N. PERRA, F. MENCZER, A. FLAMINI.*Attention on Weak Ties in Social and Communication Networks*, in "Spreading Dynamics in Social Systems", 2018, https://hal.inria.fr/hal-01675086

Scientific Popularization

[49] L. BOUNEGRU, J. GRAY, T. VENTURINI, M. MAURI.*A Field Guide to Fake News and other Information Disorders*, January 2018, The Field Guide to Fake News explores the use of digital methods to trace the production, circulation and reception of "fake news" online. This document is a sample comprising of the first three chapters of the field guide. These chapters cover recipes for tracing the circulation of "fake news" on Facebook and the web, as well as its tracker signatures and techno-commercial underpinnings. Further chapters on how to map and investigate memes as vectors of "post-truth" politics and political trolling will be released in the coming months. The field guide is a project of the Public Data Lab with support from First Draft and has been produced through a series of "data sprints" and research workshops in Amsterdam, Copenhagen and Milan, https://hal.archives-ouvertes.fr/hal-01672287

Other Publications

[50] J.-C. BERMOND, D. MAZAURIC, V. MISRA, P. NAIN.*Distributed Link Scheduling in Wireless Networks*, January 2019, working paper or preprint, https://hal.inria.fr/hal-01977266

[51] M. FOARE, N. PUSTELNIK, L. CONDAT.*Semi-Linearized Proximal Alternating Minimization for a Discrete Mumford–Shah Model*, April 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01782346

[52] L. GAUVIN, M. GÉNOIS, M. KARSAI, M. KIVELÄ, T. TAKAGUCHI, E. VALDANO, C. L. VESTERGAARD.*Randomized reference models for temporal networks*, June 2018, https://arxiv.org/abs/1806.04032 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01817633

[53] Y. KALOGA, M. FOARE, N. PUSTELNIK, P. JENSEN.*Discrete Mumford-Shah on graph for mixing matrix estimation*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01948570

[54] M. TSITSVERO, P. BORGNAT, P. GONÇALVES.*Analytic signal in many dimensions*, September 2018, https://arxiv.org/abs/1712.09350v2 - Submitted to Applied and Computational Harmonic Analysis, https://hal.inria.fr/hal-01877403

[55] Q. ZHANG, M. KARSAI, A. VESPIGNANI.*Link transmission centrality in large-scale social networks*, July 2018, https://arxiv.org/abs/1802.05337 - 19 pages, 5 figures, https://hal.inria.fr/hal-01831482

References in notes

[56] R. ALBERT, A.-L. BARABÁSI.*Statistical mechanics of complex networks*, in "Reviews of Modern Physics", 2002, vol. 74

[57] A.-L. BARABÁSI.*The origin of bursts and heavy tails in human dynamics*, in "Nature", 2005, vol. 435, 207

[58] R. R. COIFMAN, S. LAFON, A. B. LEE, M. MAGGIONI, B. NADLER, F. WARNER, S. W. ZUCKER.*Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps*, in "PNAS", 2005, vol. 102, n° 21, p. 7426-7431

[59] W. J. FITZGERALD, R. L. SMITH, A. T. WALDEN.*Nonlinear and Nonstationary Signal Processing*, Cambridge University Press, Cambridge, 2001

[60] S. FORTUNATO.*Community detection in graphs*, in "Physics Reports", 2010, vol. 486, p. 75-174

[61] G. GAUMONT, M. PANAHI, D. CHAVALARIAS.*Reconstruction of the socio-semantic dynamics of political activist Twitter networks – Method and application to the 2017 French presidential election*, in "PLOS ONE", 2018, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0201879

[62] É. GUICHARD.*La philosophie des techniques revue à l'aune de l'internet et du numérique*, in "Le numérique en débat. Des nombres, des machines et des hommes", G. CHAZAL (editor), Éditions Universitaires de Dijon. Collection Sociétés, 2017, p. 173–189, Preprint:http://barthes.enssib.fr/articles/Guichard-pensee-critique-culture-numerique-philo-technique.pdf

[63] P. HOLME, J. SARAMÄKI.*Temporal networks*, in "Physics Reports", 2012, vol. 519, p. 97-125

[64] M. KARSAI, M. KIVELÄ, R. K. PAN, K. KASKI, J. KERTÉSZ, A.-L. BARABÁSI, J. SARAMÄKI.*Small But Slow World: How Network Topology and Burstiness Slow Down Spreading*, in "Phys. Rev. E", 2011, vol. 83

[65] M. KARSAI, N. PERRA.*Control Strategies of Contagion Processes in Time-Varying Networks*, in "Temporal Network Epidemiology", N. MASUDA, P. HOLME (editors), Springer, 2017 [*DOI :* 10.1007/978-981-10-5287-3], https://hal.inria.fr/hal-01675083

[66] M. KARSAI, N. PERRA, A. VESPIGNANI.*A. Random Walks and Search in Time-Varying Networks*, 2013, arXiv:1303.5966

[67] M. KIVELÄ, R. K. PAN, K. KASKI, J. KERTÉSZ, J. SARAMÄKI, M. KARSAI.*Multiscale Analysis of Spreading in a Large Communication Network*, in "J. Stat. Mech.", 2012

[68] L. KOVANEN, M. KARSAI, K. KASKI, J. KERTÉSZ, J. SARAMÄKI.*Temporal motifs in time-dependent networks*, in "J. Stat. Mech.", 2011

[69] G. KRINGS, M. KARSAI, S. BERNHARDSSON, V. BLONDEL, J. SARAMÄKI.*Effects of time window size and placement on the structure of an aggregated communication network*, in "EPJ Data Science", 2012, vol. 1, n° 4

[70] Z. Q. LUO, M. GASTPAR, J. LIU, A. SWAMI.*Distributed Signal Processing in Sensor Networks*, in "IEEE Signal Processing Mag", 2006, vol. 23

[71] B. A. MILLER, N. T. BLISS, P. J. WOLFE.*Towards Signal Processing Theory for Graphs and Non-Euclidian Data*, in "ICASSP", Dallas, IEEE, 2010

[72] G. MIRITELLO, E. MORO, R. LARA.*Dynamical strength of social ties in information spreading*, in "Phys. Rev. E", 2011, vol. 83

[73] M. E. J. NEWMAN.*Networks: An Introduction,*, Oxford University Press, 2010

[74] N. PERRA, A. BARONCHELLI, D. MOCANU, B. GONÇALVES, R. PASTOR-SATORRAS, A. VESPIGNANI.*Random Walks and Search in Time-Varying Networks*, in "Physical review letters", 2012, vol. 109

[75] N. PERRA, B. GONÇALVES, R. PASTOR-SATORRAS, A. VESPIGNANI.*Activity driven modeling of time varying networks*, in "Scientific Reports", 2012, vol. 2, n° 469

[76] D. SHUMAN, S. NARANG, P. FROSSARD, A. ORTEGA, P. VANDERGHEYNST.*The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains*, in "Signal Processing Magazine, IEEE", May 2013, vol. 30, n° 3, p. 83 - 98

[77] G. TIBELY, L. KOVANEN, M. KARSAI, K. KASKI, J. KERTÉSZ, J. SARAMÄKI.*Communities and beyond: mesoscopic analysis of a large social network with complementary methods*, in "Phys. Rev. E", 2011, vol. 83

[78] Q. WANG, É. FLEURY, T. AYNAUD, J.-L. GUILLAUME.*Communities in evolving networks: definitions, detection and analysis techniques*, in "Dynamics of Time Varying Networks", N. GANGULY, A. MUKHERJEE, B. MITRA, F. PERUANI, M. CHOUDHURY (editors), Springer, 2012, http://hal.inria.fr/hal-00746195

[79] A. S. WILLSKY.*Multiresolution Statistical Models for Signal and Image Processing*, in "Proceedings of the IEEE", 2002, vol. 90

[80] K. ZHAO, M. KARSAI, G. BIANCONI.*Entropy of Dynamical Social Networks*, in "PLoS ONE", 2011, vol. 6, n° 12

Project-Team DATAMOVE

Data Aware Large Scale Computing

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

IN PARTNERSHIP WITH:
CNRS

Institut polytechnique de Grenoble

Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Distributed and High Performance Computing

Table of contents

Project-Team DATAMOVE

Creation of the Team: 2016 January 01, updated into Project-Team: 2017 November 01

The DataMove team is located in the IMAG building on the Campus of Univ. Grenoble Alpes.

Keywords:

Computer Science and Digital Science:

A1.1.4. - High performance computing
A1.1.5. - Exascale
A2.1.10. - Domain-specific languages
A2.6.2. - Middleware
A7.1.2. - Parallel algorithms
A8.2. - Optimization

Other Research Topics and Application Domains:

B1.1.7. - Bioinformatics
B3.3.2. - Water: sea & ocean, lake & river
B5.5. - Materials

1. Team, Visitors, External Collaborators

Research Scientists

Fanny Dufosse [Inria, Researcher]
Bruno Raffin [Team leader, Inria, Senior Researcher, HDR]
Francieli Zanon Boito [Inria, Researcher, from Nov 2018]

Faculty Members

Yves Denneulin [Institut polytechnique de Grenoble, Professor, HDR]
Pierre Francois Dutot [Univ Grenoble Alpes, Associate Professor]
Gregory Mounie [Institut polytechnique de Grenoble, Associate Professor]
Olivier Richard [Univ Grenoble Alpes, Associate Professor]
Denis Trystram [Institut polytechnique de Grenoble, Professor, HDR]
Frederic Wagner [Institut polytechnique de Grenoble, Associate Professor]
Giorgio Lucarelli [Institut polytechnique de Grenoble, Associate Professor, until Aug 2018]

Technical Staff

Tristan Ezequel [Inria, until Aug 2018]
Nicolas Michon [Inria, until Sep 2018]
Clement Labastrou [Inria, from Mar 2018 until Aug 2018]
Pierre Neyron [CNRS]
Theophile Terraz [Inria]

PhD Students

Danilo Carastan Dos Santos [Universidade Federal do ABC Brazil, from May 2017]
Estelle Dirand [CEA, until Oct 2018]
Adrien Faure [ATOS/BULL, from May 2017]
Mohammed Khatiri [Faculte des sciencesU.M.P, Morocco]
Michael Mercier [ATOS/Bull - Inria, until Jan 2019]
Clement Mommessin [Institut polytechnique de Grenoble]
Valentin Reis [Institut polytechnique de Grenoble, until Aug 2018]

Salah Zrigui [Univ Grenoble Alpes]

Administrative Assistant

Annie Simon [Inria]

2. Overall Objectives

2.1. Overall Objectives

Moving data on large supercomputers is becoming a major performance bottleneck, and the situation is expected to worsen even more at exascale and beyond. Data transfer capabilities are growing at a slower rate than processing power ones. The profusion of flops available will be difficult to use efficiently due to constrained communication capabilities. Moving data is also an important source of power consumption. The DataMove team focuses on **data aware large scale computing**, investigating approaches to reduce data movements on large scale HPC machines. We will investigate data aware scheduling algorithms for job management systems. The growing cost of data movements requires adapted scheduling policies able to take into account the influence of intra-application communications, IOs as well as contention caused by data traffic generated by other concurrent applications. At the same time experimenting new scheduling policies on real platforms is unfeasible. Simulation tools are required to probe novel scheduling policies. Our goal is to investigate how to extract information from actual compute centers traces in order to replay job allocations and executions with new scheduling policies. Schedulers need information about the jobs behavior on the target machine to actually make efficient allocation decisions. We will research approaches relying on learning techniques applied to execution traces to extract data and forecast job behaviors. In addition to traditional computation intensive numerical simulations, HPC platforms also need to execute more and more often data intensive processing tasks like data analysis. In particular, the ever growing amount of data generated by numerical simulation calls for a tighter integration between the simulation and the data analysis. The goal is to reduce the data traffic and to speed-up result analysis by processing results in-situ, i.e. as closely as possible to the locus and time of data generation. Our goal is here to investigate how to program and schedule such analysis workflows in the HPC context, requiring the development of adapted resource sharing strategies, data structures and parallel analytics schemes. To tackle these issues, we will intertwine theoretical research and practical developments to elaborate solutions generic and effective enough to be of practical interest. Algorithms with performance guarantees will be designed and experimented on large scale platforms with realistic usage scenarios developed with partner scientists or based on logs of the biggest available computing platforms. Conversely, our strong experimental expertise will enable to feed theoretical models with sound hypotheses, to twist proven algorithms with practical heuristics that could be further retro-feeded into adequate theoretical models.

3. Research Program

3.1. Motivation

Today's largest supercomputers [0] are composed of few millions of cores, with performances almost reaching 100 PetaFlops [0] for the largest machine. Moving data in such large supercomputers is becoming a major performance bottleneck, and the situation is expected to worsen even more at exascale and beyond. The data transfer capabilities are growing at a slower rate than processing power ones. The profusion of available flops will very likely be underused due to constrained communication capabilities. It is commonly admitted that data movements account for 50% to 70% of the global power consumption [0]. Thus, data movements are potentially one of the most important source of savings for enabling supercomputers to stay in the commonly adopted

[0] Top500 Ranking, http://www.top500.org

[0] 10^{15} floating point operations per second

[0] SciDAC Review, 2010

energy barrier of 20 MegaWatts. In the mid to long term, non volatile memory (NVRAM) is expected to deeply change the machine I/Os. Data distribution will shift from disk arrays with an access time often considered as uniform, towards permanent storage capabilities at each node of the machine, making data locality an even more prevalent paradigm.

The proposed DataMove team will work on **optimizing data movements for large scale computing** mainly at two related levels:

- Resource allocation
- Integration of numerical simulation and data analysis

The resource and job management system (also called batch scheduler or RJMS) is in charge of allocating resources upon user requests for executing their parallel applications. The growing cost of data movements requires adapted scheduling policies able to take into account the influence of intra-application communications, I/Os as well as contention caused by data traffic generated by other concurrent applications. Modelling the application behavior to anticipate its actual resource usage on such architecture is known to be challenging, but it becomes critical for improving performances (execution time, energy, or any other relevant objective). The job management system also needs to handle new types of workloads: high performance platforms now need to execute more and more often data intensive processing tasks like data analysis in addition to traditional computation intensive numerical simulations. In particular, the ever growing amount of data generated by numerical simulation calls for a tighter integration between the simulation and the data analysis. The challenge here is to reduce data traffic and to speed-up result analysis by performing result processing (compression, indexation, analysis, visualization, etc.) as closely as possible to the locus and time of data generation. This emerging trend called *in-situ analytics* requires to revisit the traditional workflow (loop of batch processing followed by postmortem analysis). The application becomes a whole including the simulation, in-situ processing and I/Os. This motivates the development of new well-adapted resource sharing strategies, data structures and parallel analytics schemes to efficiently interleave the different components of the application and globally improve the performance.

3.2. Strategy

DataMove targets HPC (High Performance Computing) at Exascale. But such machines and the associated applications are expected to be available only in 5 to 10 years. Meanwhile, we expect to see a growing number of petaflop machines to answer the needs for advanced numerical simulations. A sustainable exploitation of these petaflop machines is a real and hard challenge that we will address. We may also see in the coming years a convergence between HPC and Big Data, HPC platforms becoming more elastic and supporting Big Data jobs, or HPC applications being more commonly executed on cloud like architectures. This is the second top objective of the 2015 US Strategic Computing Initiative [0]: *Increasing coherence between the technology base used for modelling and simulation and that used for data analytic computing*. We will contribute to that convergence at our level, considering more dynamic and versatile target platforms and types of workloads.

Our approaches should entail minimal modifications on the code of numerical simulations. Often large scale numerical simulations are complex domain specific codes with a long life span. We assume these codes as being sufficiently optimized. We will influence the behavior of numerical simulations through resource allocation at the job management system level or when interleaving them with analytics code.

To tackle these issues, we propose to intertwine theoretical research and practical developments in an agile mode. Algorithms with performance guarantees will be designed and experimented on large scale platforms with realistic usage scenarios developed with partner scientists or based on logs of the biggest available computing platforms (national supercomputers like Curie, or the BlueWaters machine accessible through our collaboration with Argonne National Lab). Conversely, a strong experimental expertise will enable to feed theoretical models with sound hypotheses, to twist proven algorithms with practical heuristics that could be further retro-feeded into adequate theoretical models.

[0] https://www.whitehouse.gov/the-press-office/2015/07/29/executive-order-creating-national-strategic-computing-initiative

A central scientific question is to make the relevant choices for optimizing performance (in a broad sense) in a reasonable time. HPC architectures and applications are increasingly complex systems (heterogeneity, dynamicity, uncertainties), which leads to consider the **optimization of resource allocation based on multiple objectives**, often contradictory (like energy and run-time for instance). Focusing on the optimization of one particular objective usually leads to worsen the others. The historical positioning of some members of the team who are specialists in multi-objective optimization is to generate a (limited) set of trade-off configurations, called *Pareto points*, and choose when required the most suitable trade-off between all the objectives. This methodology differs from the classical approaches, which simplify the problem into a single objective one (focus on a particular objective, combining the various objectives or agglomerate them). The real challenge is thus to combine algorithmic techniques to account for this diversity while guaranteeing a target efficiency for all the various objectives.

The DataMove team aims to elaborate generic and effective solutions of practical interest. We will make our new algorithms accessible through the team flagship software tools, **the OAR batch scheduler and the in-situ processing framework FlowVR**. We will maintain and enforce strong links with teams closely connected with large architecture design and operation (CEA DAM, BULL, Argonne National Lab), as well as scientists of other disciplines, in particular computational biologists, with whom we will elaborate and validate new usage scenarios (IBPC, CEA DAM, EDF).

3.3. Research Directions

DataMove research activity is organised around three directions. When a parallel job executes on a machine, it triggers data movements through the input data it needs to read, the results it produces (simulation results as well as traces) that need to be stored in the file system, as well as internal communications and temporary storage (for fault tolerance related data for instance). Modeling in details the simulation and the target machines to analyze scheduling policies is not feasible at large scales. We propose to investigate alternative approaches, including learning approaches, to capture and model the influence of data movements on the performance metrics of each job execution to develop **Data Aware Batch Scheduling** models and algorithms (Sec. 4.1). Experimenting new scheduling policies on real platforms at scale is unfeasible. Theoretical performance guarantees are not sufficient to ensure a new algorithm will actually perform as expected on a real platform. An intermediate evaluation level is required to probe novel scheduling policies. The second research axe focuses on the **Empirical Studies of Large Scale Platforms** (Sec. 4.2). The goal is to investigate how we could extract from actual computing centers traces information to replay the job allocations and executions on a simulated or emulated platform with new scheduling policies. Schedulers need information about jobs behavior on target machines to actually be able to make efficient allocation decisions. Asking users to caracterize jobs often does not lead to reliable information. The third research direction **Integration of High Performance Computing and Data Analytics** (Sec. 4.3) addresses the data movement issue from a different perspective. New data analysis techniques on the HPC platform introduce new type of workloads, potentially more data than compute intensive, but could also enable to reduce data movements by directly enabling to pipe-line simulation execution with a live analysis of the produced results. Our goal is here to investigate how to program and schedule such analysis workflows in the HPC context.

4. Application Domains

4.1. Data Aware Batch Scheduling

Large scale high performance computing platforms are becoming increasingly complex. Determining efficient allocation and scheduling strategies that can adapt to technological evolutions is a strategic and difficult challenge. We are interested in scheduling jobs in hierarchical and heterogeneous large scale platforms. On such platforms, application developers typically submit their jobs in centralized waiting queues. The job management system aims at determining a suitable allocation for the jobs, which all compete against each other for the available computing resources. Performances are measured using different classical metrics like maximum completion time or slowdown. Current systems make use of very simple (but fast) algorithms that however rely on simplistic platform and execution models, and thus, have limited performances.

For all target scheduling problems we aim to provide both theoretical analysis and complementary analysis through simulations. Achieving meaningful results will require strong improvements on existing models (on power for example) and the design of new approximation algorithms with various objectives such as stretch, reliability, throughput or energy consumption, while keeping in focus the need for a low-degree polynomial complexity.

4.1.1. Algorithms

The most common batch scheduling policy is to consider the jobs according to the First Come First Served order (FCFS) with backfilling (BF). BF is the most widely used policy due to its easy and robust implementation and known benefits such as high system utilization. It is well-known that this strategy does not optimize any sophisticated function, but it is simple to implement and it guarantees that there is no starvation (i.e. every job will be scheduled at some moment).

More advanced algorithms are seldom used on production platforms due to both the gap between theoretical models and practical systems and speed constraints. When looking at theoretical scheduling problems, the generally accepted goal is to provide polynomial algorithms (in the number of submitted jobs and the number of involved computing units). However, with millions of processing cores where every process and data transfer have to be individually scheduled, polynomial algorithms are prohibitive as soon as the polynomial degree is too large. The model of *parallel tasks* simplifies this problem by bundling many threads and communications into single boxes, either rigid, rectangular or malleable. Especially malleable tasks capture the dynamicity of the execution. Yet these models are ill-adapted to heterogeneous platforms, as the running time depends on more than simply the number of allotted resources, and some of the common underlying assumptions on the speed-up functions (such as monotony or concavity) are most often only partially verified.

In practice, the job execution times depend on their allocation (due to communication interferences and heterogeneity in both computation and communication), while theoretical models of parallel jobs usually consider jobs as black boxes with a fixed (maximum) execution time. Though interesting and powerful, the classical models (namely, synchronous PRAM model, delay, LogP) and their variants (such as hierarchical delay), are not well-suited to large scale parallelism on platforms where the cost of moving data is significant, non uniform and may change over time. Recent studies are still refining such models in order to take into account communication contentions more accurately while remaining tractable enough to provide a useful tool for algorithm design.

Today, all algorithms in use in production systems are oblivious to communications. One of our main goals is to **design a new generation of scheduling algorithms fitting more closely job schedules according to platform topologies**.

4.1.2. Locality Aware Allocations

Recently, we developed modifications of the standard back-filling algorithm taking into account platform topologies. The proposed algorithms take into account locality and contiguity in order to hide communication patterns within parallel tasks. The main result here is to establish good lower bounds and small approximation ratios for policies respecting the locality constraints. The algorithms work in an online fashion, improving the global behavior of the system while still keeping a low running time. These improvements rely mainly on our past experience in designing approximation algorithms. Instead of relying on complex networking models and communication patterns for estimating execution times, the communications are disconnected from the execution time. Then, the scheduling problem leads to a trade-off: optimizing locality of communications on one side and a performance objective (like the makespan or stretch) on the other side.

In the perspective of taking care of locality, other ongoing works include the study of schedulers for platforms whose interconnection network is a static structured topology (like the 3D-torus of the BlueWaters platform we work on in collaboration with the Argonne National Laboratory). One main characteristic of this 3D-torus platform is to provide I/O nodes at specific locations in the topology. Applications generate and access specific data and are thus bounded to specific I/O nodes. Resource allocations are constrained in a strong and unusual way. This problem is close for actual hierarchical platforms. The scheduler needs to compute a schedule such that I/O nodes requirements are filled for each application while at the same time avoiding communication interferences. Moreover, extra constraints can arise for applications requiring accelerators that are gathered on the nodes at the edge of the network topology.

While current results are encouraging, they are however limited in performance by the low amount of information available to the scheduler. We look forward to extend ongoing work by progressively increasing application and network knowledge (by technical mechanisms like profiling or monitoring or by more sophisticated methods like learning). It is also important to anticipate on application resource usage in terms of compute units, memory as well as network and I/Os to efficiently schedule a mix of applications with different profiles. For instance, a simple solution is to partition the jobs as "communication intensive" or "low communications". Such a tag could be achieved by the users them selves or obtained by learning techniques. We could then schedule low communications jobs using leftover spaces while taking care of high communication jobs. More sophisticated options are possible, for instance those that use more detailed communication patterns and networking models. Such options would leverage the work proposed in Section 4.2 for gathering application traces.

4.1.3. Data-Centric Processing

Exascale computing is shifting away from the traditional compute-centric models to a more data-centric one. This is driven by the evolving nature of large scale distributed computing, no longer dominated by pure computations but also by the need to handle and analyze large volumes of data. These data can be large databases of results, data streamed from a running application or another scientific instrument (collider for instance). These new workloads call for specific resource allocation strategies.

Data movements and storage are expected to be a major energy and performance bottleneck on next generation platforms. Storage architectures are also evolving, the standard centralized parallel file system being complemented with local persistent storage (Burst Buffers, NVRAM). Thus, one data producer can stage data on some nodes' local storage, requiring to schedule close by the associated analytics tasks to limit data movements. This kind of configuration, often referred as *in-situ analytics*, is expected to become common as it enables to switch from the traditional I/O intensive workflow (batch-processing followed by *post mortem* analysis and visualization) to a more storage conscious approach where data are processed as closely as possible to where and when they are produced (in-situ processing is addressed in details in section 4.3). By reducing data movements and scheduling the extra processing on resources not fully exploited yet, in-situ processing is expected to have also a significant positive energetic impact. Analytics codes can be executed in the same nodes than the application, often on dedicated cores commonly called helper cores, or on dedicated nodes called stagging nodes. The results are either forwarded to the users for visualization or saved to disk through I/O nodes. In-situ analytics can also take benefit of node local disks or burst buffers to reduce data movements. Future job scheduling strategies should take into account in-situ processes in addition to the job allocation to optimize both energy consumption and execution time. On the one hand, this problem can be reduced to an allocation problem of extra asynchronous tasks to idle computing units. But on the other hand, embedding analytics in applications brings extra difficulties by making the application more heterogeneous and imposing more constraints (data affinity) on the required resources. Thus, the main point here is to develop efficient algorithms for dealing with heterogeneity without increasing the global computational cost.

4.1.4. Learning

Another important issue is to adapt the job management system to deal with the bad effects of uncertainties, which may be catastrophic in large scale heterogeneous HPC platforms (jobs delayed arbitrarly far or jobs killed). A natural question is then: *is it possible to have a good estimation of the job and platform parameters in order to be able to obtain a better scheduling ?* Many important parameters (like the number or type of required resources or the estimated running time of the jobs) are asked to the users when they submit their jobs. However, some of these values are not accurate and in many cases, they are not even provided by the end-users. In DataMove, we propose to study new methods for a better prediction of the characteristics of the jobs and their execution in order to improve the optimization process. In particular, the methods well-studied in the field of big data (in supervised Machine Learning, like classical regression methods, Support Vector

Methods, random forests, learning to rank techniques or deep learning) could and must be used to improve job scheduling in large scale HPC platforms. This topic received a great attention recently in the field of parallel and distributed processing. A preliminary study has been done recently by our team with the target of predicting the job running times (called wall times). We succeeded to improve significantly in average the reference EASY Back Filling algorithm by estimating the wall time of the jobs, however, this method leads to big delay for the stretch of few jobs. Even if we succeed in determining more precisely hidden parameters, like the wall time of the jobs, this is not enough to determine an optimized solution. The shift is not only to learn on dedicated parameters but also on the scheduling policy. The data collected from the accounting and profiling of jobs can be used to better understand the needs of the jobs and through learning to propose adaptations for future submissions. The goal is to propose extensions to further improve the job scheduling and improve the performance and energy efficiency of the application. For instance preference learning may enable to compute on-line new priorities to back-fill the ready jobs.

4.1.5. Multi-objective Optimization

Several optimization questions that arise in allocation and scheduling problems lead to the study of several objectives at the same time. The goal is then not a single optimal solution, but a more complicated mathematical object that captures the notion of trade-off. In broader terms, the goal of multi-objective optimization is not to externally arbitrate on disputes between entities with different goals, but rather to explore the possible solutions to highlight the whole range of interesting compromises. A classical tool for studying such multi-objective optimization problems is to use *Pareto curves*. However, the full description of the Pareto curve can be very hard because of both the number of solutions and the hardness of computing each point. Addressing this problem will opens new methodologies for the analysis of algorithms.

To further illustrate this point here are three possible case studies with emphasis on conflicting interests measured with different objectives. While these cases are good representatives of our HPC context, there are other pertinent trade-offs we may investigate depending on the technology evolution in the coming years. This enumeration is certainly not limitative.

Energy versus Performance. The classical scheduling algorithms designed for the purpose of performance can no longer be used because performance and energy are contradictory objectives to some extent. The scheduling problem with energy becomes a multi-objective problem in nature since the energy consumption should be considered as equally important as performance at exascale. A global constraint on energy could be a first idea for determining trade-offs but the knowledge of the Pareto set (or an approximation of it) is also very useful.

Administrators versus application developers. Both are naturally interested in different objectives: In current algorithms, the performance is mainly computed from the point of view of administrators, but the users should be in the loop since they can give useful information and help to the construction of better schedules. Hence, we face again a multi-objective problem where, as in the above case, the approximation of the Pareto set provides the trade-off between the administrator view and user demands. Moreover, the objectives are usually of the same nature. For example, *max stretch* and *average stretch* are two objectives based on the slowdown factor that can interest administrators and users, respectively. In this case the study of the norm of stretch can be also used to describe the trade-off (recall that the L_1-norm corresponds to the average objective while the L_∞-norm to the max objective). Ideally, we would like to design an algorithm that gives good approximate solutions at the same time for all norms. The L_2 or L_3-norm are useful since they describe the performance of the whole schedule from the administrator point of view as well as they provide a fairness indication to the users. The hard point here is to derive theoretical analysis for such complicated tools.

Resource Augmentation. The classical resource augmentation models, i.e. speed and machine augmentation, are not sufficient to get good results when the execution of jobs cannot be frequently interrupted. However, based on a resource augmentation model recently introduced, where the algorithm may reject a small number of jobs, some members of our team have given the first interesting results in the non-preemptive direction. In general, resource augmentation can explain the intuitive good behavior of some greedy algorithms while, more interestingly, it can give ideas for new algorithms. For example, in the rejection context we could dedicate a small number of nodes for the usually problematic rejected jobs. Some initial experiments show that this can lead to a schedule for the remaining jobs that is very close to the optimal one.

4.2. Empirical Studies of Large Scale Platforms

Experiments or realistic simulations are required to take into account the impact of allocations and assess the real behavior of scheduling algorithms. While theoretical models still have their interest to lay the groundwork for algorithmic designs, the models are necessarily reflecting a purified view of the reality. As transferring our algorithm in a more practical setting is an important part of our creed, we need to ensure that the theoretical results found using simplified models can really be transposed to real situations. On the way to exascale computing, large scale systems become harder to study, to develop or to calibrate because of the costs in both time and energy of such processes. It is often impossible to convince managers to use a production cluster for several hours simply to test modifications in the RJMS. Moreover, as the existing RJMS production systems need to be highly reliable, each evolution requires several real scale test iterations. The consequence is that scheduling algorithms used in production systems are mostly outdated and not customized correctly. To circumvent this pitfall, we need to develop tools and methodologies for alternative empirical studies, from analysis of workload traces, to job models, simulation and emulation with reproducibility concerns.

4.2.1. Workload Traces with Resource Consumption

Workload traces are the base element to capture the behavior of complete systems composed of submitted jobs, running applications, and operating tools. These traces must be obtained on production platforms to provide relevant and representative data. To get a better understanding of the use of such systems, we need to look at both, how the jobs interact with the job management system, and how they use the allocated resources. We propose a general workload trace format that adds jobs resource consumption to the commonly used SWF [0] workload trace format. This requires to instrument the platforms, in particular to trace resource consumptions like CPU, data movements at memory, network and I/O levels, with an acceptable performance impact. In a previous work we studied and proposed a dedicated job monitoring tool whose impact on the system has been measured as lightweight (0.35% speed-down) with a 1 minute sampling rate. Other tools also explore job monitoring, like TACC Stats. A unique feature from our tool is its ability to monitor distinctly jobs sharing common nodes.

Collected workload traces with jobs resource consumption will be publicly released and serve to provide data for works presented in Section 4.1. The trace analysis is expected to give valuable insights to define models encompassing complex behaviours like network topology sensitivity, network congestion and resource interferences.

We expect to join efforts with partners for collecting quality traces (ATOS/Bull, Ciment meso center, Joint Laboratory on Extreme Scale Computing) and will collaborate with the Inria team POLARIS for their analysis.

4.2.2. Simulation

Simulations of large scale systems are faster by multiple orders of magnitude than real experiments. Unfortunately, replacing experiments with simulations is not as easy as it may sound, as it brings a host of new problems to address in order to ensure that the simulations are closely approximating the execution of typical workloads on real production clusters. Most of these problems are actually not directly related to scheduling algorithms assessment, in the sense that the workload and platform models should be defined independently from the algorithm evaluations, in order to ensure a fair assessment of the algorithms' strengths and weaknesses. These research topics (namely platform modeling, job models and simulator calibration) are addressed in the other subsections.

[0] Standard Workload Format: http://www.cs.huji.ac.il/labs/parallel/workload/swf.html

We developed an open source platform simulator within DataMove (in conjunction with the OAR development team) to provide a widely distributable test bed for reproducible scheduling algorithm evaluation. Our simulator, named Batsim, allows to simulate the behavior of a computational platform executing a workload scheduled by any given scheduling algorithm. To obtain sound simulation results and to broaden the scope of the experiments that can be done thanks to Batsim, we did not chose to create a (necessarily limited) simulator from scratch, but instead to build on top of the SimGrid simulation framework.

To be open to as many batch schedulers as possible, Batsim decouples the platform simulation and the scheduling decisions in two clearly-separated software components communicating through a complete and documented protocol. The Batsim component is in charge of simulating the computational resources behaviour whereas the scheduler component is in charge of taking scheduling decisions. The scheduler component may be both a resource and a job management system. For jobs, scheduling decisions can be to execute a job, to delay its execution or simply to reject it. For resources, other decisions can be taken, for example to change the power state of a machine i.e. to change its speed (in order to lower its energy consumption) or to switch it on or off. This separation of concerns also enables interfacing with potentially any commercial RJMS, as long as the communication protocol with Batsim is implemented. A proof of concept is already available with the OAR RJMS.

Using this test bed opens new research perspectives. It allows to test a large range of platforms and workloads to better understand the real behavior of our algorithms in a production setting. In turn, this opens the possibility to tailor algorithms for a particular platform or application, and to precisely identify the possible shortcomings of the theoretical models used.

4.2.3. Job and Platform Models

The central purpose of the Batsim simulator is to simulate job behaviors on a given target platform under a given resource allocation policy. Depending on the workload, a significant number of jobs are parallel applications with communications and file system accesses. It is not conceivable to simulate individually all these operations for each job on large plaforms with their associated workload due to implied simulation complexity. The challenge is to define a coarse grain job model accurate enough to reproduce parallel application behavior according to the target platform characteristics. We will explore models similar to the BSP (Bulk Synchronous Program) approach that decomposes an application in local computation supersteps ended by global communications and a global synchronization. The model parameters will be established by means of trace analysis as discussed previously, but also by instrumenting some parallel applications to capture communication patterns. This instrumentation will have a significant impact on the concerned application performance, restricting its use to a few applications only. There are a lot of recurrent applications executed on HPC platform, this fact will help to reduce the required number of instrumentations and captures. To assign each job a model, we are considering to adapt the concept of application signatures as proposed in. Platform models and their calibration are also required. Large parts of these models, like those related to network, are provided by Simgrid. Other parts as the filesystem and energy models are comparatively recent and will need to be enhanced or reworked to reflect the HPC platform evolutions. These models are then generally calibrated by running suitable benchmarks.

4.2.4. Emulation and Reproducibility

The use of coarse models in simulation implies to set aside some details. This simplification may hide system behaviors that could impact significantly and negatively the metrics we try to enhance. This issue is particularly relevant when large scale platforms are considered due to the impossibility to run tests at nominal scale on these real platforms. A common approach to circumvent this issue is the use of emulation techniques to reproduce, under certain conditions, the behavior of large platforms on smaller ones. Emulation represents a natural complement to simulation by allowing to execute directly large parts of the actual evaluated software and system, but at the price of larger compute times and a need for more resources. The emulation approach was chosen in to compare two job management systems from workload traces of the CURIE supercomputer (80000 cores). The challenge is to design methods and tools to emulate with sufficient accuracy the platform and the workload (data movement, I/O transfers, communication, applications interference). We will also intend to leverage emulation tools like Distem from the MADYNES team. It is also important to note that the Batsim simulator also uses emulation techniques to support the core scheduling module from actual RJMS. But the integration level is not the same when considering emulation for larger parts of the system (RJMS, compute node, network and filesystem).

Replaying traces implies to prepare and manage complex software stacks including the OS, the resource management system, the distributed filesystem and the applications as well as the tools required to conduct experiments. Preparing these stacks generate specific issues, one of the major one being the support for reproducibility. We propose to further develop the concept of reconstructability to improve experiment reproducibility by capturing the build process of the complete software stack. This approach ensures reproducibility over time better than other ways by keeping all data (original packages, build recipe and Kameleon engine) needed to build the software stack.

In this context, the Grid'5000 (see Sec. 5.4) experimentation infrastructure that gives users the control on the complete software stack is a crucial tool for our research goals. We will pursue our strong implication in this infrastructure.

4.3. Integration of High Performance Computing and Data Analytics

Data produced by large simulations are traditionally handled by an I/O layer that moves them from the compute cores to the file system. Analysis of these data are performed after reading them back from files, using some domain specific codes or some scientific visualisation libraries like VTK. But writing and then reading back these data generates a lot of data movements and puts under pressure the file system. To reduce these data movements, **the in situ analytics paradigm proposes to process the data as closely as possible to where and when the data are produced**. Some early solutions emerged either as extensions of visualisation tools or of I/O libraries like ADIOS. But significant progresses are still required to provide efficient and flexible high performance scientific data analysis tools. Integrating data analytics in the HPC context will have an impact on resource allocation strategies, analysis algorithms, data storage and access, as well as computer architectures and software infrastructures. But this paradigm shift imposed by the machine performance also sets the basis for a deep change on the way users work with numerical simulations. The traditional workflow needs to be reinvented to make HPC more user-centric, more interactive and turn HPC into a commodity tool for scientific discovery and engineering developments. In this context DataMove aims at investigating programming environments for in situ analytics with a specific focus on task scheduling in particular, to ensure an efficient sharing of resources with the simulation.

4.3.1. Programming Model and Software Architecture

In situ creates a tighter loop between the scientist and her/his simulation. As such, an in situ framework needs to be flexible to let the user define and deploy its own set of analysis. A manageable flexibility requires to favor simplicity and understandability, while still enabling an efficient use of parallel resources. Visualization libraries like VTK or Visit, as well as domain specific environments like VMD have initially been developed for traditional post-mortem data analysis. They have been extended to support in situ processing with some simple resource allocation strategies but the level of performance, flexibility and ease of use that is expected requires to rethink new environments. There is a need to develop a middleware and programming environment taking into account in its fundations this specific context of high performance scientific analytics.

Similar needs for new data processing architectures occurred for the emerging area of Big Data Analytics, mainly targeted to web data on cloud-based infrastructures. Google Map/Reduce and its successors like Spark or Stratosphere/Flink have been designed to match the specific context of efficient analytics for large volumes of data produced on the web, on social networks, or generated by business applications. These systems have mainly been developed for cloud infrastructures based on commodity architectures. They do not leverage the specifics of HPC infrastructures. Some preliminary adaptations have been proposed for handling scientific data in a HPC context. However, these approaches do not support in situ processing.

Following the initial development of FlowVR, our middleware for in situ processing, we will pursue our effort to develop a programming environment and software architecture for high performance scientific data analytics. Like FlowVR, the map/reduce tools, as well as the machine learning frameworks like TensorFlow, adopted a dataflow graph for expressing analytics pipe-lines. We are convinced that this dataflow approach is both easy to understand and yet expresses enough concurrency to enable efficient executions. The graph description can be compiled towards lower level representations, a mechanism that is intensively used by Stratosphere/Flink for instance. Existing in situ frameworks, including FlowVR, inherit from the HPC way of programming with a thiner software stack and a programming model close to the machine. Though this approach enables to program high performance applications, this is usually too low level to enable the scientist to write its analysis pipe-line in a short amount of time. The data model, i.e. the data semantics level accessible at the framework level for error check and optimizations, is also a fundamental aspect of such environments. The key/value store has been adopted by all map/reduce tools. Except in some situations, it cannot be adopted as such for scientific data. Results from numerical simulations are often more structured than web data, associated with acceleration data structures to be processed efficiently. We will investigate data models for scientific data building on existing approaches like Adios or DataSpaces.

4.3.2. Resource Sharing

To alleviate the I/O bottleneck, the in situ paradigm proposes to start processing data as soon as made available by the simulation, while still residing in the memory of the compute node. In situ processings include data compression, indexing, computation of various types of descriptors (1D, 2D, images, etc.). Per se, reducing data output to limit I/O related performance drops or keep the output data size manageable is not new. Scientists have relied on solutions as simple as decreasing the frequency of result savings. In situ processing proposes to move one step further, by providing a full fledged processing framework enabling scientists to more easily and thoroughly manage the available I/O budget.

The most direct way to perform in situ analytics is to inline computations directly in the simulation code. In this case, in situ processing is executed in sequence with the simulation that is suspended meanwhile. Though this approach is direct to implement and does not require complex framework environments, it does not enable to overlap analytics related computations and data movements with the simulation execution, preventing to efficiently use the available resources. Instead of relying on this simple time sharing approach, several works propose to rely on space sharing where one or several cores per node, called *helper cores*, are dedicated to analytics. The simulation responsibility is simply to handle a copy of the relevant data to the node-local in situ processes, both codes being executed concurrently. This approach often lead to significantly beter performance than in-simulation analytics.

For a better isolation of the simulation and in situ processes, one solution consists in offloading in situ tasks from the simulation nodes towards extra dedicated nodes, usually called *staging nodes*. These computations are said to be performed *in-transit*. But this approach may not always be beneficial compared to processing on simulation nodes due to the costs of moving the data from the simulation nodes to the staging nodes.

FlowVR enables to mix these different resources allocation strategies for the different stages of an analytics pile-line. Based on a component model, the scientist designs analytics workflows by first developing processing components that are next assembled in a dataflow graph through a Python script. At runtime the graph is instantiated according to the execution context, FlowVR taking care of deploying the application on the target architecture, and of coordinating the analytics workflows with the simulation execution.

But today the choice of the resource allocation strategy is mostly ad-hoc and defined by the programmer. We will investigate solutions that enable a cooperative use of the resource between the analytics and the simulation with minimal hints from the programmer. In situ processings inherit from the parallelization scale and data distribution adopted by the simulation, and must execute with minimal perturbations on the simulation execution (whose actual resource usage is difficult to know a priori). We need to develop adapted scheduling strategies that operate at compile and run time. Because analysis are often data intensive, such solutions must take into consideration data movements, a point that classical scheduling strategies designed first for compute intensive applications often overlook. We expect to develop new scheduling strategies relying on the methodologies developed in Sec. 4.1.5. Simulations as well as analysis are iterative processes exposing a strong spatial and temporal coherency that we can take benefit of to anticipate their behavior and then take more relevant resources allocation strategies, possibly based on advanced learning algorithms or as developed in Section 4.1.

In situ analytics represent a specific workload that needs to be scheduled very closely to the simulation, but not necessarily active during the full extent of the simulation execution and that may also require to access data from previous runs (stored in the file system or on specific burst-buffers). Several users may also need to run concurrent analytics pipe-lines on shared data. This departs significantly from the traditional batch scheduling model, motivating the need for a more elastic approach to resource provisioning. These issues will be conjointly addressed with research on batch scheduling policies (Sec. 4.1).

4.3.3. Co-Design with Data Scientists

Given the importance of users in this context, it is of primary importance that in situ tools be co-designed with advanced users, even if such multidisciplinary collaborations are challenging and require constant long term investments to learn and understand the specific practices and expectations of the other domain.

We will tightly collaborate with scientists of some application domains, like molecular dynamics or fluid simulation, to design, develop, deploy and assess in situ analytics scenarios, as already done with Marc Baaden, a computational biologist from LBT.

We recently extended our collaboration network. We started in 2015 a PhD co-advised with CEA DAM to investigate in situ analytics scenarios in the context of atomistic material simulations. CEA DAM is a French energy lab hosting one of the largest european supercomputer. They gather physicists, numerical scientists as well as high performance computer engineers, making it a very interesting partner for developing new scientific data analysis solutions. We also got a national grant (2015-2018) to compute in situ statistics for multi-parametric parallel studies with the research department of French power company EDF. In this context we collaborate with statisticians and fluid simulation experts to define in situ scenarios, revisit the statistic operators to be amenable to in situ processing, and define an adapted in situ framework.

5. New Software and Platforms

5.1. FlowVR

SCIENTIFIC DESCRIPTION: FlowVR adopts the "data-flow" paradigm, where your application is divided as a set of components exchanging messages (think of it as a directed graph). FlowVR enables to encapsulate existing codes in components, interconnect them through data channels, and deploy them on distributed computing resources. FlowVR takes care of all the heavy lifting such as application deployment and message exchange.

The base entity, called a module or component, is an autonomous process, potentially multi-threaded with tools like OpenMP, TBB, or deferring computations to a GPU or Xeon Phi. This module processes data coming from input ports and write data on output ports. A module has no global insight on where the data comes from or goes to. The programming interface is designed to limit code refactoring, easing turning an existing code into a FlowVR component. The three main functions are:

wait(): Blocking function call that waits for the availability of new messages on input ports. get(): Retrieve a handle to access the message received at the previous wait() call on a given input port. put(): Notify FlowVR that a new message on a given output port is ready for dispatch. FlowVR manages data transfers. Intra-node communications between two components take place through a shared memory segment, avoiding copies. Once the sender has prepared the data in a shared memory segment, it simply handles a pointer to the destination that can directly access them. Inter-node communications extend this mechanism, FlowVR taking care of packing and transferring the data from the source shared memory segment to the destination shared memory segment.

Assembling components to build an application consists in writing a Python script, instanciate it according to the target machine. FlowVR will process it and prepare everything so that in one command line you can deploy and start your application.
FUNCTIONAL DESCRIPTION: FlowVR adopts the "data-flow" paradigm, where your application is divided as a set of components exchanging messages (think of it as a directed graph). FlowVR enables to encapsulate existing codes in components, interconnect them through data channels, and deploy them on distributed computing resources. FlowVR takes care of all the heavy lifting such as application deployment and message exchange.

- Participants: Bruno Raffin, Clément Ménier, Emmanuel Melin, Jean Denis Lesage, Jérémie Allard, Jérémy Jaussaud, Matthieu Dreher, Sébastien Limet, Sophie Robert and Valérie Gourantou
- Contact: Bruno Raffin
- URL: http://flowvr.sf.net

5.2. OAR

KEYWORDS: HPC - Cloud - Clusters - Resource manager - Light grid
SCIENTIFIC DESCRIPTION: This batch system is based on a database (PostgreSQL (preferred) or MySQL), a script language (Perl) and an optional scalable administrative tool (e.g. Taktuk). It is composed of modules which interact mainly via the database and are executed as independent programs. Therefore, formally, there is no API, the system interaction is completely defined by the database schema. This approach eases the development of specific modules. Indeed, each module (such as schedulers) may be developed in any language having a database access library.
FUNCTIONAL DESCRIPTION: OAR is a versatile resource and task manager (also called a batch scheduler) for HPC clusters, and other computing infrastructures (like distributed computing experimental testbeds where versatility is a key).

- Participants: Bruno Bzeznik, Olivier Richard and Pierre Neyron
- Partners: LIG - CNRS - Grid'5000 - CIMENT
- Contact: Olivier Richard
- URL: http://oar.imag.fr

5.3. MELISSA

Modular External Library for In Situ Statistical Analysis
KEYWORD: Sensitivity Analysis
FUNCTIONAL DESCRIPTION: Melissa is an in situ solution for sensitivity analysis. It implements iterative algorithms to compute spatio-temporal statistic fields over results of large scale sensitivity studies. Melissa relies on a client/server architecture, composed of three main modules:

Melissa Server: an independent parallel executable. It receives data from the simulations, updates iterative statistics as soon as possible, then trow data away. Melissa API: a shared library to be linked within the simulation code. It mainly transmit simulation data to Melissa Server at each timestep. The simulations of the sensitivity analysis become the clients of Melissa Server. Melissa Launcher: A Python script in charge of generating and managing the whole global sensitivity analysis.

- Authors: Théophile Terraz, Bruno Raffin, Alejandro Ribes and Bertrand Iooss
- Partner: Edf
- Contact: Bruno Raffin
- Publications: In Situ Statistical Analysis for Parametric Studies - Melissa: Large Scale In Transit Sensitivity Analysis Avoiding Intermediate Files
- URL: https://melissa-sa.github.io

5.4. Platforms

5.4.1. Grid'5000 (https://www.grid5000.fr/) and Meso Center Ciment (https://ciment.ujf-grenoble.fr)

We have been very active in promoting the factorization of compute resources at a regional and national level. We have a three level implication, locally to maintain a pool of very flexible experimental machines (hundreds of cores), regionally through the CIMENT meso center (Equipex Grant), and nationally by contributing to the Grid'5000 platform, our local resources being included in this platform. Olivier Richard is member of Grid'5000 scientific committee and Pierre Neyron is member of the technical committee. The OAR scheduler in particular is deployed on both infrastructures. We are currently preparing proposals for the next generation machines within the context of the new university association (Univ. Grenoble-Alpes).

6. New Results

6.1. Integration of High Performance Computing and Data Analytics

6.1.1. I/O Survey

First contribution is a comprehensive survey on parallel I/O in the HPC context [14]. As the available processing power and amount of data increase, I/O remains a central issue for the scientific community. This survey focuses on a traditional I/O stack, with a POSIX parallel file system. Through the comprehensive study of publications from the most important conferences and journals in a five-year time window, we discuss the state of the art of I/O optimization approaches, access pattern extraction techniques, and performance modeling, in addition to general aspects of parallel I/O research. This survey enables us to identify the general characteristics of the field and the main current and future research topics.

6.1.2. Task Based In Situ Processing

One approach to bypass the I/O bottleneck is *in situ* processing, an important research topic at DataMove. The *in situ* paradigm proposes to reduce data movement and to analyze data while still resident in the memory of the compute node by co-locating simulation and analytics on the same compute node. The simplest approach consists in modifying the simulation timeloop to directly call analytics routines. However, several works have shown that an *asynchronous* approach where analytics and simulation run concurrently can lead to a significantly better performance. Today, the most efficient approach consists in running the analytics processes on a set of dedicated cores, called helper cores, to isolate them from the simulation processes. Simulation and analytics thus run concurrently on different cores but this static isolation can lead to underused resources if the simulation or the analytics do not fully use all the assigned cores.

In this work performed in collaboration with CEA, we developed TINS, a task-based in situ framework that implements a novel *dynamic helper core* strategy. TINS relies on a work stealing scheduler and on task-based programming. Simulation and analytics tasks are created concurrently and scheduled on a set of worker threads created by a single instance of the work stealing scheduler. Helper cores are assigned dynamically: some worker threads are dedicated to analytics when analytics tasks are available while they join the other threads for processing simulation tasks otherwise, leading to a better resource usage. We leverage the good compositionality properties of task-based programming to seamlessly keep the analytics and simulation codes well separated and a plugin system enables to develop parallel analytics codes outside of the simulation code.

TINS is implemented with the Intel Threading Building Blocks (TBB) library that provides a task-based programming model and a work stealing scheduler. The experiments are conducted with the hybrid MPI+TBB ExaStamp molecular dynamics code that we associate with a set of analytics representative of computational physics algorithms. We show up to 40% performance improvement over various other approaches, including the standard helper core, on experiments on up to 14,336 Broadwell cores.

6.1.3. Stream Processing

Stream processing is the Big Data equivalent of in situ processing. It consists in analyzing on-line incoming streams of data, often produced from sensors or social networks like Twitter. We investigated the convergence between both paradigms through different directions: how the programming environment developed specifically for stream processing can applied to the data produced by large parallel simulations [18]; Proposing a dynamics data structure to keep sorted data streams [12]; Evaluating the performance of the FlameMR framework on data produced from a parallel simulation[13]. We summarize here the 2 first contributions.

6.1.3.1. Packed Memory QuadTree.

Over the past years, several in-memory big-data management systems have appeared in academia and industry. In-memory databases systems avoid the overheads related to traditional I/O disk-based systems and have made possible to perform interactive data-analysis over large amounts of data. A vast literature of systems and research strategies deals with different aspects, such as the limited storage size and a multi-level memory-hierarchy of caches. Maintaining the right data layout that favors locality of accesses is a determinant factor for the performance of in-memory processing systems. Stream processing engines like Spark or Flink support the concept of *window*, which collects the latest events without a specific data organization. It is possible to trigger the analysis upon the occurrence of a given criterion (time, volume, specific event occurrence). After a window is updated, the system shifts the processing to the next batch of events. There is a need to go one step further to keep a live window continuously updated while having a fine grain data replacement policy to control the memory footprint. The challenge is the design of dynamic data structures to absorb high rate data streams, stash away the oldest data to stay in the allowed memory budget while enabling fast queries executions to update visual representations. A possible solution is the extension of database structures like R-trees used in SpatiaLite or PostGis, or to develop dedicated frameworks like Kit based on a pyramid structure.

We developed a novel self-organized cache-oblivious data structure, called PMQ, for in-memory storage and indexing of fixed length records tagged with a spatiotemporal index. We store the data in an array with a controlled density of gaps (*i.e.*, empty slots) that benefits from the properties of the *Packed Memory Arrays*. The empty slots guarantee that insertions can be performed with a low amortized number of data movements ($O(\log^2(N))$) while enabling efficient spatiotemporal queries. During insertions, we rebalance parts of the array when required to respect density constraints, and the oldest data is stashed away when reaching the memory budget. To spatially subdivide the data, we sort the records according to their Morton index, thus ensuring spatial locality in the array while defining an implicit, recursive quadtree, which leads to efficient spatiotemporal queries. We validate PMQ for consuming a stream of tweets to answer visual and range queries. PMQ significantly outperforms the widely adopted spatial indexing data structure R-tree, typically used by relational databases, as well as the conjunction of Geohash and B$^+$-tree, typically used by NoSQL databases.

6.1.3.2. Flink based in situ Processing.

We proposed to leverage Apache Flink, a scalable stream processing engine from the Big Data domain, in this HPC context. Flink enables to program analyses within a simple window based map/reduce model, while the runtime takes care of the deployment, load balancing and fault tolerance. We build a complete in transit analytics workflow, connecting an MD simulation to Apache Flink and to a distributed database, Apache HBase, to persist all the desired data. To demonstrate the expressivity of this programming model and its suitability for HPC scientific environments, two common analytics in the Molecular Dynamics field have been implemented. We assessed the performance of this framework, concluding that it can handle simulations of sizes used in the literature while providing an effective and versatile tool for scientists to easily incorporate on-line parallel analytics in their current workflows.

6.2. Data Aware Batch Scheduling

6.2.1. Batch Scheduling for Energy

The project COSMIC [24], [22], [16], [17], in collaboration with Myriads team in Inria Rennes-Atlantique, targets the optimization of green energy usage in Clouds. The project considers a geographically distributed cloud, with each data center associated with a local photovoltaic (PV) farm. The objective is to maximize the photovoltaic energy by allocation the computing workload to the data centers according to its energy production. The production forecasting is modeled with a truncated normal law, permitting to consider the uncertainty of the forecast.

Chapter [24] considers a simple model with homogeneous Virtual Machines submitted at unpredictable rate. This study has resulted in a scheduling algorithm for task allocation. The chapter demonstrates the optimality of this algorithm at current time slot according to production forecast parameters.

Paper [22] extends these results to heterogeneous VM. Each VM is defined by its arrival date, its execution time, its memory requirement and its CPU usage. In this model, due to execution time durations, the possibility to migrate running VM was considered. An algorithm is detailed in the paper that is compared to standard algorithm through simulations.

A third study [16], [17] has carefully modeled the interactions between the Cloud and the energy supplier. Due to variability of PV production and workload submission, each data center will alternatively inject energy into the electricity grid or purchase energy. The energy model considers a virtual energy pool mitigating the surplus and deficit of the different data center, with reduced costs regarding the difference between electricity cost and electricity injection tariff. The algorithm detailed in this paper outperforms well-known round-robin approaches, as shown by simulations.

6.2.2. Learning Methods for Batch Scheduling

Most of Job Scheduling algorithms apply greedy tasks ordering, as First Come First Served (FCFS) or Shortest Processing time First (SPF). They give simple methods, highly practical with certain guarantees. They are however far from optimal. Mixed methods, combining many of this basic methods permit to improve their performance. DataMove has developed [27] a learning method permitting to adapt the Mixed method to benchmarks. An extensive experimental campaign has permitted to determine the possibilities of basic and mixed methods according to the benchmarks characteristics, enhancing the efficiency of mixed methods.

6.2.3. Reproducibility

Related to batch scheduling experimentation, DataMove has led investigations on reproducibility [23]. Existing approaches focus on repeatability, but this is only the first step to reproducibility: Continuing a scientific work from a previous experiment requires to be able to modify it. This ability is called reproducibility with Variation. We show that capturing the environment of execution is necessary but not sufficient ; we also need the environment of development. The variation also implies that those environments are subject to evolution, so the whole software development lifecycle needs to be considered. To take into account these evolutions, software environments need to be clearly defined, reconstructible with variation, and easy to share. In this context, we propose new way of seeing reproducibility through the scientific software development lifecycle. Each step in this lifecycle requires a software environment. We define a software environment by a set of applications and libraries, with all their dependencies, and their configurations, required to achieve a step in a scientific workflow.

6.2.4. Online Algorithms

Rob van Stee wrote a review of 2018 online algorithms including our recent contributions on resource augmentation [0] We quote him here:

Progress was also made on scheduling to minimize weighted flow time on unrelated machines. In ESA 2016, Giorgio Lucarelli et al. [1] had considered a version where the online algorithm can reject some $\varepsilon_r > 0$ fraction (by weight) of the jobs and have machines that are $1 + \varepsilon_s$ as fast as the offline machines, for some $\varepsilon_s > 0$. They showed that this is already enough to achieve a competitive ratio of $O(1/(\varepsilon_s\varepsilon_r))$.

[0] Rob van Stee. 2018. SIGACT News Online Algorithms Column 34: 2018 in review. SIGACT News 49, 4 (December 2018), 36-45.

In SPAA 2018, Giorgio Lucarelli et al.[20] (a superset of the previous authors) showed that it is in fact sufficient to reject a 2ε fraction of the total number of jobs to achieve a competitive ratio of $2(\frac{1+\varepsilon}{\varepsilon})$ for minimizing the total flow time. This algorithm sometimes rejects a job other than the one that has just arrived. The authors show that this is necessary, as otherwise there is a lower bound of $\Omega(\Delta)$ even on a single machine. Here Δ is the size ratio (the ratio of largest to smallest job size). (Obviously this lower bound also holds if you cannot reject jobs at all.)

They also consider the speed scaling model, in which machines can be sped up if additional energy is invested, and the goal is to minimize the total weighted flow time plus energy usage. If the power function of machine i is given by $P(s_i(t)) = s_i(t)^\alpha$, where $s_i(t)$ is the current speed of machine i, there is an algorithm which is $O((1+1/\varepsilon)^{\alpha/(\alpha-1)})$-competitive that rejects jobs of total weight at most a fraction ε of the total weight of all the jobs. They also give a positive result for jobs with hard deadlines, where the goal is to minimize the total energy usage and no job may be rejected.

In ESA 2018, the same set of authors [11] improved/generalized these results by showing that rejection alone is sufficient for an algorithm to be competitive even for weighted flow time. They presented an $O(1/\varepsilon^3)$-competitive algorithm that rejects at most $O(\varepsilon)$ of the total weight of the jobs. In this algorithm, jobs are assigned (approximately) greedily to machines, and each machine runs the jobs assigned to it using Highest Density First. A job may be rejected if it is running while much heavier jobs arrive or if it is in the queue while very many jobs arrive. The second rule simulates the resource augmentation on the speed.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- **BULL-ATOS SE (2016-2019)**. Two PhD grants (Michael Mercier and Adrien Faure). Job and resource management algorithms.
- **CEA DAM (2016-2018)**. PhD grant support contract (PhD of Estelle Dirand, funded by CEA). In situ analysis for Molecular Simulations.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

- **ANR grant GRECO (2017-2020).** Resource manager for cloud of things. Coordinator: Quarnot Computing. Partners: Quarnot Computing, Grenoble-INP, Inria.
- **ANR grant Energumen (2018-2022).** Resource management: malleable jobs for a better use of the resources along with energy optimization. Coordinator: Denis Trystram. Partners: Grenoble-INP, IRIT, Sorbonne Université.

8.1.2. Competitivity Clusters

- **PIA Avido (2015-2018)**. In situ analysis and visualization for large scale numerical simulation. Coordinator: EDF SA. Partners: EDF R&D, Total SA, Kitware SAS , Université Pierre et Marie CURIE, Inria (DataMove).
- **FUI OverMind (2015-2018)**. Task planification and asset management for the cartoon productions. Coordinator: Teamto Studio. Partners: Teamto Studio, Folimage Studio, Ecole de Gobelins, Inria (DataMove).

- **FUI IDIOM (2018-2020)**. Monitoring and optimization of I/Os. Coordinator DDN Storage. Partners: DDN Storage, Criteo, Quarnot, QuasarDB, CEA, Université de Bretagne Occidentale, Telecom SudParis, Inria (DataMove).

8.1.3. Inria

- Inria PRE COSMIC (exploratory research project), 2017-2019. Photovoltaic Energy Management for Distributed Cloud Platforms. Myriads, DataMove.
- Inria IPL HPC-BigData, 2018-2021). Convergence between HPC, Big Data and AI. Coordinator: Bruno Raffin. Partners: the Inria teams Zenith, Kerdata, Datamove, Tadaam, SequeL, Parietal, Tau, and the external partners ATOS, ANL, IBPC, ESI-Group. See https://project.inria.fr/hpcbigdata/

8.2. European Initiatives

8.2.1. Collaborations in European Programs, Except FP7 & H2020

Program: SKŁODOWSKA-CURIE ACTIONS - Individual Fellowship

Project acronym: DAMA

Project title: Extreme-Scale Data Management

Duration: November 2018 - October 2020

Coordinator: Bruno Raffin

Followship Recipient: Francieli Zanon Boito.

Abstract: This project is concerned with the I/O challenges that arise from the convergence between these two different paradigms. It is clear data analytics tools cannot simply replace their typical storage solutions for the HPC I/O stack, centered on the abstraction of files and powered by a parallel file system, because their workload is not well suited for that and would observe poor performance. Moreover, the separated storage infrastructure breaks the data affinity idea in which they are built upon. Finally, even among traditional HPC applications there is a need to minimize data movement, as it imposes high latency and increases energy consumption.

8.3. International Initiatives

8.3.1. Inria International Labs

8.3.1.1. JLESC

Title: Joint Laboratory for Extreme-Scale-Computing.

International Partners:

University of Illinois at Urbana Champaign (USA)

Argonne National Laboratory (USA),

Barcelona Supercomputing Center (Spain),

Jülich Supercomputing Centre (Germany)

Riken Advanced Institute for Computational Science (Japan)

Start year: 2009

See also: https://jlesc.github.io/

The purpose of the Joint Laboratory for Extreme Scale Computing is to be an international, virtual organization whose goal is to enhance the ability of member organizations and investigators to make the bridge between Petascale and Extreme computing. The JLESC organizes a workshop every 6 months DataMove participates to. DataMove developed several collaborations related to in situ processing with Tom Peterka group (ANL) , the Argo exascale operating system with Swann Perarnau (ANL).

8.3.2. Participation in Other International Programs

8.3.2.1. LICIA

Title: International Laboratory in High Performance and Ubiquitous Computing

International Partner (Institution - Laboratory - Researcher):

UFRGS (Brazil)

Duration: 2011 - 2018

See also: http://licia-lab.org/

The LICIA is an Internacional Laboratory and High Performance and Ubiquitous Computing born in 2011 from the common desire of members of Informatics Institute of the Federal University of Rio Grande do Sul and of Laboratoire d'Informatique de Grenoble to enhance and develop their scientific parternship that started by the end of the 1970. LICIA is an Internacional Associated Lab of the CNRS, a public french research institution. It has support from several brazilian and french research funding agencies, such as CNRS, Inria, ANR, European Union (from the french side) and CAPES, CNPq, FAPERGS (from the Brazilian side). DataMove is deeply involved in the animation of LICIA. Bruno Raffin is LICIA associate director.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

Professor visit: Alba Cristina Magalhaes Alves De Melo, Professor at University of Brasilia, visited the Datamove for one month in 2018.

PhD in progress: Danilo Carastan Dos Santos, Dynamic Scheduling of Tasks in High Performance Platforms with Machine Learning (Sao Paulo, Brasil). 1 year "sandwich" visit. Local adviser: Denis Trystram

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

President of the steering committee of Edu-Europar.

President of the steering committee of EGPGV (Eurographics Symposium on Parallel Graphics and Visualization).

Member of the steering committee of Europar.

Member of the steering committee of *Journée de visualisation scientifique*.

Member of the steering committee of HeteroPar.

Co-chair of the First HPML Workshop collocated with SBAC-PAD Lyon , October 2018.

9.1.1.2. Member of the Organizing Committees

Euro-Par Advisory Board Member

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

ISAV 2018 (Workshop on In Situ Infrastructures for Enabling Extreme-scale Analysis and Visualization) , November, Dallas, USA

EGPGV 2018 (Eurographics Symposium on Parallel Graphics and Visualization), June, Brno, Czech Republic.

LADV 2018 (IEEE Symposium on Large Data Analysis and Visualization), October, Berlin, Germany.

IPDPS 2018 (nternational Parallel and Distributed Processing Symposium), May, Vancouver, Canada.

LATIN, April 2018, Buenos Aires, Argentina.

CCgrid 2018 (18th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing), May, Washington DC.

SPAA (30th ACM Symposium on Parallelism in Algorithms and Architecture), July 2018, Vienna, Austria.

ISPDC, June 2018, Geneva, Switzerland.

CONPAS (Conference d'informatique en Parallelisme, Architecture et Systeme), July 2018, Toulouse, France.

HiPC 2018 (IEEE internat. Conf on High Performance Computing, Data and Analytics), December 2018, Bengaluru, India.

EURO 2018 (29th European Conference On Operational Research), July, Valencia.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Associate Editor of the Parallel Computing journal PARCO.

Member of the Editorial Board of Computational Methods in Science and Technology.

Member of the Editorial Board of ARIMA (revue africaine de recherche en informatique et maths appliquées).

9.1.4. Scientific Expertise

ANR project evaluation expert

9.1.5. Research Administration

Director of Pôle MSTIC of COMUE Univ. Grenoble-Alpes.

Head of the international Master program (MOSIG1) at Grenoble INP (ENSIMAG)

Steering committee of Grid'5000

Steering committee of GRICAD

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: Denis Trystram is responsible of the first year (M1) of the international Master of Science in Informatics at Grenoble (MOSIG-M1). 200 hours per year in average.

Master: Fanny Dufossé. 30 hours per year. Combinatorial scientific computing in Master at ENS Lyon and Algorithmic in Licence at Grenoble INP.

Master: Pierre-François Dutot. 226 hours per year. Licence (first and second year) at IUT2/UPMF (Institut Universitaire Technologique de Univ. Grenoble-Alpes) and 9 hours Master M2R-ISC Informatique-Systèmes-Communication at Univ. Grenoble-Alpes.

Master: Grégory Mounié. 242 hours per year. Master (M1/2nd year and M2/3rd year) at Engineering school ENSIMAG, Grenoble-INP.

Master: Bruno Raffin. 28 hours per year. Parallel System. International Master of Science in Informatics at Grenoble (MOSIG-M2).

Master: Olivier Richard is responsible of the third year of the computer science department of Grenoble INP. 222 hours per year. Master at Engineering school Polytech-Grenoble, Univ. Grenoble-Alpes.

Master: Frédéric Wagner. 220 hours per year. Engineering school ENSIMAG, Grenoble-INP (M1/2nd year and M2/3rd year).

Master: Yves Denneulin. 192 hours per year. Engineering school ENSIMAG, Grenoble-INP (M1/2nd year and M2/3rd year).

9.2.2. Supervision

PhD: Estelle Dirand, Integration of High-Performance Data Analytics and IOs for Molecular Dynamics on Exascale Computer, Univ. Grenoble-Alpes. November 2018. Bruno Raffin and Laurent Colombet (CEA).

PhD: Valentin Reis, Learning to Control Large-Scale Parallel Platforms, Univ. Grenoble-Alpes. October 2018. Advisers: Denis Trystram and Eric Gaussier

PhD in progress: Michael Mercier, Resource Management and Job Scheduling in HPC–Cloud environments towards the Big Data era, Univ. Grenoble Alpes. Started October 2016. Advisers: Olivier Richard and Bruno Raffin.

PhD in progress: Alessandro Kraemer, Scheduling in the Cloud, Univ Grenoble-Alpes and UFPR (co-tutelle). Started October 2014. Advisers: Olivier Richard and Denis Trystram.

PhD in progress: Mohammed Khatiri, Tasks scheduling on heterogeneous Multicore, Univ. Grenoble-Alpes and University Mohammed First (co-tutelle), Advisers: Denis Trystram, El Mostafa DAOUDI (University Mohammed First, Oujda, Morocco)

PhD in progress: Adrien Faure, Scheduling with Resource Augmentation, Advisers: Denis Trystram

PhD in progress: Clément Mommessin, Scheduling on heterogeneous platforms, Advisers: Denis Trystram

PhD in progress: Loris Felardos, Deep Learning for the Analytics of Molecular Systems, Advisers: Bruno Raffin, Guillaume Charpiat (Inria team Tau), Jérome Hénin (IBPC).

PhD in Progress: Salah Zrigui, Learning Scheduling Strategies, Advisers: Denis Trystram and Fanny Dufossé.

9.2.3. Juries

PhD Defense of Jonathan Sarton, Visualisations interactives haute-performance de donnes volumiques massives : une approche out-of-core multi-resolution basee GPUs, 28 Novembre 2018, Université de Reims Champagne-Ardenne. Reviewer

PhD Defense of Adrian Perez Dieguez, Parallel Prefix Operations on Heterogeneous Platforms, December 2018, Universidad da Coruna. Reviewer.

PhD Defense of Khalil Labidi, Parallelisation of hybrid metaheuristics for COP solving, 20 septembre 2018, Université de Tunis et Université Paris-Dauphine. President

PhD Defense of Stéphane Durand, Contrôle distribué et théorie des jeux : application aux systèmes auto-optimisants, 13 décembre 2018, Univ Grenoble Alpes. President

HDR Defense of Samuel Thibault, 13 décembre 2018, University Bordeaux I. Reviewer

9.3. Popularization

9.3.1. Interventions

- Talk *Des besoins en calcul de plus en plus performant* for the conference cycle "Accompagnement de l'informatique au lycée en 2018".

10. Bibliography

Major publications by the team in recent years

[1] G. LUCARELLI, N. KIM THANG, A. SRIVASTAV, D. TRYSTRAM.*Online Non-preemptive Scheduling in a Resource Augmentation Model based on Duality*, in "European Symposium on Algorithms (ESA 2016)", Aarhus, Denmark, August 2016, vol. 57, n^{o} 63, p. 1-17 [*DOI :* 10.4230/LIPICS.ESA.2016.63], http://hal.univ-grenoble-alpes.fr/hal-01334219

Publications of the year

Doctoral Dissertations and Habilitation Theses

[2] E. DIRAND.*Integration of High-Performance Task-Based In Situ for Molecular Dynamics on Exascale Computers*, Université Grenoble - Alpes, November 2018, https://hal.archives-ouvertes.fr/tel-01949170

[3] V. REIS.*Learning to control large-scale parallel platforms*, Université Grenoble Alpes (France), September 2018, https://hal.inria.fr/tel-01965150

Articles in International Peer-Reviewed Journal

[4] M. AMARIS, G. LUCARELLI, C. MOMMESSIN, D. TRYSTRAM.*Generic Algorithms for Scheduling Applications on Heterogeneous Platforms*, in "Concurrency and Computation: Practice and Experience", July 2018, p. 1-29 [*DOI :* 10.1002/CPE.4647], https://hal.inria.fr/hal-01896868

[5] E. BAMPIS, A. KONONOV, D. LETSIOS, G. LUCARELLI, M. SVIRIDENKO.*Energy Efficient Scheduling and Routing via Randomized Rounding*, in "Journal of Scheduling", February 2018, vol. 21, n^{o} 1, p. 35-51 [*DOI :* 10.1007/S10951-016-0500-2], https://hal.inria.fr/hal-01725140

[6] J. BLAZEWICZ, B. MOSELEY, E. PESCH, D. TRYSTRAM, G. ZHANG.*New challenges in scheduling theory*, in "Journal of Scheduling", December 2018, vol. 21, n^{o} 6, p. 581-582 [*DOI :* 10.1007/S10951-018-0571-3], https://hal.archives-ouvertes.fr/hal-01963221

[7] F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Further notes on Birkhoff-von Neumann decomposition of doubly stochastic matrices*, in "Linear Algebra and Applications", 2018, vol. 554, p. 68–78 [*DOI :* 10.1016/J.LAA.2018.05.017], https://hal.inria.fr/hal-01586245

[8] É. GAUSSIER, J. LELONG, V. REIS, D. TRYSTRAM.*Online Tuning of EASY-Backfilling using Queue Reordering Policies*, in "IEEE Transactions on Parallel and Distributed Systems", October 2018, vol. 29, n^{o} 10, p. 2304-2316 [*DOI :* 10.1109/TPDS.2018.2820699], https://hal.archives-ouvertes.fr/hal-01963216

[9] L. JOURDAN, S. JACQUIN, F. DUFOSSÉ.*An exact algorithm for the bi-objective timing problem*, in "Optimization Letters", 2018, vol. 12, n^{o} 4, p. 903–914 [*DOI :* 10.1007/S11590-018-1237-Y], https://hal.archives-ouvertes.fr/hal-01716581

[10] S. KEDAD-SIDHOUM, F. MONNA, G. MOUNIÉ, D. TRYSTRAM.*A Family of Scheduling Algorithms for Hybrid Parallel Platforms*, in "International Journal of Foundations of Computer Science", 2018, vol. 29, n^{o} 1, p. 63-90 [*DOI :* 10.1142/S012905411850003X], https://hal.sorbonne-universite.fr/hal-01516700

[11] G. LUCARELLI, B. MOSELEY, N. K. THANG, A. SRIVASTAV, D. TRYSTRAM.*Online Non-Preemptive Scheduling to Minimize Weighted Flow-time on Unrelated Machines*, in "Leibniz International Proceedings in Informatics ", 2018, vol. 112 [*DOI :* 10.4230/LIPICS.ESA.2018.59], https://hal.archives-ouvertes.fr/hal-01986352

[12] J. TOSS, C. A. L. PAHINS, B. RAFFIN, J. L. D. COMBA.*Packed-Memory Quadtree: a cache-oblivious data structure for visual exploration of streaming spatiotemporal big data*, in "Computers and Graphics", 2018, vol. 76, p. 117-128 [*DOI :* 10.1016/J.CAG.2018.09.005], https://hal.archives-ouvertes.fr/hal-01876579

[13] J. VEIGA, R. R. EXPÓSITO, B. RAFFIN, J. TOURINO.*Optimization of Real-World MapReduce Applications With Flame-MR: Practical Use Cases*, in "IEEE Access", 2018, vol. 6, p. 69750-69762 [*DOI :* 10.1109/ACCESS.2018.2880842], https://hal.inria.fr/hal-01955503

[14] F. ZANON BOITO, E. CAMILO INACIO, J. LUCA BEZ, P. O. A. NAVAUX, M. A. R. DANTAS, Y. DENNEULIN.*A Checkpoint of Research on Parallel I/O for High Performance Computing*, in "ACM Computing Surveys", March 2018, vol. 51, n^{o} 2, p. 23:1-23:35 [*DOI :* 10.1145/3152891], http://hal.univ-grenoble-alpes.fr/hal-01591755

International Conferences with Proceedings

[15] R. BLEUSE, K. DOGEAS, G. LUCARELLI, G. MOUNIÉ, D. TRYSTRAM.*Interference-Aware Scheduling using Geometric Constraints*, in "Euro-Par 2018 - European Conference on Parallel Processing", Torino, Italy, Euro-Par 2018: Parallel Processing, Springer, August 2018, vol. LNCS 11014, p. 205-217 [*DOI :* 10.1007/978-3-319-96983-1_15], http://hal.univ-grenoble-alpes.fr/hal-01884542

[16] B. CAMUS, A. BLAVETTE, F. DUFOSSÉ, A.-C. ORGERIE.*Self-Consumption Optimization of Renewable Energy Production in Distributed Clouds*, in "Cluster 2018 - IEEE International Conference on Cluster Computing", Belfast, United Kingdom, IEEE, September 2018, p. 1-11, https://hal.archives-ouvertes.fr/hal-01856660

[17] B. CAMUS, F. DUFOSSÉ, A. BLAVETTE, M. QUINSON, A.-C. ORGERIE.*Network-aware energy-efficient virtual machine management in distributed Cloud infrastructures with on-site photovoltaic production*, in "SBAC-PAD 2018 - International Symposium on Computer Architecture and High Performance Computing", Lyon, France, September 2018, p. 1-8, https://hal.archives-ouvertes.fr/hal-01856657

[18] E. DIRAND, L. COLOMBET, B. RAFFIN.*TINS: A Task-Based Dynamic Helper Core Strategy for In Situ Analytics*, in "SCA18 - Supercomputing Frontiers Asia 2018", Singapore, Singapore, March 2018, p. 159-178 [*DOI :* 10.1007/978-3-319-69953-0_10], https://hal.inria.fr/hal-01730910

[19] F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Approximation algorithms for maximum matchings in undirected graphs*, in "CSC 2018 - SIAM Workshop on Combinatorial Scientific Computing", Bergen, Norway, Proceedings of the Seventh SIAM Workshop on Combinatorial Scientific Computing, SIAM, June 2018, p. 56-65 [*DOI :* 10.1137/1.9781611975215.6], https://hal.archives-ouvertes.fr/hal-01740403

[20] G. LUCARELLI, B. MOSELEY, N. K. THANG, A. SRIVASTAV, D. TRYSTRAM.*Online Non-preemptive Scheduling on Unrelated Machines with Rejections*, in "Proceedings of the 30th on Symposium on Parallelism in Algorithms and Architectures", Vienna, Austria, ACM Press, July 2018, p. 291-300 [*DOI :* 10.1145/3210377.3210402], https://hal.archives-ouvertes.fr/hal-01986312

[21] H. C. ZANÚZ, B. RAFFIN, O. A. MURES, E. J. PADRÓN.*In-Transit Molecular Dynamics Analysis with Apache Flink*, in "ISAV 2018 - In Situ Infrastructures for Enabling Extreme-scale Analysis and Visualization", DALLAS, United States, November 2018, p. 1-8, https://hal.inria.fr/hal-01889939

Conferences without Proceedings

[22] B. CAMUS, A. BLAVETTE, F. DUFOSSÉ, A.-C. ORGERIE.*Harnessing the geographical flexibility of distributed computing clouds for cooperative self-consumption*, in "ISGT-Europe 2018 - 8th IEEE PES Innovative Smart Grid Technologies Conference Europe", Sarajevo, Bosnia and Herzegovina, October 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01856665

[23] M. MERCIER, A. FAURE, O. RICHARD.*Considering the Development Workflow to Achieve Reproducibility with Variation*, in "SC 2018 - Workshop: ResCuE-HPC", Dallas, United States, November 2018, p. 1-5, https://hal.inria.fr/hal-01891084

Scientific Books (or Scientific Book chapters)

[24] B. CAMUS, F. DUFOSSÉ, A.-C. ORGERIE.*The SAGITTA approach for optimizing solar energy consumption in distributed clouds with stochastic modeling*, in "Smart Cities, Green Technologies, and Intelligent Transport Systems", December 2018, p. 1-25, https://hal.archives-ouvertes.fr/hal-01945821

Research Reports

[25] F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Effective heuristics for matchings in hypergraphs*, Inria Grenoble Rhône-Alpes, November 2018, n° RR-9224, p. 1-18, https://hal.archives-ouvertes.fr/hal-01924180

[26] F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Scaling matrices and counting the perfect matchings in graphs*, Inria Grenoble Rhône-Alpes, March 2018, n° RR-9161, p. 1-22, https://hal.inria.fr/hal-01743802

[27] A. LEGRAND, D. TRYSTRAM, S. ZRIGUI.*Adapting Batch Scheduling to Workload Characteristics: What can we expect From Online Learning?*, Grenoble 1 UGA - Université Grenoble Alpe, October 2018, n° RR-9212, p. 1-23, https://hal.inria.fr/hal-01896121

Other Publications

[28] A. FAURE, M. POQUET, O. RICHARD.*Évaluation d'algorithmes d'ordonnancement par simulation réaliste*, April 2018, working paper or preprint, https://hal.inria.fr/hal-01779936

[29] P. NEYRON.*Grid'5000*, June 2018, working paper or preprint, https://hal.inria.fr/hal-01805122

Team DATASPHERE

Economie des données et des plateformes

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Security and Confidentiality

Table of contents

Team DATASPHERE

Creation of the Team: 2017 January 01

Keywords:

Computer Science and Digital Science:
- A1.5. - Complex systems
- A3.1. - Data
- A3.5. - Social networks

Other Research Topics and Application Domains:
- B3. - Environment and planet
- B8.2. - Connected city
- B8.5. - Smart society
- B9.6. - Humanities
- B9.11. - Risk management

1. Team, Visitors, External Collaborators

Research Scientist

Stephane Grumbach [Team leader, Inria, Senior Researcher, HDR]

External Collaborators

Jean Pascal Bassino [Ecole Normale Supérieure Lyon, until Aug 2018, HDR]
Jean Sylvestre Berge [Univ de Lyon, until Aug 2018, HDR]
Frederick Douzet [Univ Vincennes-Saint Denis]
Olivier Hamant [INRA, HDR]
Kave Salamatian [Univ Savoie Mont-Blanc, HDR]

Technical Staff

Pascal Carrivain [Inria, from Jun 2018 until Nov 2018]

PhD Student

Colin Gerard [Inria, from May 2018]

Administrative Assistant

Sylvie Boyer [Inria]

2. Overall Objectives

2.1. Overall Objectives

The past decade have witnessed an explosion of the amount of data harvested through digital systems and produced by human activity or from the large set of environmental sensors (IoT). These data are collected, analyzed, correlated and transformed to enable innovative services, which have strong, often disruptive, impact on societies. The datasphere is the new space resulting from these data, considered as a whole, independently of their control. It constitutes a dynamic complex system, much like the hydrosphere, where the basic constituents are bits of data in place of water molecules. Although the date of the inception of the datasphere can be debated, it is really at the turn of the century that its role became dominant and challenged the legacy organisation of societies.

The objective of the Datasphere team is to study the transformation of socio-economic and socio-ecological systems triggered by the diffusion of digital services. We propose a holistic view of the datasphere to apprehend global changes at a planetary scale, with a long term perspective, sometimes with a teleologic vision to understand the phenomena at play and model the interactions of the future. We also consider the digital transformation of socio-economic systems in relation with the challenges that the threats on the natural ecosystem of our planet impose on human societies. Both transformations happen contemporarily, and share parallel impact on the tension between local and global, vertical and horizontal.

A major goal of the project is to develop tools and methodologies in order to observe and analyse the ongoing changes induced by digital transformations. These tools are generally software systems that have to process large volume of heterogeneous data in order to harvest relevant metrics. For this purpose we are pursuing big data processing, machine learning, data visualization, cartography and graph analysis methods that are applicable to our specific needs and can be used in broader scopes.

From a methodological point of view, we aim (i) at interdisciplinary research with all relevant disciplines, and in particular social sciences, and (ii) when possible, analysis of large datasets, such as those from network activities, to investigate quantitatively global phenomena. The first aim raises classical difficulties of interdisciplinary research, but is carried on in a very favorable environment, namely the complex system institute, IXXI. For the second aim, we need to develop original data analysis techniques, new metrics on data flows related to social activities, as well as new visualisation methods to show the interdependencies between entities, from States to people and devices.

3. Research Program

3.1. Dynamics of digital transformations

The research program of the Datasphere team aims at understanding the transformations induced by digital systems on socio-economic and socio-ecological organizations. These transformations are very broad and impact a large part of society. Understanding these changes is very ambitious and would require much more resources than those of the team. Interactions with other teams in other disciplines is thus of strategic importance. The research directions we have worked in and will continue to in the coming years are the following.

- The legal and strategic implications of the development of networks, the growing global interdependencies, and the increase of digital flows beyond control.
- The geopolitics of digital systems, data flows and cyber control, the raise of new strategic imbalances, and digital powers (US, China, Russia, etc.)
- The structural consequences of the translation of governance to digital actors, their inclusion into diplomatic forums, and the weakening of sovereignty over territories.

3.2. Foundations of digital economy

- The economy of intermediation and the progressive control of all two-sided and multi-sided markets by remote digital platforms.
- The methodologies for assessing the strategic value of data and evaluating its leverage for the political economy.
- The analysis of Online Advertisement/tracking ecosystems.

3.3. Ecosystems and Anthropocene

- The interdependencies of natural ecosystems and socio-economic systems, and the role of digital systems on measuring and controlling the global natural/social system.
- The role of digital actors in the adaptation and mitigation of climate change.

- The information economy of planetary challenges related to global warming, biodiversity, health monitoring.

3.4. Large scale graph analysis

- Community analysis and extraction, spectral methods.
- Manifold based approaches to large scale graph analysis, optimal transport.
- Information/rumor/fake news propagation in social networks.

4. Application Domains

4.1. Governance

- City governance, local democracy and interaction with citizens.
- Local governance versus global norms and control.
- Strategy beyond public open data.
- Smart city governance.

4.2. CyberStrategy/CyberSecurity

- Cyber-strategy, defense and security in an evolving world shaped by the digital in particular China/Russia/US cyber-strategy.
- Data strategy for the digital economy, cross border intermediation, platform strategie.
- Strategy of Artificial Intelligence, transparency/acceptability/explainability of AI.
- Cartography of the cyberspace.
- Network, BGP security.

4.3. Anthropocene

- Adaptation to the conditions of the anthropocene, digital control of resources and homeostasis.
- Geopolitics of the environmental challenges, adaptation and mitigation.
- Contemporaneity of the digital revolution and global warming.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

Kavé Salamatian has been awarded in 2018 a President's International Fellowship of the Chinese Academy of Sciences.

6. New Software and Platforms

6.1. DNS data analysis

Data analytics tools for DNS data analysis were developed in a cooperation with ICT, Chinese Academy of Sciences in the frame of the thesis of Jingxiu SU [8].

6.1.1. BGP Geopolitics

An observatory of global BGP connectivity has been developed that is used to monitor and detect in real time BGP level attacks. In addition, a set of tools were developed to analyse the structure of information propagation over social networks.

6.1.2. Atlas of Data

A platform to visualize data flows over the planet is under construction. It can be accessed online at https://theatlasofdata.earth/.

7. New Results

7.1. Political economy

We pursued our work on digital platforms and their impact on the structure of socio-economic systems, which results from the capacity to separate data or information from the actors of the physical world. In [9], we showed how the movement above ground of the intermediation activity transforms territories. A global analysis of the geopolitics of technology was presented in [3].

7.2. Anthropocene studies

We have investigated the possible similarities between biological systems and social systems facing shortage of resources, suggesting that the digital revolution might have something to do with the Anthropocene. More comprehensive approaches that rely on digital systems to control society and nudge citizens to adapt their behavior have been developed in Asia. We analyse in particular the social scoring system in China, and Society 5.0 in Japan [6]. An investigation of the world of images and photography in the time of algorithms was conducted in [2].

7.3. Laws and digital

The emergence of digital services affects the legal system. The law is always associated to a territory, while digital systems act remotely over large regions crossing borders to reach the population, imposing new norms. In [1], we suggest that a new framework is necessary to apprehend new phenomena, such as those resulting from the conflicts between global search engines and local rules with respect to the Right to be forgotten for instance.

7.4. Network data analytics

In collaboration with the Chinese Academy of Sciences, we worked on packet processing algorithmic for high speed network measurements. In [5] a packet capture archive system is developed and described. In [4] a theoretical analysis of the TCAM updates delay that is the main shortcoming of TCAM usage in high speed packet processors is presented. Quality of service for network functions were considered in [7].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

The PhD Thesis of Colin Gerard is funded through a contract with DGA (Ministry of Defense).

9. Partnerships and Cooperations

9.1. Regional Initiatives

The team is hosted by IXXI, the Complex System Institute, at ENS Lyon, and strongly involved in the interdisciplinary cooperation promoted by IXXI. Stéphane Grumbach is vice-director of IXXI. Kavé Salamatian is in the Executive committee of the Data Institute of Grenoble Alps Institute, and of the Cyber@Alps Institute of cybersecurity.

9.2. National Initiatives

- Chaire Castex, Ecole Militaire, Paris.
- AMNECYS (Alpine Multidisciplinary NEtwork on CYber-security Studies), University of Grenoble-Alpes.
- GEODE Research team on Geopolitics.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

- RIHN, Research Institute on Humanity and Nature, Kyoto.
- Information School, UC Berkeley.
- ICT, Institute of Computing Technologies, Chinese Academy of Sciences, Beijing.
- CSIRO, Sydney.
- Center for CyberSecurity, University Macquarie, Sydney.
- Center for Internet Human Rights (CIHR), Berlin.

9.4. International Research Visitors

9.4.1. Visits to International Teams

9.4.1.1. Research Stays Abroad

Stéphane Grumbach has been visiting scientist at the Research Institute on Humanity and Nature, RIHN, in Kyoto.

10. Dissemination

10.1. Promoting Scientific Activities

Stéphane Grumbach has been co-director of IXXI since 2014. He is also involved in the Anthropocene Group at ENS Lyon, which promotes interdisciplinary research and teaching activities on issues related to the adaptation to the changes of the natural ecosystem. He is involved in various initiatives to promote scientific knowledge to a wider audience, as well as in cooperation with public administrations (local and national) to face the challenges of the digital revolution.

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Kavé Salamatian Organised the 3rd French-Japan CyberSecurity workshop in Annecy in April 2018.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Kavé Salamatian is professor at Université de Savoie.

Stéphane Grumbach is lecturer at SciPo Paris, where he teaches Master courses (M1, M2) on the Economy of Data. He also regularly gives lectures in universities, including ENA, ENS Lyon, Ecole centrale, Insa Lyon, etc.

10.2.2. Supervision

PhD in progress: Jingxiu Su, DNS data analysis, 2016, directeur de thèse Kave Salamatian

PhD in progress: Colin GERARD, Stratégies d'influence de la Russie sur les réseaux sociaux, 2018 PhD in progress with Institut Français de Géopolitique, sponsored by DGA, Director: Frederick Douzet

10.3. Popularization

Various publications have appeared in journals accessible to a larger audience [3], [2].

11. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] J.-S. Bergé, S. Grumbach, V. Zeno-Zencovich.*Datasphere, Data Flows beyond Control: Challenges for Law and Governance*, in "European Journal of Comparative Law and Governance", 2018, vol. 5, n^o 2, p. 144-178 [*DOI :* 10.1163/22134514-00502001], https://hal-univ-lyon3.archives-ouvertes.fr/hal-01846878

[2] S. Grumbach.*Dans les griffes des algorithmes*, in "Infra-minces: revue de photographie", June 2018, https://hal.inria.fr/hal-01927620

[3] S. Grumbach, C. Renaud.*Géopolitique des technologies*, in "Questions internationales", June 2018, https://hal.inria.fr/hal-01927622

[4] P. He, W. Zhang, H. Guan, K. Salamatian, G. Xie.*Partial Order Theory for Fast TCAM Updates*, in "IEEE/ACM Transactions on Networking", February 2018, vol. PP, n^o 99, p. 1 - 14 [*DOI :* 10.1109/TNET.2017.2776565], https://hal.archives-ouvertes.fr/hal-01666244

[5] J. Su, Z. Li, S. Grumbach, M. Ikram, K. Salamatian, G. Xie.*A cartography of web tracking using DNS records*, in "Computer Communications", December 2018, vol. 134, p. 83-95 [*DOI :* 10.1016/j.comcom.2018.11.008], https://hal.archives-ouvertes.fr/hal-01950449

Invited Conferences

[6] S. Grumbach.*Digital Control and the Earth EcosystemWill the governance of the anthropocene be designed in East Asia?*, in "RIHN - 2018 13th International Symposium Humanities on the Ground: Confronting the Anthropocene in Asia", Kyoto, Japan, December 2018, p. 1-13, https://hal.inria.fr/hal-01927612

International Conferences with Proceedings

[7] A. MARANDI, T. BRAUN, K. SALAMATIAN, N. THOMOS.*A Comparative Analysis of Bloom Filter-based Routing Protocols for Information-Centric Networks*, in "IEEE Symposium on Computer Communications (ISCC'18)", Natal, Brazil, June 2018, https://hal.inria.fr/hal-01872303

[8] J. SU, Z. LI, S. GRUMBACH, K. SALAMATIAN, G. XIE.*Toward Accurate Inference of Web Activities from Passive DNS Data*, in "IWQoS'18 - IEEE/ACM International Symposium on Quality of Service", Banff, Canada, June 2018, https://hal.inria.fr/hal-01872310

Scientific Books (or Scientific Book chapters)

[9] S. GRUMBACH.*De nouveaux intermédiaires : les plates-formes extraterritoriales*, in "L'ère du numérique", J.-P. CHAMOUX (editor), ISTE éditions, June 2018, vol. Vol. 2: L'économie politique à l'épreuve, https://hal.inria.fr/hal-01669282

Project-Team DRACULA

Multi-scale modelling of cell dynamics : application to hematopoiesis

IN COLLABORATION WITH: Institut Camille Jordan

IN PARTNERSHIP WITH:
CNRS

Université Claude Bernard (Lyon 1)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team DRACULA

Creation of the Team: 2010 January 01, updated into Project-Team: 2011 January 01

Keywords:

Computer Science and Digital Science:

A6.1. - Methods in mathematical modeling
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A6.1.4. - Multiscale modeling
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.3.1. - Inverse problems

Other Research Topics and Application Domains:

B1.1.2. - Molecular and cellular biology
B1.1.5. - Immunology
B1.1.7. - Bioinformatics
B1.1.8. - Mathematical biology
B1.1.10. - Systems and synthetic biology
B2.2.1. - Cardiovascular and respiratory diseases
B2.2.3. - Cancer
B2.2.5. - Immune system diseases
B2.2.6. - Neurodegenerative diseases

1. Team, Visitors, External Collaborators

Research Scientists

Mostafa Adimy [Team leader, Inria, Senior Researcher, HDR]
Samuel Bernard [CNRS, Researcher, HDR]
Fabien Crauste [CNRS, Researcher, HDR]
Olivier Gandrillon [CNRS, Senior Researcher, HDR]
Thomas Lepoutre [Inria, Researcher, HDR]
Vitaly Volpert [CNRS, Senior Researcher, HDR]

Faculty Members

Laurent Pujo Menjouet [Univ de Claude Bernard, Associate Professor, HDR]
Léon Tine [Univ de Claude Bernard, Associate Professor]
Celine Vial [Univ de Claude Bernard, Associate Professor, HDR]

PhD Students

Arnaud Bonnaffoux [The Cosmo Company, until Oct 2018]
Loïs Boullu [Univ de Claude Bernard, until Aug 2018, then researcher (ATER UCBL)]
Aurélien Canet [Univ de Claude Bernard]
Ronan Duchesne [Ecole Normale Supérieure Lyon]
Simon Girel [Univ de Lyon, until Aug 2018, then researcher (ATER UCBL)]

Ulysse Herbach [Univ de Claude Bernard, until Sep 2018]
Alexey Koshkin [Inria, from Sep 2018]

Post-Doctoral Fellow

Chloé Audebert [Inria, until Aug 2018]

Administrative Assistant

Claire Sauer [Inria]

2. Overall Objectives

2.1. Presentation

Dracula is a joint research team between Inria, UniversitÃ© Claude Bernard Lyon 1 (UCBL) and CNRS (Institut Camille-Jordan (ICJ, UMR 5208) and Laboratoire de Biologie et Modélisation de la Cellule (LBMC, UMR 5239)).

The Dracula project is devoted to multi-scale modeling in biology and medicine, and more specifically to the development of tools and methods to describe multi-scale processes in biology and medicine. Applications include normal and pathological hematopoiesis (for example leukemia), immune response, and other biological processes, like: tissue renewal, morphogenesis, atherosclerosis, prion disease, hormonal regulation of food intake, and so on. Multi-scale modeling implies simultaneous modeling of several levels of descriptions of biological processes: intra-cellular networks (molecular level), cell behavior (cellular level), dynamics of cell populations (organ or tissue) with the control by other organs (organism) (see Figure 1). Such modeling represents one of the major challenges in modern science due to its importance and because of the complexity of biological phenomena and of the presence of very different interconnected scales.

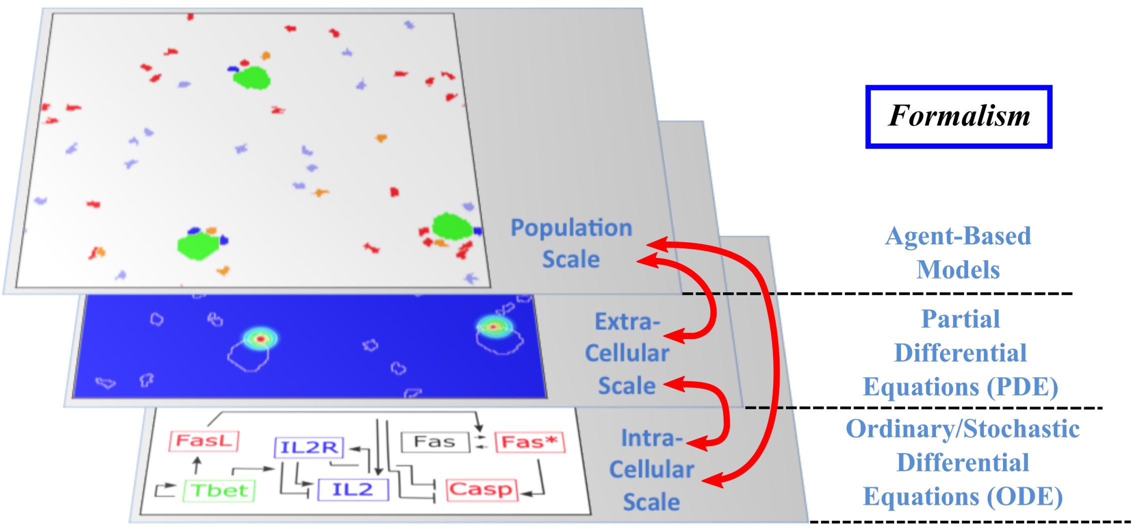


Figure 1. Scheme of multi-scale models of cell dynamics

Although multi-scale modeling holds a great potential for biology and medicine, and despite the fact that a variety of techniques exists to deal with such problems, the complexity of the systems poses new challenges and needs the development of new tools. Moreover, different biological questions usually require different types of multi-scale modeling. The expected results of these studies are numerous. On one hand, they will shed new light on the understanding of specific biological and medical questions (for instance, what is the behavior of hematopoietic stem cells under pathological conditions? Or how to efficiently stimulate an immune response in order to design new vaccines?). On the other hand, the modeling methods developed here for specific processes are relevant to study other complex biological systems. We pay a special attention on developing methods that are not restricted to one or two applications.

An important part of our researches is performed in close collaboration with biologists and physicians in order to stay in contact with the biological and medical goals. The presence, within the project, of a biologist (Olivier Gandrillon) who has acquired over the years the know-how required for interacting with mathematicians is probably one of the main assets of the project. He participates actively in many tasks of our program, stimulates interactions between members of the project and biologists, and everyone benefits from his expertise in molecular and cell biology.

2.2. Keywords

Multi-scale modeling; Hybrid modeling; Mathematical Biology; Computational Biology; Immune response modeling; Normal and pathological hematopoiesis; Multi-scale cancer modeling; Regulatory networks; Reaction-diffusion equation; Structured partial differential equations; Delay differential equations; Agent-based modeling; Dynamical systems.

2.3. Research axis 1: Mathematical modeling for cell population dynamics

2.3.1. Executive summary

Stem cells are essential for development and keep the maintenance of many tissues homeostasis. They are characterized by their ability to self-renew as well as to produce differentiated cells. They vary enormously, for each organ, in their proliferation capacity, their potency to produce different cell lineage and their response to various environmental cues. How a cell will react to a given external signal does not depend only on its current state but also on its environment. Understanding the effect of cell-to-cell heterogeneity and the spatial organization of cell populations is therefore necessary to help keeping the normal function of an organ.

We develop mathematical tools and methods to study cell population dynamics and other biological processes: stability of steady sates, existence of bifurcations, kinetic properties, spatial organization, in finely detailed cell populations. The main tools we use are hybrid discrete-continuous models, reaction-diffusion equations, structured models (in which the population is endowed with relevant structures or traits), delay differential systems, agent-based models. Our team has acquired an international expertise in the fields of analysis of reaction-diffusion and structured equations, particularly integro-differential and delay differential equations.

The mathematical methods we develop are not restricted to hematopoietic system (Research axis 2), and immune response (Research axis 3), rather we apply them in many other biological phenomena, for example: tissue renewal, morphogenesis, prion disease, atherosclerosis, hormonal regulation of food intake, cancer, and others.

2.3.2. Project-team positioning

The focus of this objective is the development, analysis and application of hybrid discrete-continuous, reaction-diffusion and structured partial differential models. The structured equations allow a fine description of a population as some structures (age, maturity, intracellular content) change with time. In many cases, structured equations can be partially integrated to yield integro-differential equations (ordinary or partial differential equations involving non-local integral terms), time-delay differential or time-delay partial differential, or coupled differential-difference models. Analysis of integro-differential and time-delay systems deals with existence of solutions and their stability. Applications are found in the study of normal and pathological hematopoietic system (Research axis 2), immune response (Research axis 3), morphogenesis, prion disease, cancer development and treatment, and generally in tissue renewal problems. Models based on structured equations are especially useful to take into account the effect of finite time cells take to divide, die or become mature. Reaction-diffusion equations are used in order to describe spatial distribution of cell populations. It is a well developed area of research in our team which includes qualitative properties of travelling waves for reaction-diffusion systems with or without delay, and complex nonlinear dynamics.

Our team has developed a solid expertise in mathematical analysis of reaction-diffusion with or without delay and structured equations (in particular, delay differential equations) and one of the most prolific. Other major groups are the teams of Benoit Perthame (Pierre et Marie CURIE University and Mamba, Paris, https://www.inria.fr/en/teams/mamba), Emmanuel Grenier (Ecole normale supérieure de Lyon and NUMED, https://www.inria.fr/en/teams/numed), Odo Diekmann (Utrecht University, The Netherlands, https://www.uu.nl/staff/ODiekmann), Avner Friedman (The Ohio State University, USA, https://people.math.osu.edu/friedman.158/), Jianhong Wu (York University, Canada, http://liam.lab.yorku.ca/), Glenn Webb (Vanderbilt University, Nashville, USA, https://as.vanderbilt.edu/math/bio/glenn-webb), Philip K. Maini (University of Oxford, England, https://people.maths.ox.ac.uk/maini/), Mark Chaplain (University of St Andrews, Scotland, http://www.mcs.st-andrews.ac.uk/~majc/), Nicola Bellomo (University of Turin, Italy, http://staff.polito.it/nicola.bellomo/index.html). Most of the members of all these groups and of our team belong to the same mathematical community working on partial differential equations and dynamical systems with applications to biology and medicine.

2.3.3. Collaborations

- University of Toronto, Canada; Mathematical analysis and applications of reaction-diffusion equations (more than 30 joint papers).
- Institute of Problems of Mechanical Engineering, St.Petersburg, Russia; Dynamics of cell renewal (more than 10 joint papers).
- Department of Cell and Molecular Biology and Department of Forensic Medicine, Stockholm, Sweden; Dynamics of cell generation and turnover (3 joint papers).
- Universities of Tlemcen (Algeria) and Marrakech (Morocco); Delay differential equations (7 joint papers)

2.4. Research axis 2: Multi-scale modeling of hematopoiesis and leukemia

2.4.1. Executive summary

Hematopoiesis is a complex process that begins with hematopoietic stem cells (HSCs) and results in formation of mature cells: red blood cells, white cells and platelets. Blood cells are produced in the bone marrow, from where mature cells are released into the blood stream. Hematopoiesis is based on a balance between cell proliferation (including self-renewal), differentiation and apoptosis. The choice between these three possibilities is determined by intra-cellular regulatory networks and by numerous control mechanisms in the bone marrow or carried out by other organs. Intra-cellular regulatory networks are complex biochemical reactions involving proteins, enzymes and signalling molecules. The deregulation of hematopoiesis can result in numerous blood diseases including leukemia (a cancer of blood cells). One important type of leukemia is Chronic Myeloid Leukemia (CML). The strong tyrosine kinase activity of the BCR-ABL protein is the basis for the main cell effects that are observed in CML: significant proliferation, anti-apoptotic effect, disruption of stroma adhesion properties. This explains the presence in CML blood of a very important number of cells belonging to the myeloid lineage, at all stages of maturation.

Multi-scale modeling in hematopoiesis holds a great potential. A variety of techniques exists to deal with this problem. However, the complexity of the system poses new difficulties and leads to the development of new tools. The expected results of this study are numerous. On one hand, it will shed new light on the different physiological mechanisms that converge toward the continuous regeneration of blood cells, for example: the understanding of deregulation of erythropoiesis (the process of red blood cell production) under drug treatments (this can lead to lack of red blood cells (anemia), or a surplus of red blood cells), the dynamic of leukemic cells under the action of drugs and the control of their resistance to these treatments.

2.4.2. Project team positioning

Multi-scale modeling of hematopoiesis is one of the key points of the project that has started in the early stage of the Dracula team. Investigated by all the team members, it took many years of close discussion with biologists to get the best understanding of the key role played by the most important molecules, hormones, kinase cascade, cell communication up to the latest knowledge. One of the important questions here is to identify particular biological mechanisms (intracellular regulation, control mechanisms) and to integrate them in the different models. Our main work consisted in the development of a hybrid (continuous/discrete) model for red blood cell progenitor proliferation, survival/death, differentiation, and migration. Cells are modeled as discrete objects, and the extracellular medium is described by continuous equations for extracellular concentrations. This is to our knowledge the most complete model for erythropoiesis to date, and the only one using a multi-scale formalism. Other models published by our group and others for hematopoiesis are population-based models, mostly population structured equations (transport partial differential equations or delay differential equations). The interest in modeling hematopoiesis dates back to the 70's and two groups have been responsible for most of development in the past 40 years: Markus Loeffer's team in Leipzig, Germany (Wichmann et al. 1976, in Mathematical Models in Medicine) and Michael Mackey's team at McGill University, Montreal, Canada (Mackey 1978, Blood). Our model differs from population based models in that the regulation is directly modeled at the molecular level (See Figure 1) rather than acting on rates at the population level. Thus we can take into account non-predictable effects of interactions between different molecular pathways and between cells that would otherwise be lost in the global population rates.

Regarding modeling leukemia, we concentrated on Chronic Myeloid Leukemia (CML) and its treatment. We considered models based on ordinary differential equations for the action of the main proteins involved in CML (as BCR-ABL protein), and of transport equations (with or without delay, physiologically structured or not) to represent healthy and leukemic cell populations, take into account many interactions between proteins (especially BCR-ABL), cells (anti-apoptotic effect, etc.). The development of models for CML allowed us to interact with Franck Nicolini in Lyon (Centre Hospitalier de Lyon) and Doron Levy (Maryland University, http://www.math.umd.edu/~dlevy/). Different schools developed models for CML and its treatment. The three leading groups are the ones of Franziska Michor (Harvard School of public health, http://michorlab.dfci.harvard.edu/), Ingo Roeder (Institute for Medical Informatics and Biometry, Dresden, https://tu-dresden.de/med/mf/imb/das-institut) and Michael Mackey (McGill University, http://www.mcgill.ca/mathematical-physiology-lab/).

2.4.3. Collaborations

Members of the team have worked for several years in collaboration with biologists (FranÃ§ois MorlÃ©, University Lyon 1) and hematologists (Charles Dumontet, Lyon and Mark Koury, Nashville, http://www.hematology.org/Thehematologist/Authors/298.aspx) on the Modelling of normal and pathological hematopoiesis .

The work on modeling Leukemia is based on two major collaborations: firstly, an ongoing (since 2011) mathematical collaboration with the University of Maryland through the program Associate Teams Inria project, "Modelling Leukemia" (http://dracula.univ-lyon1.fr/modelling_leukemia.php). Secondly, an ongoing (since 2012) collaboration with a clinician from Hospices Civils de Lyon (Dr. F.E. Nicolini). In this framework, we shall have soon access to the data of the clinical trial PETALs (2×100 patients).

2.5. Research axis 3: Multi-scale modeling of the immune response

2.5.1. Executive summary

Vaccination represents a worldwide health, social and economical challenge as it has allowed the eradication or the strong containment of several devastating diseases over the past century. However to date, most of the effective vaccines rely on the generation of neutralizing antibody responses and such vaccines have proven largely unsuccessful in the prevention against some pathogens, such as HIV or malaria. In such cases, vaccines geared towards the generation of CD8 T cell immunity may provide a better protection. The generation of memory CD8 T cells following antigenic immunization is a long process (lasting up to month in murine preclinical models), therefore strongly slowing the process of vaccine monitoring in preclinical studies. Thus, the dynamical modeling of the CD8 T cell immune response both at the cellular and molecular levels should provide an important tool to better understand the dynamics of the response and to speed-up the process and reduce costs of vaccine development.

However, currently published cellular models of the immune response are either over-simplified, not predicting important parameters of this response, or too complicated for most of their parameters to be accessible for experimental measurements, thus impeding their biological validation. Dynamical models of the CD8 T cell response at the molecular level are very scarce and there is no multi-scale model of the immune response giving insights into both the regulation at the molecular scale and the consequences on cell population dynamics.

The objective of this research axis is therefore to develop a predictive multi-scale model of the CD8 T cell response, by confronting the model at different stages to in vivo-acquired experimental data, in order to be able to investigate the influence of early molecular events on cell population dynamics few days or weeks later.

2.5.2. Project-team positioning

We are aiming at building and analyzing a multi-scale model of the CD8 T cell immune response, from the molecular to the cellular and potentially organismal scale. This consists in describing the dynamics at each scale with relevant formalisms as well as the careful description of the couplings between scales.

Only few research groups are actually working on the CD8 T cell immune response around the world, and none of them deals with multi-scale modeling of this response. A network developed around Alan Perelson's work in theoretical immunology in the last decades, at Los Alamos National Laboratory, and involves mainly people in various US universities or institutes. In Europe, Rob De Boer's group (http://theory.bio.uu.nl/rdb/) of theoretical immunology in Utrecht, Netherlands, is the historical leader in the CD8 T cell dynamics modeling. We considered the models developed in these groups when we started our project, and we contributed to improve them by using nonlinearities accounting for cell population interactions to regulate the response. Also, our initial focus was on the generation of memory cells associated with vaccine development so we modeled CD8 T cell responses against influenza and vaccinia viruses, whereas other groups usually consider LCMV in its chronic form.

Ron Germain's group at the NIH, and GrÃ©goire Altan-Bonnet in subsequent works, focused on the molecular regulation of the CD4 and CD8 T cell immune responses. In particular, they built the *Simmune* software, which allows the modeling and simulation of molecular interactions (https://www.niaid.nih.gov/research/simmune-project). This software is not really devoted to multi-scale modeling yet it provides an interesting tool to describe molecular interactions. Since our aim is to couple molecular and cellular scales at the tissue level, and we do not want to consider large networks but rather small-simplified informative interaction networks, we are confident that our approach is complementary of these works.

Within Inria project-teams, NUMED develops multi-scale approaches for biological problems, and MAMBA and MONC (https://www.inria.fr/en/teams/monc) mention models of cancer progression and treatment including immune responses. In the first case the methodology is similar, and collaborations between NUMED and DRACULA already exist (both teams are located in Lyon), but applications differ. In the second case, MAMBA and MONC are mainly focused on cancer modeling and up to now are motivated by including an action of the immune system in the fight against cancer, which is very different from what we are developing. However, both modeling approaches are complementary and could lead to interactions, in particular in the light of recent advances in medical research pointing towards an important role - and high expectations - of the immune reaction in fighting cancers. Finally, SISTM (https://www.inria.fr/en/teams/sistm) also focuses on the modeling of the immune response, mainly against HIV, but the motivation is very similar to ours: the objective is to provide tools and methods in order to efficiently develop vaccines. They consider the CD4 T cell response instead of the CD8 T cell response, and biostatistics to achieve their goals instead of multi-scale models, yet even though there is no interaction between SISTM and DRACULA at this moment our methods and objectives are close enough to foreshadow future collaborations.

2.5.3. Collaborations

On this topic our main collaborators are members of Jacqueline Marvel's team in Lyon in the CIRI (Centre International de Recherche en Infectiologie INSERM U1111): Dr. Jacqueline Marvel, head of the team, Dr. Christophe Arpin (CR CNRS), and other technicians and engineers of the team. They are all immunologists, specialists of the CD8 T cell response and of the generation of memory CD8 T cells.

We also interact with private companies: AltraBio (http://www.altrabio.com/), that provides tools for data analysis, and CosmoTech, that develops a modeling and simulating platform that should allow transferring our model on an easy-to-use platform devoted to commercial uses.

2.6. Evolution of research direction during the last evaluation

2.6.1. Reminder of the objectives given for the last evaluation

The aim of this project is the development of modern tools for multi-scale modeling in biological phenomena. During the period 2014-2017, the objectives we had fixed were to develop modern tools for multi-scale modeling of biological phenomena, as detailed hereafter:

1. **Multi-scale modeling of erythropoiesis**, the process of red blood cell production, in order to describe normal, stress, and pathological erythropoiesis, using mathematical and computational models. This led to:
2. **The modeling of hemoglobin instability** in dialysis patients: Thomas Lepoutre has been progressively taking part in this theme through a collaboration with P. Kim (University of Sydney, Australia);
3. **Multi-scale modeling of the CD8 T cell immune response**, in order to develop a predictive model of the CD8 T cell response, by confronting the model at different stages to in vivo-acquired experimental data;
4. **Population dynamics modeling**, with the aim to develop general mathematical tools to study them. The main tools we were using were structured equations, in which the cell population is endowed with relevant structures, or traits. We identified limitations in using these formalisms, this is why we started developing multi-scale approaches;
5. **Modeling of Chronic Myeloid Leukemia** (CML) **treatment**, using ordinary differential equations models. Our team had already developed a first model of mutant leukemic cells being resistant to chemotherapy. A next step would be to identify the parameters using experimental data;
6. **Multi-scale modeling** carried out **on the basis of hybrid discrete-continuous models**, where dissipative particle dynamics (DPD) are used in order to describe individual cells and relatively small cell populations, partial differential equations (PDE) are used to describe concentrations of biochemical substances in the extracellular matrix, and ordinary differential equations for intracellular regulatory networks (Figure 1). An emphasis would be made on developing codes that are both flexible and powerful enough to implement variants of the model, perform simulations, produce desired outputs, and provide tools for analysis; to do so:
7. We planned to contribute to a recent project named *chronos*, whose code (written in C++) represents heterogeneous populations of individual cells evolving in time and interacting physically and biochemically, and the objective is to make the code flexible enough to implement different formalisms within the same model, so that different components of the model can be represented in the most appropriate way;
8. **Partial differential equations** (PDE) **analysis**, with a focus on reaction-diffusion equations, transport equations (hyperbolic PDEs) in which the structure can be age, maturity, protein concentration, etc., with particular cases where transport equations are reduced to delay differential equations (DDE).

2.6.2. Comments on these objectives over the evaluation period

We have had strong contributions to objectives 1, 3, 4, 5, and consequently to objective 6, as well as to objective 8, as mentioned in previous sections. These contributions represented the core of the team's research activity over the evaluation period, as stressed by our publications. It is however noticeable that multi-scale modeling of the immune response and of pathological hematopoiesis (leukemia) has come to represent a proportionally more important part of our activity.

Objective 2 has been cancelled few months after the previous evaluation, following meetings with clinicians who did not show any particular interest in our approaches. The modeling of chronic myeloid leukemia instead took a bigger part of the team's research activity, both project being at the time coordinated by Thomas Lepoutre.

Objective 7 has been pursued, the project *chronos* evolved to a better defined project *SiMuScale* that is currently being developed and aims at structuring the team's activity and providing a simulation platform that could be adapted to various biological questions necessitating multi-scale modeling.

2.6.3. Objectives for the next four years

The main objectives for the next four years are to continue to improve the 3 previous points: **1)** Mathematical and computational modeling for cell population dynamics; **2)** Multi-scale modeling of hematopoiesis and leukemia; **3)** Multi-scale modeling of the immune response. In addition, we will pursue our effort to develop a simulation platform for multi-scale models (*SiMuScale*) and we intend to develop the use of mixed effect models and other statistical approaches to deal with the challenges offered by modern biology, in particular the generation of single cell data.

3. Research Program

3.1. Mixed-effect models and statistical approaches

Most of biological and medical data our team has to deal with consist in time series of experimental measurements (cell counts, gene expression level, etc.). The intrinsic variability of any biological system complicates its confrontation to models. The trivial use of means, eliminating the data variance, is but a second-best solution. Furthermore, the amount of data that can be experimentally generated often limits the use of classical mathematical approaches because model's identifiability or parameter identifiability cannot be obtained. In order to overcome this issue and to efficiently take advantage of existing and available data, we plan to use mixed effect models for various applications (for instance: leukemia treatment modeling, immune response modeling). Such models were initially developed to account for individual behaviors within a population by characterizing distributions of parameter values instead of a unique parameter value. We plan to use those approaches both within that frame (for example, taking into account longitudinal studies on different patients, or different mice) but also to extend its validity in a different context: we will consider different ex vivo experiments as being “different individuals”: this will allow us to make the most of the experience-to-experience variations.

Such approaches need expertise in statistics to be correctly implemented, and we will rely on the presence of Céline Vial in the team to do so. Céline Vial is an expert in applied statistics and her experience already motivated the use of better statistical methods in various research themes. The increasing use of single cell technologies in biology make such approaches necessary and it is going to be critical for the project to acquire such skills.

3.2. Development of a simulation platform

We have put some effort in developing the *SiMuScale* platform, a software coded in C++ dedicated to exploring multiscale population models, since 2014. In order to answer the challenges of multi-scale modeling it is necessary to possess an all-purpose, fast and flexible modeling tool, and *SiMuScale* is the choice we made.

Since it is based on a core containing the simulator, and on plug-ins that contain the biological specifications of each cell, this software will make it easier for members of the team – and potentially other modelers – to focus on the model and to capitalize on existing models, which all share the same framework and are compatible with each other. Within the next four years, *SiMuScale* should be widely accessible and daily used in the team for multi-scale modeling. It will be developed into a real-case context, the modeling of the hematopoietic stem cell niche, in collaboration with clinicians (Eric Solary, INSERM) and physicists (Bertrand Laforge, UPMC).

3.3. Mathematical and computational modeling

Multi-scale modeling of hematopoiesis is one of the key points of the project that has started in the early stage of the Dracula team. Investigated by the team members, it took many years of close discussion with biologists to get the best understanding of the key role played by the most important molecules, hormones, kinase cascade, cell communication up to the latest knowledge. An approach that we used is based on hybrid discrete-continuous models, where cells are considered as individual objects, intracellular regulatory networks are described with ordinary differential equations, extracellular concentrations with diffusion or diffusion-convection equations (see Figure 1). These modeling tools require the expertise of all team members to get the most qualitative satisfactory model. The obtained models will be applied particularly to describe normal and pathological hematopoiesis as well as immune response.

3.4. From hybrid dynamics to continuum mechanics

Hybrid discrete-continuous methods are well adapted to describe biological cells. However, they are not appropriate for the qualitative investigation of the corresponding phenomena. Therefore, hybrid model approach should be combined with continuous models. If we consider cell populations as a continuous medium, then cell concentrations can be described by reaction-diffusion systems of equations with convective terms. The diffusion terms correspond to a random cell motion and the reaction terms to cell proliferation, differentiation and death. We will continue our studies of stability, nonlinear dynamics and pattern formation. Theoretical investigations of reaction-diffusion models will be accompanied by numerical simulations and will be applied to study cell population dynamic.

3.5. Structured partial differential equations

Hyperbolic problems are also of importance when describing cell population dynamics. They are structured transport partial differential equations, in which the structure is a characteristic of the considered population, for instance age, size, maturity, etc. In the scope of multi-scale modeling, protein concentrations as structure variables can precisely indicate the nature of cellular events cells undergo (differentiation, apoptosis), by allowing a representation of cell populations in a multi-dimensional space. Several questions are still open in the study of this problem, yet we will continue our analysis of these equations by focusing in particular on the asymptotic behavior of the system (stability, oscillations) and numerical simulations.

3.6. Delay differential equations

The use of age structure in PDE often leads to a reduction (by integration over the age variable) to delay differential equations. Delay differential equations are particularly useful for situations where the processes are controlled through feedback loops acting after a certain time. For example, in the evolution of cell populations the transmission of control signals can be related to some processes as division, differentiation, maturation, apoptosis, etc. Delay differential equations offer good tools to study the behavior of the systems. Our main investigation will be the effect of perturbations of the parameters, as cell cycle duration, apoptosis, differentiation, self-renewal, etc., on the behavior of the system, in relation for instance with some pathological situations. The mathematical analysis of delay differential equations is often complicated and needs the development of new criteria to be performed.

3.7. Multi-scale modeling of the immune response

The main objective of this part is to develop models that make it possible to investigate the dynamics of the adaptive CD8 T cell immune response, and in particular to focus on the consequences of early molecular events on the cellular dynamics few days or weeks later: this would help developing predictive tools of the immune response in order to facilitate vaccine development and reduce costs. This work requires a close and intensive collaboration with immunologist partners.

We recently published a model of the CD8 T cell immune response characterizing differentiation stages, identified by biomarkers, able to predict the quantity of memory cells from early measurements ([32]). In parallel, we improved our multiscale model of the CD8 T cell immune response, by implementing a full differentiation scheme, from naïve to memory cells, based on a limited set of genes and transcription factors.

Our first task will be to infer an appropriate gene regulatory network (GRN) using single cell data analysis (generate transcriptomics data of the CD8 T cell response to diverse pathogens), the previous biomarkers we identified and associated to differentiation stages, as well as piecewise-deterministic Markov processes (Ulysse Herbach's PhD thesis, ongoing).

Our second task will be to update our multiscale model by first implementing the new differentiation scheme we identified ([32]), and second by embedding CD8 T cells with the GRN obtained in our first task (see above). This will lead to a multi-scale model incorporating description of the CD8 T cell immune response both at the molecular and the cellular levels (Simon Girel's PhD thesis, ongoing).

In order to further develop our multiscale model, we will consider an agent-based approach for the description of the cellular dynamics. Yet, such models, coupled to continuous models describing GRN dynamics, are computationally expensive, so we will focus on alternative strategies, in particular on descriptions of the cellular dynamics through both continuous and discrete models, efficiently coupled. Using discrete models for low cell numbers and continuous (partial differential equations) models for large cell numbers, with appropriate coupling strategies, can lead to faster numerical simulations, and consequently can allow performing intense parameter estimation procedures that are necessary to validate models by confronting them to experimental data, both at the molecular and cellular scales.

The final objective will be to capture CD8 T cell responses in different immunization contexts (different pathogens, tumor) and to predict cellular outcomes from molecular events.

3.8. Dynamical network inference from single-cell data

Up to now, all of our multiscale models have incorporated a dynamical molecular network that was build "by hand" after a thorough review of the literature. It would be highly valuable to infer it directly from gene expression data. However, this remains very challenging from a methodological point of view. We started exploring an original solution for such inference by using the information contained within gene expression distributions. Such distributions can be acquired through novel techniques where gene expression levels are quantified at the single cell level. We propose to view the inference problem as a fitting procedure for a mechanistic gene network model that is inherently stochastic and takes not only protein, but also mRNA levels into account. This approach led to very encouraging results [34] and we will actively pursue in that direction, especially in the light of the foreseeable explosion of single cell data.

3.9. Leukemia modeling

Imatinib and other tyrosine kinase inhibitors (TKIs) have marked a revolution in the treatment of Chronic Myelogenous Leukemia (CML). Yet, most patients are not cured, and must take their treatment for life. Deeper mechanistic understanding could improve TKI combination therapies to better control the residual leukemic cell population. In a collaboration with the Hospital Lyon Sud and the University of Maryland, we have developed mathematical models that integrate CML and an autologous immune response ([29], [30] and [31]). These studies have lent theoretical support to the idea that the immune system plays a rôle in maintaining remission over long periods. Our mathematical model predicts that upon treatment discontinuation, the immune system can control the disease and prevent a relapse. There is however a possibility for relapse via a sneak-though mechanism [29]. Research in the next four years will focus in the Phase III PETALS trial. In the PETALS trial (https://clinicaltrials.gov/ct2/show/NCT02201459), the second generation TKI Nilotinib is combined with Peg-IFN, an interferon that is thought to enhance the immune response. We plan to: 1) Adapt the model to take into account the early dynamics (first three months). 2) Use a mixed-effect approach to analyse the effect of the combination, and find population and individual parameters related to treatment efficacy and immune system response. 3) Optimise long-term treatment strategies to reduce or cease treatment and make personalised predictions based on mixed-effect parameters, to minimise the long-term probability of relapse.

4. New Software and Platforms

4.1. CelDyn

KEYWORDS: Modeling - Bioinformatics - Biology
FUNCTIONAL DESCRIPTION: Software "Celdyn" is developed in order to model cell population dynamics for biological applications. Cells are represented either as soft spheres or they can have more complex structure. Cells can divide, move, interact with each other or with the surrounding medium. Different cell types can be introduced. When cells divide, the types of daughter cells are specified. A user interface is developed.

- Participants: Alen Tosenberger, Laurent Pujo-Menjouet, Nikolai Bessonov and Vitaly Volpert
- Contact: Vitaly Volpert

5. New Results

5.1. Oscillations and asymptotic convergence for a delay differential equation modeling platelet production

In [13], a model for platelet production is introduced for which the platelet count is described by a delay differential equation $P'(t) = -\gamma P(t) + f(P(t))g(P(t-r))$ where f and g are positive decreasing functions. First, the authors study the oscillation of the solutions around the unique equilibrium of the equation above, obtaining an inequality implying such an oscillation. They also obtain provide a condition such that this inequality is necessary and sufficient for oscillation. This result is compared to already existing results and the biological meaning of the inequality is studied. The authors also present a result on the asymptotic convergence of the solutions. This result depends on the behavior of the solution for $t \in [0, r]$, and the authors provide an analysis of the link between this behavior and the initial conditions in the case of a simpler model.

5.2. Meningioma growth dynamics assessed by radiocarbon retrospective birth dating

It is not known how long it takes from the initial neoplastic transformation of a cell to the detection of a tumor, which would be valuable for understanding tumor growth dynamics. We have assessed the age and growth dynamics in patients with WHO grade I meningiomas by combining retrospective birth-dating of cells by analyzing incorporation of nuclear-bomb-test-derived 14C, analysis of cell proliferation, cell density, MRI imaging and mathematical modeling. We provide an integrated model of the growth dynamics of benign meningiomas. The mean age of WHO grade I meningiomas was 22.1$\pm$ 6.5 years. We conclude that WHO grade I meningiomas are very slowly growing brain tumors, which are resected in average two decades after time of origination. [18]

5.3. Existence and stability of periodic solutions of an impulsive differential equation and application to CD8 T-cell differentiation

In this article [16], we study a scalar impulsive differential equation (IDE) with the aim of studying the effects of uneven molecular partitioning upon cell mitosis on CD8 T-cell differentiation. To do so, we introduce mathematical results that stand for a more general class of IDE, then apply them to our IDE and discuss those results with regard to the initial biological problem.

5.4. Investigating the role of the experimental protocol in phenylhydrazine-induced anemia on mice recovery

Erythropoiesis, the process of production of red blood cells, is performed through complex regulatory processes. We proposed an earlier model describing stress erythropoiesis in mice [33]. This model, based on the description of erythroid progenitor and erythrocyte dynamics using delay equations, led us to conclude on the quantitative importance of self-renewal. In [6], we refined this previous approaches by taking into account a more mechanistic description of the induction of anemia via phenylhydrazine injection. This led us to revisit some of our initial hypothesis regarding self-renewal regulation.

5.5. Generalizing a mathematical model of prion aggregation allows strain coexistence and co-stability by including a novel misfolded species

Prions are proteins capable of adopting misfolded conformations and transmitting these conformations to other normally folded proteins. A distinct feature of prion propagation is the existence of different phenotypical variants, called strains. In order to conform to biological observations of strain coexistence and co-stability, we develop in [19] an extension of the classical model by introducing a novel prion species consistent with biological studies.

5.6. Analysis and Numerical Simulation of a Polymerization Model with Possible Agglomeration Process

The purpose of [20] is to provide analytical and numerical results for a general polymerization model with lengthening process by agglomeration. 2D spatial diffusion of monomers is taken into account for the mass transfer between monomers and polymers. The analysis of the model is performed thanks to a double fixed point theorem. Adequate numerical scheme based on a generalization of the anti-dissipative method developed in Goudon (Math. Models Methods Appl. Sci. 23:1177–1215, 2013

5.7. The Origin of Species by Means of Mathematical Modelling

Darwin described biological species as groups of morphologically similar individuals. These groups of individuals can split into several subgroups due to natural selection, resulting in the emergence of new species. Some species can stay stable without the appearance of a new species, some others can disappear or evolve. In [10] we have developed a model which allows us to reproduce the principal patterns in Darwin's diagram. Some more complex evolutionary patterns are also observed. The relation between Darwin's definition of species, stated above, and Mayr's definition of species (group of individuals that can reproduce) is also discussed.

5.8. Improved duality estimates in the time discrete case for cross diffusion models

In [28], time discrete versions of the duality estimates derived by Canizo et al. for parabolic systems have been obtained. They allow the construction of solution with superquadratic reactions terms for cross diffusion models with bounded pressure.

6. Partnerships and Cooperations

6.1. National Initiatives

6.1.1. ANR

- ANR SinCity "Single cell transcriptomics on genealogically identified differentiating cells", 2017-2020.
  Participant: Olivier Gandrillon [Coordinator].
- Olivier Gandrillon participates in the ANR MEMOIRE (head Jacqueline Marvel) dedicated to "MultiscalE MOdeling of CD8 T cell Immune REsponses". 2018-2021.
- Fabien Crauste participates in the ANR MEMOIRE (head Jacqueline Marvel) dedicated to "MultiscalE MOdeling of CD8 T cell Immune REsponses". 2018-2021.
- Thomas Lepoutre is a member of the ANR KIBORD (head L. Desvillettes) dedicated to “kinetic and related models in biology". 2014-2018: https://www.ljll.math.upmc.fr/kibord/.

6.1.2. Other projects

- Association France Alzheimer Sciences Médicales: PAMELA “Prion et Alzheimer : Modélisation et Expérimentation d’une Liaison Agressive”, 2014-2017 (https://www.youtube.com/watch?v=X0mLf8IJhV4).
  Participants: Mostafa Adimy, Samuel Bernard, Thomas Lepoutre, Laurent Pujo Menjouet [Coordinator], Léon Tine.
- Thomas Lepoutre is a member of the ERC MESOPROBIO (head V. Calvez) dedicated to "Mesoscopic models for propagation in biology". 2015-2020: .

6.2. European Initiatives

6.2.1. FP7 & H2020 Projects

Fabien Crauste and Olivier Gandrillon participates in the EU RTN network COSMIC (head Antpoine. van Kampen) dedicated to "Combatting disorders of adaptive immunity with systems medicine"). 2018-2021. https://cosmic-h2020.eu

6.3. International Initiatives

6.3.1. MODELLING_LEUKEMIA

Title: Modeling quiescence and drug resistance in Chronic Myeloid Leukemia

International Partner (Institution - Laboratory - Researcher):

University of Maryland (United States) - Center for Scientific Computation and Mathematical Modeling (CSCAMM) - Levy Doron

Start year: 2016

See also: http://dracula.univ-lyon1.fr/modelling_leukemia.php

This project is dedicated to the mathematical modelling of chronic myeloid leukemia and treatment effects. We focus especially on the interplay between the immune response and treatment. This has a potential impact on the study of treatment cessation. This work is conducted in close collaboration with a clinician.

6.3.2. Participation in Other International Programs

6.3.2.1. Indo-French Center of Applied Mathematics

Mathematical modeling of hematopoiesis process in application to chronic and acute myelogenous leukemia

Title: Mathematical modeling of hematopoiesis process in application to chronic and acute myelogenous leukemia

International Partner (Institution - Laboratory - Researcher):

(India)- Subhas Khajanchi

Duration: 2018 - 2021

Start year: 2018

6.4. International Research Visitors

6.4.1. Visits of International Scientists

Antone dos Santos Benedito, PHD student on Adding temperature and anthropogenic actions in the study of spatial-temporal behavior of insectplague Chrysodeixis Includens. Institute of Biosciences, São Paulo State University (UNESP), Botucatu, Brazil not a team member but visiting for 6 months (from September 1st 2018 to February 28, 2019)

6.4.2. Visits to International Teams

Paul Lemarre is visiting University of Merced in 2018-2019.

7. Dissemination

7.1. Promoting Scientific Activities

7.1.1. Scientific Events Organisation

7.1.1.1. General Chair, Scientific Chair

- Olivier Gandrillon International Conference of Systems Biology 2018, Lyon, France http://icsb2018-france.com/.

7.1.1.2. Member of the Organizing Committees

- Fabien Crauste, International Conference of Systems Biology 2018, Lyon, France http://icsb2018-france.com/.
- Laurent Pujo Menjouet, co-organization of the minisymposium « Hematopoiesis and its disease » at ECMTB 2018, 11th European Conference on Mathematical and Theoretical Biology, Lisbon, Portugal, 23 – 27 July , 2018

7.1.2. Scientific Events Selection

7.1.2.1. Reviewer

M .Adimy : 2nd International Conference on Applied Mathematics (ICAM 2018), 2018 Fez-Morocco,

7.1.3. Journal

7.1.3.1. Member of the Editorial Boards

- O. Gandrillon Associate editor for BMC research notes
- M. Adimy Journal of Nonlinear Systems and Applications; Chinese Journal of Mathematics.
- L. Pujo Menjouet :Associate editor of PLOS ONE Journal of Thoerical Biology, Mathematical modelling of natural phenomena

7.1.3.2. Reviewer - Reviewing Activities

- O. Gandrillon : BioEssays, BioTechniques, BMC Bioinformatics, BMC Genomics, Cell Biology and Toxicology, Genes, Genomics, Journal of the Royal Society Interface, Journal of Theoretical Biology, Molecular BioSystems, Nature Communications, npj Systems Biology and Applications, Plos Computational Biology and Progress in Biophysics and Molecular Biology
- L. Tine Journal of Theoretical Biology
- T. Lepoutre Applied Mathematical Letters, Journal of Differential Equations, Siam Journal of Mathematical Analysis, Marrow
- L. Pujo Menjouet Journal of Theoretical Biology Bulletin of mathematical biology Journal of mathematical biology Plos computational biology
- S. Bernard J Theor Biol, Comput and Appl Math, J Roy Soc Interface

7.1.4. Invited Talks

- S Bernard, Modélisation mathématique de la croissance et de la réparation tissulaire, Fondation Les Treilles, Tourtour FR, 12-17/11/2018
- S Bernard, Workshop Mathematics of Biological rhythms, Northunbria University, Newcastle, 3-5/12/2018
- Thomas Lepoutre : Mathematical perspetives in the biology and therapeutics of cancer, Cirm, Luminy France, https://mathscancer.sciencesconf.org/
- Thomas Lepoutre : Mathematical Challenges in the Analysis of Continuum Models for Cancer Growth, Evolution and Therapy, CMO, Oaxaca MEXICO, https://www.birs.ca/events/2018/5-day-workshops/18w5115
- Mostafa Adimy, 2nd International Conference on Applied Mathematics (ICAM 2018), Fez-Morocco.
- Mostafa Adimy, 2ème Colloque des Mathématiciens Marocains à l’étranger, Marrakech-Morocco.
- Mostafa Adimy, 8th Mathematical and Biological School (Programa Argentino BIOMAT), Córdoba-Argentina.
- Mostafa Adimy, Mathematical Challenges in the Analysis of Continuum Models for Cancer Growth, Evolution and Therapy, Oaxaca, Mexico. https://www.birs.ca/events/2018/5-day-workshops/18w5115
- Laurent Pujo Menjouet, talk at the minisymposium « Topics in structured population dynamics » ECMTB 2018, 11th European Conference on Mathematical and Theoretical Biology, Lisbon, Portugal, 23 – 27 Juillet , 2018
- Laurent Pujo Menjouet, LIAM-IRC-MfPH 2018 Symposium in Structured population models: theory, numerics and applications, 27-29 août 2016

7.1.5. Leadership within the Scientific Community

- O . Gandrillon : Director of BioSyL, the Federative Research Structure for Systems Biology attached to University of Lyon
- T. Lepoutre : Head of the Groupe de Recherches CNRS MAMOVI on applied mathematical modelling in Life Sciences.

7.1.6. Scientific Expertise

- L. Pujo Menjouet: reviewer for ANR

7.1.7. Research Administration

- L.M. Tine :Membre conseil du département de Mathématiques, Lyon 1
- L.M. Tine Co-responsable de l'enseignement TMB (Techniques Mathématiques de Base) du portail PCSI.
- M. Adimy : Comité scientifique (COS) du centre Rhône-Alpes.
- M. Adimy : Comité scientifique (CS) de l'Institut Camille Jordan.
- M. Adimy : Comité des thèses de l'Institut Camille Jordan.
- M. Adimy : Comité d'évaluation de l'Université de Guyane.
- L. Pujo- Menjouet : Responsable de la filière mathématiques pour la biologie et la médecine pour le master 2 math en actions à l'université Claude Bernard Lyon 1,
- L. Pujo- Menjouet : correspondant mobilité international pour le département de mathématiques à l'université Claude Bernard Lyon 1,
- L. Pujo- Menjouet : directeur du portail mathématiques et informatique à l'université Claude Bernard Lyon 1,
- T. Lepoutre member of the CORDI-S commission.

7.2. Teaching - Supervision - Juries

7.2.1. Teaching

Licence: Samuel Bernard: Algèbre Linéaire, 15h, L3, INSA

Licence: Samuel Bernard: EDO-EDP, 15h, L3, INSA

Licence : Ronan Duchesne, Anglais, 5h, L3

Licence : Ronan Duchesne, Modélisation des systèmes biologiques, 22h, L3

Licence : Ronan Duchesne, Bioinformatique, 20h, L3

Licence : Ronan Duchesne, Développement, 4h, L3

Licence : Simon Girel, Fondamentaux des mathématiques 1, 86h, L1, Université Lyon 1, France

Licence : Simon Girel, Introduction à l'analyse numérique, 10h, L1, Université Lyon 1, France

Licence: Laurent Pujo Menjouet, Fondamentaux des mathématiques I, 138h EQTD, L1, UCBL 1, FRANCE

Licence: Laurent Pujo Menjouet, Introduction à l'analyse numérique, 62h EQTD, L2, UCBL 1, FRANCE

Licence: Laurent Pujo Menjouet, bio-mathématiques et modélisation BISM, 10.5h EQTD, L3 UCBL 1, FRANCE

Licence: Laurent Pujo Menjouet, 3ème année biosciences BIM: EDO, 35h EQTD, INSA Lyon, FRANCE

Licence : L. M. Tine, Techniques mathématiques de base, 53h (EqTD), niveau L0, Lyon 1, France.

Licence : L. M. Tine, Techniques mathématiques de base, 62h (EqTD), niveau L1, Lyon 1, France.

Licence : L. M. Tine, Initiation LaTex+ stage, 12h (EqTD), niveau L3, Lyon 1, France.

Master : Samuel Bernard, Population Dynamics, 36h ETD, M2, UCBL, Lyon.

Master : Mostafa Adimy, Population Dynamics, 9h ETD, M2, UCBL, Lyon.

Master : Mostafa Adimy, Epidemiology, 12h ETD, M2, UCBL, Lyon.

Master : Ronan Duchesne, Biologie du développement, 12h, M1

Master : Ronan Duchesne, Statistiques, 2h, M1

Master : Ronan Duchesne, Adaptation, 2h, M1

Master : Ronan Duchesne, Practicals in statistics and modelling for the biosciences, 28h, M2

Master: Thomas Lepoutre, préparation à l'option pour l'agrégation, 45 h eq TD, M2 UCBL 1, FRANCE

Master: Laurent Pujo Menjouet, maths appliquées et statistiques: Systèmes dynamiques, 78h EQTD, M1, UCBL 1, FRANCE,

Master: Laurent Pujo Menjouet, master modélisation des systèmes complexes: modelling biology and medicine, M2, 9h EQTD, ENS-Lyon, FRANCE

Master: Laurent Pujo Menjouet,INSA 4ème année biosciences BIM: ED-EDP, 22h EQTD, M1 INSA Lyon, FRANCE

Master: L. M. Tine, Maths en action, Remise à niveau analyse, 12h (EqTD), niveau M2, Lyon 1, France.

Master: L. M. Tine, Maths en action, épidémiologie, 18h (EqTD), niveau M2, Lyon 1, France.

7.2.2. Supervision

PhD in progress: Aurélien Canet, "Contribution à l'étude de la quantification de la réponse d'une tumeur solide après un traitement par radiothérapie", Université Lyon, since January 2016, encadrants: Larry Bodgi, Nicolas Foray and Laurent Pujo Menjouet.

PhD in progress: Kyriaki Dariva , "Modélisation mathématique des interactions avec le système immunitaire en leucémie myéloide chronique". Université Lyon 1, September 2018, supervisor : Thomas Lepoutre

PhD in progress: Ronan Duchesne, "Vers un modèle multi-échelle de la différentiation cellulaire : Application à la différentiation érythrocytaire", École normale supérieure de Lyon and Université Lyon 1, since September 2016, supervisors: Olivier Gandrillon and Fabien Crauste.

PhD in progress: Alexey Koshkin, Inferring gene regulatory networks from single cell data , ENS de Lyon, since September 2018, supervisors : Olivier Gandrillon and Fabien Crauste.

PhD in progress : Paul Lemarre, " Modélisation des souches de prions". Université Lyon 1, since May 2017, supervisors Laurent Pujo Menjouet et Suzanne Sindi (University of California, Merced)

PhD : Arnaud Bonnafoux, Vers une inférence automatique de réseaux de gènes dynamiques à partir de « mégadonnées » temporelles discrètes acquises sur cellules uniques, Université Lyon 1, October 2018, Olivier Gandrillon (CIFRE with the COSMOTECH).

PhD : Loïs Boullu, Modélisation de la mégacaryopoïèse et applications aux maladies liées â la production des plaquettes, Université Lyon 1, November 2018, Laurent Pujo Menjouet and Jacques Bélair (co-tutelle avec l'Université de Montréal).

PhD : Simon Girel, Modélisation de la réponse immunitaire T CD8 : analyse mathématique et modèles multi-échelles, Université de Lyon, November 2018 , encadrant: Fabien Crauste.

PhD : Ulysse Herbach, Modèles graphiques probabilistes pour l'inférence de réseaux de gènes, Université Lyon 1, September 2018, Olivier Gandrillon, Thibault Espinasse (ICJ) and Anne-Laure Fougères (ICJ).

7.2.3. Juries

We separate the juries of team members from external participations. PhD Defense within the team in 2018

- Arnaud Bonnaffoux (O. Gandrillon supervisor),
- Lois Boullu (L. Pujo Menjouet supervisor, M. Adimy and F. Crauste examiners),
- Ulysse Herbach (O. Gandrillon supervisor)
- Simon Girel (F. Crauste supervisor L. Pujo Menjouet examiner).

- M. Adimy: PhD of Linlin Li, Mathematical analysis of a model of partial differential equations describing the adaptation of mosquitoes facing the usage of insecticides, University of Bordeaux, reviewer.
- M. Adimy: PhD of Zhengyang Zhang, A class of state-dependent delay differential equations and applications to forest growth, University of Bordeaux, examiner.
- V. Volpert: PhD of Guillaume Cantin, Étude de réseaux complexes de systèmes dynamiques dissipatifs ou conservatifs en dimension finie ou infinie. Application à l'analyse des comportements humains en situation de catastrophe. Université Le Havre Normandie, examiner.
- S. Bernard : Charles Rocabert, Etude de l'évolution de micro-organismes bactériens par des approches de modélisation et de simulation informatique, INSA Lyon, (examiner)

7.3. Popularization

7.3.1. Articles and contents

- L. Pujo Menjouet : interview dans Causette (hors série Juillet-Août 2018)
- L. Pujo Menjouet : interview dans le magazine Society (fin juillet 2018)

7.3.2. Interventions

- T. Lepoutre: Visite de l'inria pour les stagiaires de 3e de l'ICJ (plusieurs fois dans l'année, une matinée en général).
- T. Lepoutre: présentation de l'antenne et des équipes à deux classes lors du congrès Maths en Jeans (mars 2018)
- T. Lepoutre participation to MathaLyon interventions
- L. Pujo Menjouet : participation à Math en Jeans édition 2018 (collège de la côte Roannaise, Renaison)
- L. Pujo Menjouet : présentation de deux conférences au collège de Saint Bonnet le château le 21 juin 2018

8. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. BONNAFFOUX.*Inferring gene regulatory networks from dynamic multi-scale data*, Université de Lyon, October 2018, https://hal.archives-ouvertes.fr/tel-01920262

[2] L. BOULLU.*Study of delay differential equations with applications to the regulation of blood platelet production*, Université de Lyon, November 2018, https://hal.inria.fr/tel-01948726

[3] S. GIREL.*Modeling the CD8 T-cell Immune Response : Mathematical Analysis and Multiscale Models*, Université de Lyon, November 2018, https://hal.archives-ouvertes.fr/tel-01941850

[4] U. HERBACH.*From stochastic modelling of gene expression to inference of regulatory networks*, Université de Lyon, September 2018, https://tel.archives-ouvertes.fr/tel-01930398

Articles in International Peer-Reviewed Journal

[5] O. ANGULO, F. CRAUSTE, J. LÓPEZ-MARCOS.*Numerical integration of an erythropoiesis model with explicit growth factor dynamics*, in "Journal of Computational and Applied Mathematics", March 2018, vol. 330, p. 770 - 782 [*DOI :* 10.1016/J.CAM.2017.01.033], https://hal.inria.fr/hal-01646786

[6] O. ANGULO, O. GANDRILLON, F. CRAUSTE.*Investigating the role of the experimental protocol in phenylhydrazine-induced anemia on mice recovery*, in "Journal of Theoretical Biology", January 2018, vol. 437, p. 286 - 298 [*DOI :* 10.1016/J.JTBI.2017.10.031], https://hal.inria.fr/hal-01646792

[7] L. BARBARROUX, P. MICHEL, M. ADIMY, F. CRAUSTE.*A multiscale model of the CD8 T cell immune response structured by intracellular content*, in "Discrete and Continuous Dynamical Systems - Series B (DCDS-B)", November 2018, vol. 23, n^{o} 9, p. 3969-4002, https://hal.inria.fr/hal-01969237

[8] V. A. BELYAEV, J. L. DUNSTER, J. M. GIBBINS, M. PANTELEEV, V. VOLPERT.*Modeling thrombosis in silico: Frontiers, challenges, unresolved problems and milestones*, in "Physics of Life Reviews", November 2018, vol. 26-27, p. 57-95 [*DOI :* 10.1016/J.PLREV.2018.02.005], https://hal.archives-ouvertes.fr/hal-01941469

[9] N. BESSONOV, A. BEUTER, S. TROFIMCHUK, V. VOLPERT.*Estimate of the travelling wave speed for an integro-differential equation*, in "Applied Mathematics Letters", February 2019, vol. 88, p. 103-110 [*DOI :* 10.1016/J.AML.2018.07.037], https://hal.archives-ouvertes.fr/hal-01941455

[10] N. BESSONOV, N. REINBERG, M. BANERJEE, V. VOLPERT.*The Origin of Species by Means of Mathematical Modelling*, in "Acta Biotheoretica", December 2018, vol. 66, n^{o} 4, p. 333-344 [*DOI :* 10.1007/S10441-018-9328-9], https://hal.archives-ouvertes.fr/hal-01941452

[11] A. BEUTER, A. BALOSSIER, S. TROFIMCHUK, V. VOLPERT.*Modeling of post-stroke stimulation of cortical tissue*, in "Mathematical Biosciences", November 2018, vol. 305, p. 146-159 [*DOI :* 10.1016/J.MBS.2018.08.014], https://hal.archives-ouvertes.fr/hal-01941460

[12] G. BOCHAROV, A. MEYERHANS, N. BESSONOV, S. TROFIMCHUK, V. VOLPERT.*Interplay between reaction and diffusion processes in governing the dynamics of virus infections*, in "Journal of Theoretical Biology", November 2018, vol. 457, p. 221-236 [*DOI :* 10.1016/J.JTBI.2018.08.036], https://hal.archives-ouvertes.fr/hal-01941458

[13] L. BOULLU, M. ADIMY, F. CRAUSTE, L. PUJO-MENJOUET.*Oscillations and asymptotic convergence for a delay differential equation modeling platelet production*, in "Discrete and Continuous Dynamical Systems - Series B", 2018, p. 1-27, https://hal.inria.fr/hal-01835435

[14] I. S. CIUPERCA, M. DUMONT, A. LAKMECHE, P. MAZZOCCO, L. PUJO-MENJOUET, H. REZAEI, L. M. TINE.*Alzheimer's disease and prion: analysis of an in vitro mathematical model*, in "Discrete and Continuous Dynamical Systems - Series B (DCDS-B)", 2018, https://hal.inria.fr/hal-01368862

[15] T. GALOCHKINA, M. MARION, V. VOLPERT.*Initiation of reaction–diffusion waves of blood coagulation*, in "Physica D: Nonlinear Phenomena", August 2018, vol. 376-377, p. 160-170 [*DOI :* 10.1016/J.PHYSD.2017.11.006], https://hal.archives-ouvertes.fr/hal-01941465

[16] S. GIREL, F. CRAUSTE.*Existence and stability of periodic solutions of an impulsive differential equation and application to CD8 T-cell differentiation*, in "Journal of Mathematical Biology", June 2018, vol. 76, n^{o} 7,

p. 1765-1795, https://arxiv.org/abs/1803.01674 [*DOI :* 10.1007/S00285-018-1220-3], https://hal.archives-ouvertes.fr/hal-01721728

[17] M. HELAL, A. IGEL-EGALON, A. LAKMECHE, P. MAZZOCCO, A. PERRILLAT-MERCEROT, L. PUJO-MENJOUET, H. REZAEI, L. MATAR TINE.*Stability analysis of a steady state of a model describing Alzheimer's disease and interactions with prion proteins*, in "Journal of Mathematical Biology", August 2018, p. 1-25 [*DOI :* 10.1007/S00285-018-1267-1], https://hal.archives-ouvertes.fr/hal-01946610

[18] H. B. HUTTNER, O. BERGMANN, M. SALEHPOUR, R. E. CHEIKH, M. NAKAMURA, A. TORTORA, P. HEINKE, R. CORAS, E. ENGLUND, I. Y. EYÜPOGLU, J. B. KURAMATSU, S. S. ROEDER, S. P. KLOSKA, I. MUEHLEN, A. DOERFLER, S. SCHWAB, G. POSSNERT, S. BERNARD, J. FRISÉN.*Meningioma growth dynamics assessed by radiocarbon retrospective birth dating*, in "EBioMedicine", 2018, vol. 27, p. 176–181 [*DOI :* 10.1016/J.EBIOM.2017.12.020], https://hal.archives-ouvertes.fr/hal-01671339

[19] P. LEMARRE, L. PUJO-MENJOUET, S. S. SINDI.*Generalizing a mathematical model of prion aggregation allows strain coexistence and co-stability by including a novel misfolded species*, in "Journal of Mathematical Biology", 2018 [*DOI :* 10.1007/S00285-018-1280-4], https://hal.archives-ouvertes.fr/hal-01858724

[20] L. MATAR TINE, B. LÈYE.*Analysis and numerical simulation of a polymerization model with aggregation and Ostwald ripening for the growth phenomenon*, in "Acta Applicandae Mathematicae", October 2018, p. 1-28 [*DOI :* 10.1007/S10440-018-0222-X], https://hal.archives-ouvertes.fr/hal-01933553

[21] M. SEN, V. VOLPERT, V. VOUGALTER.*Existence of waves for a reaction–diffusion–dispersion system*, in "Nonlinear Analysis: Theory, Methods and Applications", March 2019, vol. 180, p. 52-66 [*DOI :* 10.1016/J.NA.2018.09.011], https://hal.archives-ouvertes.fr/hal-01941461

[22] S. TROFIMCHUK, V. VOLPERT.*Traveling Waves for a Bistable Reaction-Diffusion Equation with Delay*, in "SIAM Journal on Mathematical Analysis", January 2018, vol. 50, n^o 1, p. 1175-1199 [*DOI :* 10.1137/17M1115587], https://hal.archives-ouvertes.fr/hal-01941474

[23] G. VOGIN, T. BASTOGNE, L. BODGI, J. GILLET-DAUBIN, A. CANET, S. PEREIRA, N. FORAY.*The pATM Immunofluorescence assay: a high-performance radiosensitivity assay to predict post radiotherapy overreactions*, in "International Journal of Radiation Oncology - Biology - Physics", July 2018, vol. 101, n^o 3, p. 690–693 [*DOI :* 10.1016/J.IJROBP.2018.03.047], https://hal.archives-ouvertes.fr/hal-01757771

[24] V. VOUGALTER, V. VOLPERT.*On the Existence in the Sense of Sequences of Stationary Solutions for Some Systems of Non-Fredholm Integro-differential Equations*, in "Mediterranean Journal of Mathematics", October 2018, vol. 15, n^o 5, https://hal.archives-ouvertes.fr/hal-01941463

[25] V. VOUGALTER, V. VOLPERT.*Solvability of Some Integro-Differential Equations with Anomalous Diffusion in Two Dimensions*, in "Journal of Mathematical Sciences", December 2018, vol. 235, n^o 3, p. 243-255 [*DOI :* 10.1007/978-3-319-47512-7_9], https://hal.archives-ouvertes.fr/hal-01941457

Scientific Books (or Scientific Book chapters)

[26] G. BOCHAROV, V. VOLPERT, B. LUDEWIG, A. MEYERHANS.*Mathematical Immunology of Virus Infections*, Springer International Publishing, 2018, XV, 245 [*DOI :* 10.1007/978-3-319-72317-4], https://hal.archives-ouvertes.fr/hal-01941475

Other Publications

[27] I. S. Ciuperca, M. Dumont, A. Lakmeche, P. Mazzocco, L. Pujo-Menjouet, H. Rezaei, L. Matar Tine.*Alzheimer's disease and prion: an in vitro mathematical model*, February 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01708659

[28] T. Lepoutre.*Improved duality estimates: time discrete case and applications to a class of cross-diffusion systems*, September 2018, https://arxiv.org/abs/1812.01296 - working paper or preprint, https://hal.inria.fr/hal-01943576

References in notes

[29] A. Besse, G. D. Clapp, S. Bernard, F. E. Nicolini, D. Levy, T. Lepoutre.*Stability Analysis of a Model of Interaction Between the Immune System and Cancer Cells in Chronic Myelogenous Leukemia*, in "Bulletin of Mathematical Biology", 2017, p. 1–27 [*DOI :* 10.1007/s11538-017-0272-7], https://hal.archives-ouvertes.fr/hal-01539364

[30] A. Besse, T. Lepoutre, S. Bernard.*Long-term treatment effects in chronic myeloid leukemia*, in "Journal of Mathematical Biology", 2017, p. 1-26 [*DOI :* 10.1007/s00285-017-1098-5], https://hal.archives-ouvertes.fr/hal-01447781

[31] G. D. Clapp, T. Lepoutre, R. El Cheikh, S. Bernard, J. Ruby, H. Labussière-Wallet, F. E. Nicolini, D. Levy.*Implication of the Autologous Immune System in BCR-ABL Transcript Variations in Chronic Myelogenous Leukemia Patients Treated with Imatinib.*, in "Cancer Research", October 2015, vol. 75, n^{o} 19, p. 4053-62 [*DOI :* 10.1158/0008-5472.CAN-15-0611], https://hal.inria.fr/hal-01251396

[32] F. Crauste, J. Mafille, L. Boucinha, S. Djebali, O. Gandrillon, J. Marvel, C. Arpin.*Identification of Nascent Memory CD8 T Cells and Modeling of Their Ontogeny*, in "Cell Systems", February 2017, vol. 4, n^{o} 3, p. 306-317 [*DOI :* 10.1016/j.cels.2017.01.014], https://hal.inria.fr/hal-01409637

[33] F. Crauste, L. Pujo-Menjouet, S. Génieys, C. Molina, O. Gandrillon.*Adding self-renewal in committed erythroid progenitors improves the biological relevance of a mathematical model of erythropoiesis*, in "Journal of theoretical biology", 2008, vol. 250, n^{o} 2, p. 322–338

[34] U. Herbach, A. Bonnaffoux, T. Espinasse, O. Gandrillon.*Inferring gene regulatory networks from single-cell data: a mechanistic approach*, in "BMC Systems Biology", November 2017, vol. 11, n^{o} 1, 105 [*DOI :* 10.1186/s12918-017-0487-0], https://hal.archives-ouvertes.fr/hal-01646910

Team ELAN

modELing the Appearance of Nonlinear phenomena

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Numerical schemes and simulations

Table of contents

Team ELAN

Creation of the Team: 2017 October 01

Keywords:

Computer Science and Digital Science:

A2.5. - Software engineering
A5.5.4. - Animation
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.4. - Multiscale modeling
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.5. - Numerical Linear Algebra
A6.2.6. - Optimization
A6.2.8. - Computational geometry and meshes
A6.3.1. - Inverse problems
A6.5. - Mathematical modeling for physical sciences
A9.2. - Machine learning

Other Research Topics and Application Domains:

B3.3.1. - Earth and subsoil
B5.5. - Materials
B9.2.2. - Cinema, Television
B9.5.3. - Physics
B9.5.5. - Mechanics

1. Team, Visitors, External Collaborators

Research Scientist

Florence Descoubes [Team leader, Inria, Researcher, HDR]

Technical Staff

Laurence Boissieux [SED Inria, 20%]
Eric Madaule [Inria, until Jul 2018]
Victor Romero Gramegna [Inria]

PhD Students

Raphael Charrondiere [Univ Grenoble Alpes, from Sep 2018]
Mickael Ly [Inria]
Abdullah Haroon Rasheed [Inria]

Administrative Assistant

Diane Courtiol [Inria]

2. Overall Objectives

2.1. Overall Objectives

ELAN is a new joint team of Inria and Laboratoire Jean Kuntzmann (UMR 5224), with an original positioning across Computer Graphics and Computational Mechanics. The team is focussed on the design of predictive, robust, efficient, and controllable numerical models for capturing the shape and motion of visually rich mechanical phenomena, such as the coiling of a viscous rod, the buckling of an elastic plate, or the entangling of large fiber assemblies. Target applications encompass the digital entertainment industry (e.g., feature film animation, special effects), as well as virtual prototyping for the mechanical engineering industry (e.g., aircraft manufacturing, cosmetology); though very different, these two application fields require predictive and scalable models for capturing complex mechanical phenomena at the macroscopic scale. An orthogonal objective is the improvement of our understanding of natural physical and biological processes involving slender structures, through active collaborations with soft matter physicists. To achieve its goals, the team strives to master as finely as possible the entire modeling pipeline, involving a pluridisciplinary combination of scientific skills across Mechanics and Physics, Applied Mathematics, and Computer Science.

3. Research Program

3.1. Discrete modeling of slender elastic structures

For the last 15 years, we have investigated new discrete models for solving the Kirchhoff dynamic equations for thin elastic rods [7], [9], [12]. All our models share a curvature-based spatial discretization, allowing them to capture inextensibility of the rod intrinsically, without the need for adding any kinematic constraint. Moreover, elastic forces boil down to linear terms in the dynamic equations, making them well-suited for implicit integration. Interestingly, our discretization methodology can be interpreted from two different points-of-views. From the finite-elements point-of-view, our strain-based discrete schemes can be seen as discontinuous Galerkin methods of zero and first orders. From the multibody system dynamics point of view, our discrete models can be interpreted as deformable Lagrangian systems in finite dimension, for which a dedicated community has started to grow recently [33]. We note that adopting the multibody system dynamics point of view helped us formulate a linear-time integration scheme [8], which had only be investigated in the case of multibody rigid bodies dynamics so far.

3.1.1. High-order spatial discretization schemes for rods, ribbons and shells

Our goal is to investigate similar high-order modeling strategies for surfaces, in particular for the case of inextensible ribbons and shells. Elastic ribbons have been scarcely studied in the past, but they are nowadays drawing more and more the attention from physicists [22], [31]. Their numerical modeling remains an open challenge. In contrast to ribbons, a huge litterature exists for shells, both from a theoretical and numerical viewpoints (see, e.g., [26], [13]). However, no real consensus has been obtained so far about a unified nonlinear shell theory able to support large displacements. In [10] we have started building an inextensible shell patch by taking as degrees of freedom the curvatures of its mid-surface, expressed in the local frame. As in the super-helix model, we show that when taking curvatures uniform over the element, each term of the equations of motion may be computed in closed-form; besides, the geometry of the element corresponds to a cylinder patch at each time step. Compared to the 1D (rod) case however, some difficulties arise in the 2D (plate/shell) case, where compatibility conditions are to be treated carefully.

3.1.2. Numerical continuation of rod equilibria in the presence of unilateral constraints

In Alejandro Blumentals' PhD thesis [11], we have adopted an optimal control point of view on the static problem of thin elastic rods, and we have shown that direct discretization methods [0] are particularly well-suited for dealing with scenarios involving both bilateral and unilateral constraints (such as contact). We would like to investigate how our formulations extend to continuation problems, where the goal is to follow a certain branch of equilibria when the rod is subject to some varying constraints (such as one fixed end being applied a constant rotation). To the best of our knowledge, classical continuation methods used for rods [23] are not able to deal with non-persistent or sliding contact.

3.2. Discrete modeling of frictional contact

Most popular approaches in Computer Graphics and Mechanical Engineering consist in assuming that the objects in contact are locally compliant, allowing them to slightly penetrate each other. This is the principle of penalty-based methods (or molecular dynamics), which consists in adding mutual repulsive forces of the form $k\,f(\delta)$, where δ is the penetration depth detected at current time step [14], [30]. Though simple to implement and computationally efficient, the penalty-based method often fails to prevent excessive penetration of the contacting objects, which may prove fatal in the case of thin objects as those may just end up traversing each other. One solution might be to set the stiffness factor k to a large enough value, however this causes the introduction of parasitical high frequencies and calls for very small integration steps [6]. Penalty-based approaches are thus generally not satisfying for ensuring robust contact handling.

In the same vein, the friction law between solid objects, or within a yield-stress fluid (used to model foam, sand, or cement, which, unlike water, cannot flow beyond a certain threshold), is commonly modeled using a regularized friction law (sometimes even with simple viscous forces), for the sake of simplicity and numerical tractability (see e.g., [32], [25]). Such a model cannot capture the threshold effect that characterizes friction between contacting solids or within a yield-stress fluid. The nonsmooth transition between sticking and sliding is however responsible for significant visual features, such as the complex patterns resting on the outer surface of hair, the stable formation of sand piles, or typical stick-slip instabilities occurring during motion.

The search for a realistic, robust and stable frictional contact method encouraged us to depart from those, and instead to focus on rigid contact models coupled to the exact nonsmooth Coulomb law for friction (and respectively, to the exact nonsmooth Drucker-Prager law in the case of a fluid), which better integrate the effects of frictional contact at the macroscopic scale. This motivation was the sense of the hiring of F. Bertails-Descoubes in 2007 in the Inria/LJK BIPOP team, specialized in nonsmooth mechanics and related convex optimization methods. In the line of F. Bertails-Descoubes's work performed in the BIPOP team, the ELAN team keeps on including some active research on the finding of robust frictional contact algorithms specialized for slender deformable structures.

3.2.1. Optimized algorithms for large nodal systems in frictional contact

In the fiber assembly case, the resulting mass matrix M is block-diagonal, so that the Delassus operator can be computed in an efficient way by leveraging sparse-block computations [15]. This justifies solving the reduced discrete frictional contact problem where primary unknowns are forces, as usually advocated in nonsmooth mechanics [28]. For cloth however, where primal variables (nodal velocities of the cloth mesh) are all interconnected via elasticity through implicit forces, the method developed above is computationally inefficient. Indeed, the matrix M (only block-sparse, but not block-diagonal) is costly to invert for large systems and its inverse is dense. Recently, we have leveraged the fact that generalized velocities of the system are 3D velocities, which simplifies the discrete contact problem when contacts occur at the nodes. Combined with a multiresolution strategy, we have devised an algorithm able to capture exact Coulomb friction constraints at contact, while retaining computational efficiency [29]. This work also supports cloth self-contact and cloth multilayering. How to enrich the interaction model with, e.g., cohesion, remains an open question. The experimental validation of our frictional contact model is also one of our goals in the medium run.

[0] Within this optimal control framework, our previous curvature-based methods can actually be interpreted as a special case of direct single shooting methods.

3.2.2. Continuum modeling of large fiber assemblies

Though we have recently made progress on the continuum formulation and solving of granular materials in Gilles Daviet's PhD thesis [19], [17], [16], we are still far from a continuum description of a macroscopic dry fibrous medium such as hair. One key ingredient that we have not been considering in our previous models is the influence of air inside divided materials. Typically, air plays a considerable role in hair motion. To advance in that direction, we have started to look at a diphasic fluid representation of granular matter, where a Newtonian fluid and the solid phase are fully coupled, while the nonsmooth Drucker-Prager rheology for the solid phase is enforced implicitly [18]. This first approach could be a starting point for modeling immersed granulars in a liquid, or ash clouds, for instance.

A long path then remains to be achieved, if one wants to take into account long fibers instead of isotropic grains in the solid phase. How to couple the fiber elasticity with our current formulation remains a challenging problem.

3.3. Inverse design of slender elastic structures [ERC Gem]

With the considerable advance of automatic image-based capture in Computer Vision and Computer Graphics these latest years, it becomes now affordable to acquire quickly and precisely the full 3D geometry of many mechanical objects featuring intricate shapes. Yet, while more and more geometrical data get collected and shared among the communities, there is currently very little study about how to infer the underlying mechanical properties of the captured objects merely from their geometrical configurations.

An important challenge consists in developing a non-invasive method for inferring the mechanical properties of complex objects from a minimal set of geometrical poses, in order to predict their dynamics. In contrast to classical inverse reconstruction methods, our claim is that 1/ the mere geometrical shape of physical objects reveals a lot about their underlying mechanical properties and 2/ this property can be fully leveraged for a wide range of objects featuring rich geometrical configurations, such as slender structures subject to contact and friction (e.g., folded cloth or twined filaments).

In addition to significant advances in fast image-based measurement of diverse mechanical materials stemming from physics, biology, or manufacturing, this research is expected in the long run to ease considerably the design of physically realistic virtual worlds, as well as to boost the creation of dynamic human doubles.

To achieve this goal, we shall develop an original inverse modeling strategy based upon the following research topics:

3.3.1. Design of well-suited discrete models for slender structures

We believe that the quality of the upstream, reference physics-based model is essential to the effective connection between geometry and mechanics. Typically, such a model should properly account for the nonlinearities due to large displacements of the structures, as well as to the nonsmooth effects typical of contact and friction.

It should also be parameterized and discretized in such a way that inversion gets simplified mathematically, possibly avoiding the huge cost of large and nonconvex optimization. In that sense, unlike concurrent methods which impose inverse methods to be compatible with a generic physics-based model, we instead advocate the design of specific physics-based models which are tailored for the inversion process.

More precisely, from our experience on fiber modeling, we believe that reduced Lagrangian models, based on a minimal set of coordinates and physical parameters (as opposed to maximal coordinates models such as mass-springs), are particularly well-suited for inversion and physical interpretation of geometrical data [21], [20]. Furthermore, choosing a high-order coordinate system (e.g., curvatures instead of angles) allows for a precise handling of curved boundaries and contact geometry, as well as the simplification of constitutive laws (which are transformed into a linear equation in the case of rods). We are currently investigating high-order discretization schemes for elastic ribbons and developable shells [10].

3.3.2. Static inversion of physical objects from geometrical poses

We believe that pure static inversion may by itself reveal many insights regarding a range of parameters such as the undeformed configuration of the object, some material parameters or contact forces.

The typical settings that we consider is composed of, on the one hand, a reference mechanical model of the object of interest, and on the other hand a single or a series of complete geometrical poses corresponding each to a static equilibrium. The core challenge consists in analyzing theoretically and practically the amount of information that can be gained from one or several geometrical poses, and to understand how the fundamental under-determinacy of the inverse problem can be reduced, for each unknown quantity (parameter or force) at play. Both the equilibrium condition and the stability criterion of the equilibrium are leveraged towards this goal. On the theoretical side, we have recently shown that a given 3D curve always matches the centerline of an isotropic suspended Kirchhoff rod at equilibrium under gravity, and that the natural configuration of the rod is unique once material parameters (mass, Young modulus) are fixed [1]. On the practical side, we have recently devised a robust algorithm to find a valid natural configuration for a discrete shell to match a given surface under gravity and frictional contact forces [3]. Unlike rods however, shells can have multiple inverse (natural) configurations. Choosing among the multiple solutions based on some selection criteria is an open challenge. Another open issue, in all cases, is the theoretical characterization of material parameters allowing the equilibrium to be stable.

3.3.3. Dynamic inversion of physical objects from geometrical poses

To refine the solution subspaces searched for in the static case and estimate dynamic parameters (e.g., some damping coefficients), a dynamic inversion process accounting for the motion of the object of interest is necessary.

In contrast to the static case where we can afford to rely on exact geometrical poses, our analysis in the dynamic case will have to take into account the imperfect quality of input data with possible missing parts or outliers. One interesting challenge will be to combine our high-order discretized physics-based model together with the acquisition process in order to refine both the parameter estimation and the geometrical acquisition.

3.3.4. Experimental validation with respect to real data

The goal will be to confront the theories developed above to real experiments. Compared to the statics, the dynamic case will be particularly involving as it will be highly dependent on the quality of input data as well as the accuracy of the motion predicted by our physics-based simulators. Such experiments will not only serve to refine our direct and inverse models, but will also be leveraged to improve the 3D geometrical acquisition of moving objects. Besides, once validation will be performed, we shall work on the setting up of new non-invasive measurement protocols to acquire physical parameters of slender structures from a minimal amount of geometrical configurations.

4. Application Domains

4.1. Mechanical Engineering

Many physicists and mathematicians have strived for centuries to understand the principles governing those complex mechanical phenomena, providing a number of continuous models for slender structures, granular matter, and frictional contact. In the XX^{th} century, industrial applications such as process automatization and new ways of transportation have boosted the fields of Mechanical Engineering and Computer-Aided Design, where material strength, reliability of mechanisms, and safety, standed for the main priorities. Instead, large displacements of structures, buckling, tearing, or entanglement, and even dynamics, were long considered as undesirable behaviors, thus restraining the search for corresponding numerical models.

Only recently, the engineering industry has shown some new and growing interest into the modeling of dynamic phenomena prone to large displacements, contact and friction. For instance, the cosmetology industry is more and more interested in understanding the nonlinear deformation of hair and skin, with the help of simulation. Likewise, auto and aircraft manufacturers are facing new challenges involving buckling or entanglement of thin structures such as carbon or optical fibers; they clearly lack predictive, robust and efficient numerical tools for simulating and optimizing their new manufacturing process, which share many common features with the large-scale simulation scenarii traditionally studied in Computer Graphics applications.

4.2. Computer Graphics

In contrast, Computer Graphics, which has emerged in the 60's with the advent of modern computers, was from the very beginning eager to capture such peculiar phenomena, with the sole aim to produce spectacular images and create astonishing stories. At the origin, Computer Graphics thus drastically departed from other scientific fields. Everyday-life phenomena such as cloth buckling, paper tearing, or hair fluttering in the wind, mostly ignored by other scientists at that time, became actual topics of interest, involving a large set of new research directions to be explored, both in terms of modelling and simulation. Nowadays, although the image production still remains the core activity of the Computer Graphics community, more and more research studies are directed through the virtual and real prototyping of mechanical systems, notably driven by a myriad of new applications in the virtual try on industry (e.g., hairstyling and garment fitting). Furthermore, the advent of additive fabrication is currently boosting research in the free design of new mechanisms or systems for various applications, from architecture design and fabrication of metamaterials to the creation of new locomotion modes in robotics. Some obvious common interests and approaches are thus emerging between Computer Graphics and Mechanical Engineering, yet the two communities remain desperately compartmentalized.

4.3. Soft Matter Physics

From the physics-based viewpoint, since a few decades a new generation of physicists became interested again in the understanding of such visually fascinating phenomena, and started investigating the tight links between geometry and elasticity [0]. Common objects such as folded or torn paper, twined plants, coiled honey threads, or human hair have thus regained some popularity among the community in Nonlinear Physics [0]. In consequence, phenomena of interest have become remarkably close to those of Computer Graphics, since scientists in both places share the common goal to model complex and integrated mechanical phenomena at the macroscopic scale. Of course, the goals and employed methodologies differ substantially from one community to the other, but showcase some evident complementarity: while computer scientists are eager to learn and understand new physical models, physicists get more and more interested in the numerical tools, in which they perceive not only a means to confirm predictions afterwards, but also a support for testing new hypothesis and exploring scenarios that would be too cumbersome or even impossible to investigate experimentally. Besides, numerical exploration starts becoming a valuable tool for getting insights into the search for analytic solutions, thus fully participating to the modeling stage and physical understanding. However, physicists may be limited to a blind usage of numerical black boxes, which may furthermore not be dedicated to their specific needs. According to us, promoting a science of modeling in numerical physics would thus be a promising and rich avenue for the two research fields. Unfortunately, very scarce cooperation currently exists between the two communities, and large networks of collaboration still need to be set up.

5. New Software and Platforms

5.1. Argus-distribution

KEYWORDS: Frictional contact - Cloth dynamics - Mesh adaptation

[0] In France this new trend was particularly stimulated by the work of Yves Pomeau, who convinced many young scientists to study the nonlinear physics of common objects such as paper, plants, or hair [24].

[0] It is however amusing to observe that research in these areas is quite successful in obtaining the IG Nobel prize [5], [27], thus still being considered as an exotic research topic by physicists.

SCIENTIFIC DESCRIPTION: The Argus-distribution software exactly replicates all the results published in the SIGGRAPH 2018 paper entitled "An Implicit Frictional Contact Solver for Adaptive Cloth Simulation", by Li et al. This paper presents the first method able to account for cloth contact with exact Coulomb friction, treating both cloth self-contacts and contacts occurring between the cloth and an underlying character. The key contribution is to observe that for a nodal system like cloth, the frictional contact problem may be formulated based on velocities as primary variables, without having to compute the costly Delassus operator. Then, by reversing the roles classically played by the velocities and the contact impulses, conical complementarity solvers of the literature can be adapted to solve for compatible velocities at nodes. To handle the full complexity of cloth dynamics scenarios, this base algorithm has been extended in two ways: first, towards the accurate treatment of frictional contact at any location of the cloth, through an adaptive node refinement strategy, second, towards the handling of multiple constraints at each node, through the duplication of constrained nodes and the adding of pin constraints between duplicata. This method allows to handle the complex cloth-cloth and cloth-body interactions in full-size garments with an unprecedented level of realism compared to former methods, while maintaining reasonable computational timings.allows to simulate cloth dynamics subject to frictional contact.

FUNCTIONAL DESCRIPTION: Adaptive cloth simulation in the presence of frictional contact. Reference software for the paper "An Implicit Frictional Contact Solver for Adaptive Cloth Simulation", Li et al. 2018, ACM Transactions on Graphics (SIGGRAPH'18).

- Participants: Jie Li, Gilles Daviet, Rahul Narain, Florence Descoubes, Matthew Overby, George Brown and Laurence Boissieux
- Partners: Department of Computer Science and Engineering, University of Minnesota - IIT Delhi
- Contact: Florence Descoubes
- Publication: An Implicit Frictional Contact Solver for Adaptive Cloth Simulation
- URL: http://www-users.cselabs.umn.edu/~lixx4611/contact_friction.html

6. New Results

6.1. Inverse design of a suspended elastic rod

Participants: Florence Bertails-Descoubes, Victor Romero.

In collaboration with Alexandre Derouet-Jourdan (OLM Digital, Japan) and Arnaud Lazarus (UPMC, Laboratoire Jean le Rond d'Alembert), we have investigated the inverse design problem of a suspended elastic subject to gravity. We have proved that given an arbitrary space curve, there exists a unique solution for the natural configuration of the rod, which is independent of the initial framing of the input curve. Moreover, this natural configuration can be easily computed by solving three linear ODEs in sequence, starting from any input framing. This work has been published in Roy. Soc. Proc A [1] and physical aspects of this study have been communicated about in a mechanical congress [4].

6.2. Simulation of cloth contact with exact Coulomb friction

Participants: Florence Bertails-Descoubes, Laurence Boissieux.

In collaboration with Gilles Daviet (Weta Digital, New Zealand) and Rahul Narain's group (University of Minnesota and IIT Delhi), we have developed a new implicit solver for taking into account contact in cloth with Coulomb friction. Our key idea stems from the observation that for a nodal system like cloth, and in the case where each node is subject to at most one contacting constraint (either an external or self-contact), the frictional contact problem may be formulated based on velocities as primary variables, without having to compute the costly Delassus operator; then, by reversing the roles classically played by the velocities and the contact impulses, conical complementarity solvers of the literature may be leveraged to solve for compatible velocities at nodes. To handle the full complexity of cloth dynamics scenarios, we have extended this base algorithm in two ways: first, towards the accurate treatment of frictional contact at any location of the cloth, through an adaptive node refinement strategy; second, towards the handling of multiple constraints at each node, through the duplication of constrained nodes and the adding of pin constraints between duplicata. Our method proves to be both fast and robust, allowing us to simulate full-size garments with an unprecedent level of realism compared to former methods, while maintaining similar computational timings. Our work has been published at ACM Transactions on Graphics (ACM SIGGRAPH 2018) [2].

6.3. Inverse design of thin elastic shells

Participants: Mickaël Ly, Florence Bertails-Descoubes, Laurence Boissieux.

In collaboration with Romain Casati (former PhD student of F. Bertails-Descoubes) and Mélina Skouras (EPI IMAGINE), we have proposed an inverse strategy for modeling thin elastic shells physically, just from the observation of their geometry. Our algorithm takes as input an arbitrary target mesh, and interprets this configuration automatically as a stable equilibrium of a shell simulator under gravity and frictional contact constraints with a given external object. Unknowns are the natural shape of the shell (i.e., its shape without external forces) and the frictional contact forces at play, while the material properties (mass density, stiffness, friction coefficients) can be freely chosen by the user. Such an inverse problem formulates as an ill-posed nonlinear system subject to conical constraints. To select and compute a plausible solution, our inverse solver proceeds in two steps. In a first step, contacts are reduced to frictionless bilateral constraints and a natural shape is retrieved using the adjoint method. The second step uses this result as an initial guess and adjusts each bilateral force so that it projects onto the admissible Coulomb friction cone, while preserving global equilibrium. To better guide minimization towards the target, these two steps are applied iteratively using a degressive regularization of the shell energy. We validate our approach on simulated examples with reference material parameters, and show that our method still converges well for material parameters lying within a reasonable range around the reference, and even in the case of arbitrary meshes that are not issued from a simulation. We finally demonstrate practical inversion results on complex shell geometries freely modeled by an artist or automatically captured from real objects, such as posed garments or soft accessories. Our work has been published at ACM Transactions on Graphics (ACM SIGGRAPH Asia 2018) [3] and has been selected for a Press Release of the ACM.

7. Partnerships and Cooperations

7.1. European Initiatives

7.1.1. FP7 & H2020 Projects

7.1.1.1. GEM

Title: from GEometry to Motion, inverse modeling of complex mechanical structures

Programm: H2020

Type: ERC

Duration: September 2015 - August 2021

Coordinator: Inria

Inria contact: Florence BERTAILS-DESCOUBES

With the considerable advance of automatic image-based capture in Computer Vision and Computer Graphics these latest years, it becomes now affordable to acquire quickly and precisely the full 3D geometry of many mechanical objects featuring intricate shapes. Yet, while more and more geometrical data get collected and shared among the communities, there is currently very little study about how to infer the underlying mechanical properties of the captured objects merely from their geometrical configurations. The GEM challenge consists in developing a non-invasive method for inferring the mechanical properties of complex objects from a minimal set of geometrical poses, in order to predict their dynamics. In contrast to classical inverse reconstruction methods, my proposal is built upon the claim that 1/ the mere geometrical shape of physical objects reveals a lot about their underlying mechanical properties and 2/ this property can be fully leveraged for a wide range of objects featuring rich geometrical configurations, such as slender structures subject to frictional contact (e.g., folded cloth or twined filaments). To achieve this goal, we shall develop an original inverse modeling strategy based upon a/ the design of reduced and high-order discrete models for slender mechanical structures including rods, plates and shells, b/ a compact and well-posed mathematical formulation of our nonsmooth inverse problems, both in the static and dynamic cases, c/ the design of robust and efficient numerical tools for solving such complex problems, and d/ a thorough experimental validation of our methods relying on the most recent capturing tools. In addition to significant advances in fast image-based measurement of diverse mechanical materials stemming from physics, biology, or manufacturing, this research is expected in the long run to ease considerably the design of physically realistic virtual worlds, as well as to boost the creation of dynamic human doubles.

7.2. International Initiatives

7.2.1. Inria International Partners

7.2.1.1. Declared Inria International Partners

- Long-term partnership with Rahul Narain (University of Minnesota, USA, and IIT Delhi, INDIA) and Rahul Narain's PhD student Jie Li (University of Minnesota, USA).
- Long-term partnership with Alexandre-Derouet-Jourdan (OLM Digital, JAPAN).

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Member of the Conference Program Committees

- Florence Bertails-Descoubes, member of the ACM SIGGRAPH Technical Program Committee in 2018 and 2019, and the Eurographics Technical Program Committee in 2018 and 2019.

8.1.2. Journal

8.1.2.1. Reviewer - Reviewing Activities

- Florence Bertails-Descoubes, Reviewer in 2018 for ACM Transaction on Graphics, ACM SIGGRAPH Asia 2018, Soft Matter, Royal Society Open Science.

8.1.3. Invited Talks

- Florence Bertails-Descoubes, invited talk at IUSTI, Marseille, June 2018 (contact: O. Pouliquen, équipe Écoulements de Particules).

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Licence : Raphaël Charrondière, Calcul matriciel et fonctions de plusieurs variables, 36h éq TD, L2, Université Grenoble Alpes

Licence : Florence Bertails-Descoubes, Méthodes Numériques, 18h éq TD, L3, ENSIMAG 1A, Grenoble INP.

8.2.2. Supervision

PhD in progress : Mickaël Ly, Static inverse modelling of cloth, 01 octobre 2017, Florence Bertails-Descoubes and Mélina Skouras.

PhD in progress : Haroon Rasheed, Inverse dynamic modeling of cloth, 01 novembre 2017, Florence Bertails-Descoubes, Jean-Sébastien Franco, and Stefanie Wuhrer

PhD in progress : Raphaël Charrondière, Modeling and numerical simulation of elastic inextensible surfaces, 01 septembre 2018, Florence Bertails-Descoubes and Sébastien Neukirch.

8.2.3. Juries

Florence Bertails-Descoubes, member (Rapportrice) of Ph.D. Thesis committee of E. Cottenceau (10 avril 2018), ENSAM Lille (directeur de thèse : O. Thomas)

Florence Bertails-Descoubes, member (Rapportrice) of Ph.D. Thesis committee of S. Salamone (4 juillet 2018), Institut Charles Sadron à l'Université de Strasbourg (directeur de thèse : T. Charitat)

Florence Bertails-Descoubes, member (Rapportrice) of Ph.D. Thesis committee of S. Poincloux (15 octobre 2018), Laboratoire de Physique Statistique de l'ENS Paris (directeurs de thèse : F. Léchenault et M. Adda-Bedia).

8.3. Popularization

8.3.1. Articles and contents

- Interview about our past work on hair and cloth for a press article in Dauphiné Libéré des Enfants in November 2018.
- Press interview about our SIGGRAPH Asia paper on inverse shell design for an ACM Press Release in November 2018.

9. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] F. Bertails-Descoubes, A. Derouet-Jourdan, V. Romero, A. Lazarus. *Inverse design of an isotropic suspended Kirchhoff rod: theoretical and numerical results on the uniqueness of the natural shape*, in "Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences", April 2018, vol. 474, n° 2212, p. 1-26 [*DOI :* 10.1098/rspa.2017.0837], https://hal.inria.fr/hal-01827887

[2] J. Li, G. Daviet, R. Narain, F. Bertails-Descoubes, M. Overby, G. Brown, L. Boissieux. *An Implicit Frictional Contact Solver for Adaptive Cloth Simulation*, in "ACM Transactions on Graphics", August 2018, vol. 37, n° 4, p. 1-15 [*DOI :* 10.1145/3197517.3201308], https://hal.inria.fr/hal-01834705

[3] M. Ly, R. Casati, F. Bertails-Descoubes, M. Skouras, L. Boissieux. *Inverse Elastic Shell Design with Contact and Friction*, in "ACM Transactions on Graphics", November 2018, vol. 37, n° 6, p. 1-16 [*DOI :* 10.1145/3272127.3275036], https://hal.inria.fr/hal-01883655

International Conferences with Proceedings

[4] V. Romero, F. Bertails-Descoubes, A. Derouet-Jourdan, A. Lazarus. *Inverse design of a suspended Kirchhoff rod: From theory to practice*, in "ESMC10 2018 - 10th European Solid Mechanics Conference", Bologne, Italy, July 2018, vol. 25, 184302, https://hal.inria.fr/hal-01778474

References in notes

[5] B. AUDOLY, S. NEUKIRCH.*Fragmentation of Rods by Cascading Cracks: Why Spaghetti Does Not Break in Half*, in "Physical Review Letters", 2005, vol. 95, n^o 9, 095505

[6] D. BARAFF.*Analytical methods for dynamic simulation of non-penetrating rigid bodies*, in "Computer Graphics Proceedings (Proc. ACM SIGGRAPH'89)", New York, NY, USA, ACM, 1989, p. 223–232

[7] F. BERTAILS, B. AUDOLY, M.-P. CANI, B. QUERLEUX, F. LEROY, J.-L. LÉVÊQUE.*Super-Helices for Predicting the Dynamics of Natural Hair*, in "ACM Transactions on Graphics (Proc. ACM SIGGRAPH'06)", 2006, vol. 25, p. 1180–1187 [*DOI :* 10.1145/1141911.1142012], http://www-evasion.imag.fr/Publications/2006/BACQLL06

[8] F. BERTAILS.*Linear Time Super-Helices*, in "Computer Graphics Forum (Proc. Eurographics'09)", apr 2009, vol. 28, n^o 2, http://www-ljk.imag.fr/Publications/Basilic/com.lmc.publi.PUBLI_Article@1203901df78_1d3cdaa/

[9] F. BERTAILS-DESCOUBES.*Super-Clothoids*, in "Computer Graphics Forum (Proc. Eurographics'12)", 2012, vol. 31, n^o 2pt2, p. 509–518 [*DOI :* 10.1111/J.1467-8659.2012.03030.X], http://www.inrialpes.fr/bipop/people/bertails/Papiers/superClothoids.html

[10] A. BLUMENTALS, F. BERTAILS-DESCOUBES, R. CASATI.*Dynamics of a developable shell with uniform curvatures*, in "The 4th Joint International Conference on Multibody System Dynamics", Montréal, Canada, May 2016, https://hal.inria.fr/hal-01311559

[11] A. BLUMENTALS.*Numerical modelling of thin elastic solids in contact*, Université de Grenoble Alpes, July 2017

[12] R. CASATI, F. BERTAILS-DESCOUBES.*Super Space Clothoids*, in "ACM Transactions on Graphics (Proc. ACM SIGGRAPH'13)", July 2013, vol. 32, n^o 4, p. 48:1–48:12, http://doi.acm.org/10.1145/2461912.2461962

[13] D. CHAPELLE, K. BATHE.*The Finite Element Analysis of Shells - Fundamentals - Second Edition*, Computational Fluid and Solid Mechanics, Springer, 2011, 410 [*DOI :* 10.1007/978-3-642-16408-8]

[14] P. CUNDALL.*A computer model for simulating progressive large scale movements of blocky rock systems. In Proceedings of the Symposium of the International Society of Rock Mechanics*, in "Proceedings of the Symposium of the International Society of Rock Mechanics", 1971, vol. 1, p. 132–150

[15] G. DAVIET, F. BERTAILS-DESCOUBES, L. BOISSIEUX.*A hybrid iterative solver for robustly capturing Coulomb friction in hair dynamics*, in "ACM Transactions on Graphics (Proc. ACM SIGGRAPH Asia'11)", 2011, vol. 30, p. 139:1–139:12, http://www.inrialpes.fr/bipop/people/bertails/Papiers/hybridIterativeSolverHairDynamicsSiggraphAsia2011.html

[16] G. DAVIET, F. BERTAILS-DESCOUBES.*A semi-implicit material point method for the continuum simulation of granular materials*, in "ACM Transactions on Graphics", July 2016, vol. 35, n^o 4, 13 [*DOI :* 10.1145/2897824.2925877], https://hal.inria.fr/hal-01310189

[17] G. DAVIET, F. BERTAILS-DESCOUBES.*Nonsmooth simulation of dense granular flows with pressure-dependent yield stress*, in "Journal of Non-Newtonian Fluid Mechanics", April 2016, vol. 234, p. 15-35 [*DOI :* 10.1016/J.JNNFM.2016.04.006], https://hal.inria.fr/hal-01236488

[18] G. DAVIET, F. BERTAILS-DESCOUBES.*Simulation of Drucker–Prager granular flows inside Newtonian fluids*, February 2017, working paper or preprint, https://hal.inria.fr/hal-01458951

[19] G. DAVIET.*Modèles et algorithmes pour la simulation du contact frottant dans les matériaux complexes : application aux milieux fibreux et granulaires*, Grenoble Alpes Universités, December 2016

[20] A. DEROUET-JOURDAN, F. BERTAILS-DESCOUBES, G. DAVIET, J. THOLLOT.*Inverse Dynamic Hair Modeling with Frictional Contact*, in "ACM Trans. Graph.", November 2013, vol. 32, n^o 6, p. 159:1–159:10, http://doi.acm.org/10.1145/2508363.2508398

[21] A. DEROUET-JOURDAN, F. BERTAILS-DESCOUBES, J. THOLLOT.*Stable Inverse Dynamic Curves*, in "ACM Transactions on Graphics (Proc. ACM SIGGRAPH Asia'10)", December 2010, vol. 29, p. 137:1–137:10, http://doi.acm.org/10.1145/1882261.1866159

[22] M. A. DIAS, B. AUDOLY.*"Wunderlich, Meet Kirchhoff": A General and Unified Description of Elastic Ribbons and Thin Rods*, in "Journal of Elasticity", Apr 2015, vol. 119, n^o 1, p. 49–66, https://doi.org/10.1007/s10659-014-9487-0

[23] E. J. DOEDEL, R. C. PFAFFENROTH, A. R. CHAMBODUT, T. F. FAIRGRIEVE, Y. A. KUZNETSOV, B. E. OLDEMAN, B. SANDSTEDE, X. WANG.*AUTO 2000: Continuation and Bifurcation Software for Ordinary Differential Equations (with HomCont)*, March 2006

[24] ESPCI.*Rencontre en l'honneur de Yves Pomeau, octobre 2016*, 2016, https://www.sfpnet.fr/rencontre-celebrant-la-medaille-boltzmann-d-yves-pomeau, https://www.sfpnet.fr/rencontre-celebrant-la-medaille-boltzmann-d-yves-pomeau

[25] I. FRIGAARD, C. NOUAR.*On the usage of viscosity regularisation methods for visco-plastic fluid flow computation*, in "Journal of Non-Newtonian Fluid Mechanics", April 2005, vol. 127, n^o 1, p. 1–26 [*DOI :* 10.1016/J.JNNFM.2005.01.003]

[26] A. L. GOL'DENVEIZER.*Theory of Elastic Thin Shells*, Pergamon Press, 1961

[27] R. E. GOLDSTEIN, P. B. WARREN, R. C. BALL.*Shape of a Ponytail and the Statistical Physics of Hair Fiber Bundles*, in "Phys. Rev. Lett.", Feb 2012, vol. 108, 078101, https://link.aps.org/doi/10.1103/PhysRevLett.108.078101

[28] M. JEAN.*The Non Smooth Contact Dynamics Method*, in "Computer Methods in Applied Mechanics and Engineering", 1999, vol. 177, p. 235-257, Special issue on computational modeling of contact and friction, J.A.C. Martins and A. Klarbring, editors

[29] J. LI, G. DAVIET, R. NARAIN, F. BERTAILS-DESCOUBES, M. OVERBY, G. BROWN, L. BOISSIEUX.*An Implicit Frictional Contact Solver for Adaptive Cloth Simulation*, in "ACM Trans. Graph.", August 2018, vol. 37, n^o 4

[30] M. MOORE, J. WILHELMS.*Collision detection and response for computer animation*, in "Computer Graphics Proceedings (Proc. ACM SIGGRAPH'88)", 1988, p. 289–298

[31] D. MOULTON, P. GRANDGEORGE, S. NEUKIRCH.*Stable elastic knots with no self-contact*, in "Journal of the Mechanics and Physics of Solids", 2018, vol. 116, p. 33–53 [*DOI :* 10.1016/J.JMPS.2018.03.019], http://www.sciencedirect.com/science/article/pii/S0022509617310104

[32] J. SPILLMANN, M. TESCHNER.*An Adaptive Contact Model for the Robust Simulation of Knots*, in "Computer Graphics Forum", 2008, vol. 27, n^{o} 2, Proc. Eurographics'08

[33] H. SUGIYAMA, J. GERTSMAYR, A. MIKKOLA.*Flexible Multibody Dynamics — Essential for Accurate Modeling in Multibody System Dynamics*, in "Journal of Computational Nonlinear Dynamics", November 2013, vol. 9

Project-Team ERABLE

European Research team in Algorithms and Biology, formaL and Experimental

IN COLLABORATION WITH: Laboratoire de Biométrie et Biologie Evolutive (LBBE)

IN PARTNERSHIP WITH:
Centrum Wiskunde & Informatica

Institut national des sciences appliquées de Lyon

Université Claude Bernard (Lyon 1)

Université de Rome la Sapienza

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Computational Biology

Table of contents

Project-Team ERABLE

Creation of the Team: 2015 January 01, updated into Project-Team: 2015 July 01

ERABLE is a European Inria team gathering French researchers together with researchers in Italy under the banner of the Sapienza University of Rome and researchers in the Netherlands under the banner of the CWI.

Keywords:

Computer Science and Digital Science:

A3. - Data and knowledge
A3.1. - Data
A3.1.1. - Modeling, representation
A3.1.4. - Uncertain data
A3.3. - Data and knowledge analysis
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A7. - Theory of computation
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.7. - Graph theory
A8.8. - Network science
A8.9. - Performance evaluation

Other Research Topics and Application Domains:

B1. - Life sciences
B1.1. - Biology
B1.1.1. - Structural biology
B1.1.2. - Molecular and cellular biology
B1.1.4. - Genetics and genomics
B1.1.6. - Evolutionnary biology
B1.1.7. - Bioinformatics
B1.1.10. - Systems and synthetic biology
B2. - Health
B2.2. - Physiology and diseases
B2.2.3. - Cancer
B2.2.4. - Infectious diseases, Virology
B2.3. - Epidemiology

1. Team, Visitors, External Collaborators

Research Scientists

Marie-France Sagot [Inria, Team leader, Senior Researcher, HDR]
Blerina Sinaimeri [Inria, Researcher]
Fabrice Vavre [CNRS, Researcher, HDR]
Alain Viari [Inria, Senior Researcher]

Alexander Schönhuth [CWI, The Netherlands, Senior Researcher, also since Oct 2017 part-time Professor at Univ of Utrecht]

Faculty Members

Hubert Charles [INSA Lyon, Associate Professor, HDR]
Roberto Grossi [Univ Pisa, Italy, Professor]
Pierluigi Crescenzi [Univ Florence, Florence, Italy, Professor]
Giuseppe Francesco Italiano [Univ Tor Vergata, Rome, Italy, Professor]
Vincent Lacroix [Univ de Claude Bernard, Associate Professor]
Alberto Marchetti Spaccamela [Sapienza Univ Rome, Italy, Professor]
Arnaud Mary [Univ de Claude Bernard, Associate Professor]
Nadia Pisanti [Univ Pisa, Italy, Associate Professor]
Andrea Marino [Univ Pisa, Italy, Assistant Professor]
Leen Stougie [CWI & Free Univ Amsterdam, The Netherlands, Professor]
Cristina Vieira [Univ de Claude Bernard, Professor, HDR]

External Collaborators

Laurent Jacob [LBBE UMR5558, Researcher, external collaborator]
Susana Vinga [IST Lisbon, Researcher, external collaborator]

Technical Staff

Clara Benoît-Pilven [INSERM, until Sep 2018]
Martin Wannagat [Inria, until Feb 2018]

PhD Students

Marianne Borderes [MaatPharma and Claude Bernard Lyon, since Feb 2018]
Audric Cologne [Inria]
Mattia Gastaldello [Sapienza Univ of Rome and Univ Claude Bernard Lyon, until Feb 2018]
Leandro Ishi Soares de Lima [CNPq Brazil & Univ Claude Bernard Lyon]
Carol Moraga Quinteros [Conicyt & Univ Claude Bernard Lyon]
Henri Pusa [Inria]
Camille Sessegolo [Univ Claude Bernard Lyon]
Yishu Wang [Univ Claude Bernard Lyon]
Irene Ziska [Inria]

Post-Doctoral Fellows

Ricardo de Andrade Abrantes [Univ São Paulo, Brazil, and Univ Claude Bernard Lyon]
Alex Di Genova [Inria]
Mariana Galvão Ferrarini [Insa Lyon & Univ Claude Bernard Lyon, Mar 2018]
Clara Benoît-Pilven [Univ Claude Bernard Lyon, from Oct 2018]

Administrative Assistant

Claire Sauer [Inria]

2. Overall Objectives

2.1. Overall Objectives

Cells are seen as the basic structural, functional and biological units of all living systems. They represent the smallest units of life that can replicate independently, and are often referred to as the building blocks of life. Living organisms are then classified into unicellular ones – this is the case of most bacteria and archea – or multicellular – this is the case of animals and plants. Actually, multicellular organisms, such as for instance human, may be seen as composed of native (human) cells, but also of extraneous cells represented by the diverse bacteria living inside the organism. The proportion in the number of the latter in relation to the number of native cells is believed to be high: this is for example of 90% in humans. Multicellular organisms have thus been described also as "superorganisms with an internal ecosystem of diverse symbiotic microbiota and parasites" (Nicholson *et al.*, Nat Biotechnol, 22(10):1268-1274, 2004)) where symbiotic means that the extraneous unicellular organisms (cells) live a close, and in this case, long-term relation both with the multicellular organisms they inhabit and among themselves. On the other hand, bacteria sometimes group into colonies of genetically identical individuals which may acquire both the ability to adhere together and to become specialised for different tasks. An example of this is the cyanobacterium *Anabaena sphaerica* who may group to form filaments of differentiated cells, some – the heterocysts – specialised for nitrogen fixation while the others are capable of photosynthesis. Such filaments have been seen as first examples of multicellular patterning.

At its extreme, one could then see life as one collection, or a collection of collections of genetically identical or distinct self-replicating cells who interact, sometimes closely and for long periods of evolutionary time, with same or distinct functional objectives. The interaction may be at equilibrium, meaning that it is beneficial or neutral to all, or it may be unstable meaning that the interaction may be or become at some time beneficial only to some and detrimental to other cells or collections of cells. The interaction may involve other living systems, or systems that have been described as being at the edge of life such as viruses, or else genetic or inorganic material such as, respectively, transposable elements and chemical compounds.

The application goal of ERABLE is, through the use of mathematical models and algorithms, to better understand such close and often persistent interactions, with a longer term objective of becoming able in some cases to suggest the means of controlling for or of re-establishing equilibrium in an interacting community by acting on its environment or on its players, how they play and who plays. This goal requires to identify who are the partners in a closely interacting community, who is interacting with whom, how and by which means. Any model is a simplification of reality, but once selected, the algorithms to explore such model should address questions that are precisely defined and, whenever possible, be exact in the answer as well as exhaustive when more than one exists in order to guarantee an accurate interpretation of the results within the given model. This fits well the mathematical and computational expertise of the team, and drives the methodological goal of ERABLE which is to substantially and systematically contribute to the field of exact enumeration algorithms for problems that most often will be hard in terms of their complexity, and as such to also contribute to the field of combinatorics in as much as this may help in enlarging the scope of application of exact methods.

The key objective is, by constantly crossing ideas from different models and types of approaches, to look for and to infer "patterns", as simple and general as possible, either at the level of the biological application or in terms of methodology. This objective drives which biological systems are considered, and also which models and in which order, going from simple discrete ones first on to more complex continuous models later if necessary and possible.

3. Research Program

3.1. Two main goals

ERABLE has two main goals, one related to biology and the other to methodology (algorithms, combinatorics, statistics). In relation to biology, the main goal of ERABLE is to contribute, through the use of mathematical models and algorithms, to a better understanding of close and often persistent interactions between "collections of genetically identical or distinct self-replicating cells" which will correspond to organisms/species or to actual cells. The first will cover the case of what has been called symbiosis, meaning when the interaction involves different species, while the second will cover the case of a (cancerous) tumour which may be seen as a collection of cells which suddenly disrupts its interaction with the other (collections of) cells in an organism by starting to grow uncontrollably.

Such interactions are being explored initially at the molecular level. Although we rely as much as possible on already available data, we intend to also continue contributing to the identification and analysis of the main genomic and systemic (regulatory, metabolic, signalling) elements involved or impacted by an interaction, and how they are impacted. We started going to the population and ecological levels by modelling and analysing the way such interactions influence, and are or can be influenced by the ecosystem of which the "collections of cells" are a part. The key steps are:

- identifying the molecular elements based on so-called omics data (genomics, transcriptomics, metabolomics, proteomics, etc.): such elements may be gene/proteins, genetic variations, (DNA/RNA/protein) binding sites, (small and long non coding) RNAs, etc.
- simultaneously inferring and analysing the network that models how these molecular elements are physically and functionally linked together for a given goal, or find themselves associated in a response to some change in the environment;
- modelling and analysing the population and ecological network formed by the "collections of cells in interaction", meaning modelling a network of networks (previously inferred or as already available in the literature).

One important longer term goal of the above is to analyse how the behaviour and dynamics of such a network of networks might be controlled by modifying it, including by subtracting some of its components from the network or by adding new ones.

In relation to methodology, the main goal is to provide those enabling to address our main biological objective as stated above that lead to the best possible interpretation of the results within a given pre-established model and a well defined question. Ideally, given such a model and question, the method is exact and also exhaustive if more than one answer is possible. Three aspects are thus involved here: establishing the model within which questions can and will be put; clearly defining such questions; exactly answering to them or providing some guarantee on the proximity of the answer given to the "correct" one. We intend to continue contributing to these three aspects:

- at the modelling level, by exploring better models that at a same time are richer in terms of the information they contain (as an example, in the case of metabolism, using hypergraphs as models for it instead of graphs) and are susceptible to an easier treatment:
  - these two objectives (rich models that are at the same time easy to treat) might in many cases be contradictory and our intention is then to contribute to a fuller characterisation of the frontiers between the two;
  - even when feasible, the richer models may lack a full formal characterisation (this is for instance the case of hypergraphs) and our intention is then to contribute to such a characterisation;
- at the question level, by providing clear formalisations of those that will be raised by our biological concerns;
- at the answer level:
  - to extend the area of application of exact algorithms by: (i) a better exploration of the combinatorial properties of the models, (ii) the development of more efficient data structures, (iii) a smarter traversal of the space of solutions when more than one solution exists;
  - when exact algorithms are not possible, or when there is uncertainty in the input data to an algorithm, to improve the quality of the results given by a deeper exploration of the links between different algorithmic approaches: combinatorial, randomised, stochastic.

3.2. Different research axes

The goals of the team are biological and methodological, the two being intrinsically linked. Any division into axes along one or the other aspect or a combination of both is thus somewhat artificial. Following the evaluation of the team at the end of 2017, four main axes were identified, with the last one being the more recently added one. This axis is specifically oriented towards health in general, human or animal. The first three axes are: genomics, metabolism and post-transcriptional regulation, and (c)evolution.

Notice that the division itself is based on the biological level (genomic, metabolic/regulatory, evolutionary) or main current Life Science purpose (health) rather than on the mathematical or computational methodology involved. Any choice has its part of arbitrariness. Through the one we made, we wished to emphasise the fact that the area of application of ERABLE is important for us. *It does not mean that the mathematical and computational objectives are not equally important*, but only that those are, most often, motivated by problems coming from or associated to the general Life Science goal. Notice that such arbitrariness also means that some Life Science topics will be artificially split into two different Axes. One example of this is genomics and the two main health areas currently addressed that are intrinsically inter-related for now.

Axis 1: Genomics

Intra and inter-cellular interactions involve molecular elements whose identification is crucial to understand what governs, and also what might enable to control such interactions. For the sake of clarity, the elements may be classified in two main classes, one corresponding to the elements that allow the interactions to happen by moving around or across the cells, and another that are the genomic regions where contact is established. Examples of the first are non coding RNAs, proteins, and mobile genetic elements such as (DNA) transposons, retro-transposons, insertion sequences, etc. Examples of the second are DNA/RNA/protein binding sites and targets. Furthermore, both types (effectors and targets) are subject to variation across individuals of a population, or even within a single (diploid) individual. Identification of these variations is yet another topic that we wish to cover. Variations are understood in the broad sense and cover single nucleotide polymorphisms (SNPs), copy-number variants (CNVs), repeats other than mobile elements, genomic rearrangements (deletions, duplications, insertions, inversions, translocations) and alternative splicings (ASs). All three classes of identification problems (effectors, targets, variations) may be put under the general umbrella of genomic functional annotation.

Axis 2: Metabolism and post-transcriptional regulation

As increasingly more data about the interaction of molecular elements (among which those described above) becomes available, these should then be modelled in a subsequent step in the form of networks. This raises two main classes of problems. The first is to accurately infer such networks. Assuming such a network, integrated or "simple", has been inferred for a given organism or set of organisms, the second problem is then to develop the appropriate mathematical models and methods to extract further biological information from such networks.

The team has so far concentrated its efforts on two main aspects concerning such interactions: metabolism and post-transcriptional regulation by small RNAs. The more special niche we have been exploring in relation to metabolism concerns the fact that the latter may be seen as an organism's immediate window into its environment. Finely understanding how species communicate through those windows, or what impact they may have on each other through them is thus important when the ultimate goal is to be able to model communities of organisms, for understanding them and possibly, on a longer term, for control. While such communication has been explored in a number of papers, most do so at a too high level or only considered couples of interacting organisms, not larger communities. The idea of investigating consortia, and in the case of synthetic biology, of using them, has thus started being developed in the last decade only, and was motivated by the fact that such consortia may perform more complicated functions than could single populations, as well as be more robust to environmental fluctuations. Another originality of the work that the team has been doing in the last decade has also been to fully explore the combinatorial aspects of the structures used (graphs or directed hypergraphs) and of the associated algorithms. As concerns post-transcriptional regulation, the team has essentially been exploring the idea that small RNAs may have an important role in the dialog between different species.

Axis 3: (Co)Evolution

Understanding how species that live in a close relationship with others may (co)evolve requires understanding for how long symbiotic relationships are maintained or how they change through time. This may have deep implications in some cases also for understanding how to control such relationships, which may be a way of controlling the impact of symbionts on the host, or the impact of the host on the symbionts and on the environment (by acting on its symbiotic partner(s)). These relationships, also called *symbiotic associations*, have however not yet been very widely studied, at least not at a large scale.

One of the problems is getting the data, meaning the trees for hosts and symbionts but even prior to that, determining with which symbionts the present-day hosts are associated (or are “infected” by as may be the term used in some contexts) which is a big enterprise in itself. The other problem is measuring the stability of the association. This has generally been done by concomitantly studying the phylogenies of hosts and symbionts, that is by doing what is called a *cophylogeny* analysis, which itself is often realised by performing what is called a *reconciliation* of two phylogenetic trees (in theory, it could be more than two but this is a problem that has not yet been addressed by the team), one for the symbionts and one for the hosts with which the symbionts are associated. This consists in mapping one of the trees (usually, the symbiont tree) to the other. Cophylogeny inherits all the difficulties of phylogeny, among which the fact that it is not possible to check the result against the “truth” as this is now lost in the past. Cophylogeny however also brings new problems of its own which are to estimate the frequency of the different types of events that could lead to discrepant evolutionary histories, and to estimate the duration of the associations such events may create.

Axis 4: Human, animal and plant health
As indicated above, this is a recent axis in the team and concerns various applications to human and animal health. In some ways, it overlaps with the three previous axes as well as with Axis 5 on the methodological aspects, but since it gained more importance in the past few years, we decided to develop more these particular applications. Most of them started through collaborations with clinicians. Such applications are currently focused on three different topics: (i) Infectiology, (ii) Rare diseases, and (iii) Cancer.

Infectiology is the oldest one. It started by a collaboration with Arnaldo Zaha from the Federal University of Rio Grande do Sul in Brazil that focused on pathogenic bacteria living inside the respiratory tract of swines. Since our participation in the H2020 ITN MicroWine, we started interested in infections affecting plants this time, and more particularly vine plants. Rare Diseases on the other hand started by a collaboration with clinicians from the Centre de Recherche en Neurosciences of Lyon (CNRL) and is focused the Taybi-Linder Syndrome (TALS) and on abnormal splicing of U12 introns, while Cancer rests on a collaboration with the Centre Léon Bérard (CLB) and Centre de Recherche en Cancérologie of Lyon (CRCL) which is focused on Breast and Prostate carcinomas and Gynaecological carcinosarcomas.

The latter collaboration was initiated through a relationship between a member of ERABLE (Alain Viari) and Dr. Gilles Thomas who had been friends since many years. G. Thomas was one of the pioneers of Cancer Genomics in France. After his death in 2014, Alain Viari took the (part time) responsibility of his team at CLB and pursued the main projects he had started.

Within Inria and beyond, the first two applications (Infectiology and Rare Diseases) may be seen as unique because of their specific focus (resp. respiratory tract of swines / vine plants on one hand, and TALS on the other). In the first case, such uniqueness is also related to the fact that the work done involves a strong computational part but also experiments *performed within ERABLE itself*.

4. Application Domains

4.1. Biology and Health

The main areas of application of ERABLE are: (1) biology understood in its more general sense, with a special focus on symbiosis and on intracellular interactions, and (2) health with a special emphasis for now on infectious diseases, rare diseases, and cancer.

5. New Software and Platforms

5.1. C3Part/Isofun

KEYWORDS: Bioinformatics - Genomics

FUNCTIONAL DESCRIPTION: The C3PART / ISOFUN package implements a generic approach to the local alignment of two or more graphs representing biological data, such as genomes, metabolic pathways or protein-protein interactions, in order to infer a functional coupling between them.

- Participants: Alain Viari, Anne Morgat, Frédéric Boyer, Marie-France Sagot and Yves-Pol Deniélou
- Contact: Alain Viari
- URL: http://www.inrialpes.fr/helix/people/viari/lxgraph/index.html

5.2. Cassis

KEYWORDS: Bioinformatics - Genomics
FUNCTIONAL DESCRIPTION: Implements methods for the precise detection of genomic rearrangement breakpoints.

- Participants: Christian Baudet, Christian Gautier, Claire Lemaitre, Eric Tannier and Marie-France Sagot
- Contact: Marie-France Sagot
- URL: http://pbil.univ-lyon1.fr/software/Cassis/

5.3. Coala

CO-evolution Assessment by a Likelihood-free Approach
KEYWORDS: Bioinformatics - Evolution
SCIENTIFIC DESCRIPTION: Despite an increasingly vaster literature on cophylogenetic reconstructions for studying host-parasite associations, understanding the common evolutionary history of such systems remains a problem that is far from being solved. Many of the most used algorithms do the host-parasite reconciliation analysis using an event-based model, where the events include in general (a subset of) cospeciation, duplication, loss, and host-switch. All known event-based methods then assign a cost to each type of event in order to nd a reconstruction of minimum cost. The main problem with this approach is that the cost of the events strongly inuence the reconciliation obtained.

To deal with this problem, we developed an algorithm, called Coala, for estimating the frequency of the events based on an approximate Bayesian computation approach.
FUNCTIONAL DESCRIPTION:COALA stands for "COevolution Assessment by a Likelihood-free Approach". It is thus a likelihood-free method for the co-phylogeny reconstruction problem which is based on an Approximate Bayesian Computation (ABC) approach.

- Participants: Beatrice Donati, Blerina Sinaimeri, Catherine Matias, Christian Baudet, Christian Gautier, Marie-France Sagot and Pierluigi Crescenzi
- Contact: Blerina Sinaimeri
- URL: http://coala.gforge.inria.fr/

5.4. CSC

KEYWORDS: Genomics - Algorithm
FUNCTIONAL DESCRIPTION: Given two sequences x and y, CSC (which stands for Circular Sequence Comparison) finds the cyclic rotation of x (or an approximation of it) that minimises the blockwise q-gram distance from y.

- Contact: Nadia Pisanti
- URL: https://github.com/solonas13/csc

5.5. Cycads

KEYWORDS: Systems Biology - Bioinformatics

FUNCTIONAL DESCRIPTION: Annotation database system to ease the development and update of enriched BIOCYC databases. CYCADS allows the integration of the latest sequence information and functional annotation data from various methods into a metabolic network reconstruction. Functionalities will be added in future to automate a bridge to metabolic network analysis tools, such as METEXPLORE. CYCADS was used to produce a collection of more than 22 arthropod metabolism databases, available at ACYPICYC (http://acypicyc.cycadsys.org) and ARTHROPODACYC (http://arthropodacyc.cycadsys.org). It will continue to be used to create other databases (newly sequenced organisms, Aphid biotypes and symbionts...).

- Participants: Augusto Vellozo, Hubert Charles, Marie-France Sagot and Stefano Colella
- Contact: Hubert Charles
- URL: http://www.cycadsys.org/

5.6. DBGWAS

KEYWORDS: Graph algorithmics - Genomics
FUNCTIONAL DESCRIPTION:DBGWAS is a tool for quick and efficient bacterial GWAS. It uses a compacted De Bruijn Graph (cDBG) structure to represent the variability within all bacterial genome assemblies given as input. Then cDBG nodes are tested for association with a phenotype of interest and the resulting associated nodes are then re-mapped on the cDBG. The output of DBGWAS consists of regions of the cDBG around statistically significant nodes with several informations related to the phenotypes, offering a representation helping in the interpretation. The output can be viewed with any modern web browser, and thus easily shared.

- Contact: Leandro Ishi Soares De Lima
- URL: https://gitlab.com/leoisl/dbgwas

5.7. Eucalypt

KEYWORDS: Bioinformatics - Evolution
FUNCTIONAL DESCRIPTION:EUCALYPT stands for “EnUmerator of Coevolutionary Associations in PoLYnomial-Time delay”. It is an algorithm for enumerating all optimal (possibly time-unfeasible) mappings of a symbiont tree unto a host tree.

- Participants: Beatrice Donati, Blerina Sinaimeri, Christian Baudet, Marie-France Sagot and Pierluigi Crescenzi
- Contact: Blerina Sinaimeri
- URL: http://eucalypt.gforge.inria.fr/

5.8. Fast-SG

KEYWORDS: Genomics - Algorithm - NGS
FUNCTIONAL DESCRIPTION:FAST-SG enables the optimal hybrid assembly of large genomes by combining short and long read technologies.

- Contact: Alex Di Genova
- URL: https://github.com/adigenova/fast-sg

5.9. Gobbolino-Touché

KEYWORDS: Bioinformatics - Graph algorithmics - Systems Biology
FUNCTIONAL DESCRIPTION: Designed to solve the metabolic stories problem, which consists in finding all maximal directed acyclic subgraphs of a directed graph G whose sources and targets belong to a subset of the nodes of G, called the black nodes.

- Participants: Etienne Birmelé, Fabien Jourdan, Ludovic Cottret, Marie-France Sagot, Paulo Vieira Milreu, Pierluigi Crescenzi, Vicente Acuna Aguayo and Vincent Lacroix
- Contact: Marie-France Sagot
- URL: http://gforge.inria.fr/projects/gobbolino

5.10. HapCol

KEYWORDS: Bioinformatics - Genomics
FUNCTIONAL DESCRIPTION: A fast and memory-efficient DP approach for haplotype assembly from long reads that works until 25x coverage and solves a constrained minimum error correction problem exactly.

- Contact: Nadia Pisanti
- URL: http://hapcol.algolab.eu/

5.11. HgLib

HyperGraph Library
KEYWORDS: Graph algorithmics - Hypergraphs
FUNCTIONAL DESCRIPTION: The open-source library hglib is dedicated to model hypergraphs, which are a generalisation of graphs. In an *undirected* hypergraph, an hyperedge contains any number of vertices. A *directed* hypergraph has hyperarcs which connect several tail and head vertices. This library, which is written in C++, allows to associate user defined properties to vertices, to hyperedges/hyperarcs and to the hypergraph itself. It can thus be used for a wide range of problems arising in operations research, computer science, and computational biology.

RELEASE FUNCTIONAL DESCRIPTION: Initial version

- Participants: Martin Wannagat, David Parsons, Arnaud Mary and Irene Ziska
- Contact: Arnaud Mary
- URL: https://gitlab.inria.fr/kirikomics/hglib

5.12. KissDE

KEYWORDS: Bioinformatics - NGS
FUNCTIONAL DESCRIPTION:KISSDE is an R Package enabling to test if a variant (genomic variant or splice variant) is enriched in a condition. It takes as input a table of read counts obtained from an NGS data pre-processing and gives as output a list of condition-specific variants.

RELEASE FUNCTIONAL DESCRIPTION: This new version improved the recall and made more precise the size of the effect computation.

- Participants: Camille Marchet, Aurélie Siberchicot, Audric Cologne, Clara Benoît-Pilven, Janice Kielbassa, Lilia Brinza and Vincent Lacroix
- Contact: Vincent Lacroix
- URL: http://kissplice.prabi.fr/tools/kissDE/

5.13. KisSplice

KEYWORDS: Bioinformatics - Bioinfirmatics search sequence - Genomics - NGS
FUNCTIONAL DESCRIPTION: Enables to analyse RNA-seq data with or without a reference genome. It is an exact local transcriptome assembler, which can identify SNPs, indels and alternative splicing events. It can deal with an arbitrary number of biological conditions, and will quantify each variant in each condition.

RELEASE FUNCTIONAL DESCRIPTION: Improvements : KissReads module has been modified and sped up, with a significant impact on run times. Parameters : –timeout default now at 10000: in big datasets, recall can be increased while run time is a bit longer. Bugs fixed : Reads containing only 'N': the graph construction was stopped if the file contained a read composed only of 'N's. This is was a silence bug, no error message was produced. Problems compiling with new versions of MAC OSX (10.8+): KisSplice is now compiling with the new default C++ compiler of OSX 10.8+.

- Participants: Alice Julien-Laferrière, Leandro Ishi Soares De Lima, Vincent Miele, Rayan Chikhi, Pierre Peterlongo, Camille Marchet, Gustavo Akio Tominaga Sacomoto, Marie-France Sagot and Vincent Lacroix
- Contact: Vincent Lacroix
- URL: http://kissplice.prabi.fr/

5.14. KisSplice2RefGenome

KEYWORDS: Bioinformatics - NGS - Transcriptomics
FUNCTIONAL DESCRIPTION:KISSPLICE identifies variations in RNA-seq data, without a reference genome. In many applications however, a reference genome is available. KISSPLICE2REFGENOME enables to facilitate the interpretation of the results of KISSPLICE after mapping them to a reference genome.

- Participants: Audric Cologne, Camille Marchet, Camille Sessegolo, Alice Julien-Laferrière and Vincent Lacroix
- Contact: Vincent Lacroix
- URL: http://kissplice.prabi.fr/tools/kiss2refgenome/

5.15. KisSplice2RefTranscriptome

KEYWORDS: Bioinformatics - NGS - Transcriptomics
FUNCTIONAL DESCRIPTION:KISSPLICE2REFTRANSCRIPTOME enables to combine the output of KISSPLICE with the output of a full length transcriptome assembler, thus allowing to predict a functional impact for the positioned SNPs, and to intersect these results with condition-specific SNPs. Overall, starting from RNA-seq data only, we obtain a list of condition-specific SNPs stratified by functional impact.

- Participants: Helene Lopez Maestre, Mathilde Boutigny and Vincent Lacroix
- Contact: Vincent Lacroix
- URL: http://kissplice.prabi.fr/tools/kiss2rt/

5.16. MetExplore

KEYWORDS: Systems Biology - Bioinformatics
SCIENTIFIC DESCRIPTION: MetExplore stores metabolic networks of 160 organisms into a relational database. Information about metabolic networks mainly come from BioCyc-like databases. Two BioCyc-like databases contain information about several organisms: PlantCyc and MetaCyc. MetExplore contains also the information about metabolites stored in Metabolome.jp. Note that there is no information about reactions in this database and is only useful to identify compounds from masses. Several genome-scale models designed for Flux Balance Analysis have also been imported into MetExplore. The table below gives details about the sources of the metabolic networks present in MetExplore.
FUNCTIONAL DESCRIPTION: Web-server that allows to build, curate and analyse genome-scale metabolic networks. METEXPLORE is also able to deal with data from metabolomics experiments by mapping a list of masses or identifiers onto filtered metabolic networks. Finally, it proposes several functions to perform Flux Balance Analysis (FBA). The web-server is mature, it was developed in PHP, JAVA, Javascript and Mysql. METEXPLORE was started under another name during Ludovic Cottret's PhD in Bamboo, and is now maintained by the METEXPLORE group at the Inra of Toulouse.

- Participants: Fabien Jourdan, Hubert Charles, Ludovic Cottret and Marie-France Sagot
- Contact: Fabien Jourdan
- URL: https://metexplore.toulouse.inra.fr/index.html/

5.17. Mirinho

KEYWORDS: Bioinformatics - Computational biology - Genomics - Structural Biology
FUNCTIONAL DESCRIPTION: Predicts, at a genome-wide scale, microRNA candidates.

- Participants: Christian Gautier, Christine Gaspin, Cyril Fournier, Marie-France Sagot and Susan Higashi
- Contact: Marie-France Sagot
- URL: http://mirinho.gforge.inria.fr/

5.18. Momo

Multi-Objective Metabolic mixed integer Optimization
KEYWORDS: Metabolism - Metabolic networks - Multi-objective optimisation
FUNCTIONAL DESCRIPTION:MOMO is a multi-objective mixed integer optimisation approach for enumerating knockout reactions leading to the overproduction and/or inhibition of specific compounds in a metabolic network.

- Contact: Marie-France Sagot
- URL: http://momo-sysbio.gforge.inria.fr

5.19. MultiPus

KEYWORDS: Systems Biology - Algorithm - Graph algorithmics - Metabolic networks - Computational biology
SCIENTIFIC DESCRIPTION: Synthetic biology has boomed since the early 2000s when it started being shown that it was possible to efficiently synthetise compounds of interest in a much more rapid and effective way by using other organisms than those naturally producing them. However, to thus engineer a single organism, often a microbe, to optimise one or a collection of metabolic tasks may lead to difficulties when attempting to obtain a production system that is efficient, or to avoid toxic effects for the recruited microorganism. The idea of using instead a microbial consortium has thus started being developed in the last decade. This was motivated by the fact that such consortia may perform more complicated functions than could single populations and be more robust to environmental fluctuations. Success is however not always guaranteed. In particular, establishing which consortium is best for the production of a given compound or set thereof remains a great challenge. The algorithm MultiPus is based on an initial model that enables to propose a consortium to synthetically produce compounds that are either exogenous to it, or are endogenous but where interaction among the species in the consortium could improve the production line.
FUNCTIONAL DESCRIPTION:MULTIPUS (for “MULTIple species for the synthetic Production of Useful biochemical Substances”) is an algorithm that, given a microbial consortium as input, identifies all optimal sub-consortia to synthetically produce compounds that are either exogenous to it, or are endogenous but where interaction among the species in the sub-consortia could improve the production line.

- Participants: Alberto Marchetti-Spaccamela, Alice Julien-Laferrière, Arnaud Mary, Delphine Parrot, Laurent Bulteau, Leen Stougie, Marie-France Sagot and Susana Vinga
- Contact: Marie-France Sagot
- URL: http://multipus.gforge.inria.fr/

5.20. Pitufolandia

KEYWORDS: Bioinformatics - Graph algorithmics - Systems Biology
FUNCTIONAL DESCRIPTION: The algorithms in PITUFOLANDIA (PITUFO / PITUFINA / PAPAPITUFO) are designed to solve the minimal precursor set problem, which consists in finding all minimal sets of precursors (usually, nutrients) in a metabolic network that are able to produce a set of target metabolites.

- Contact: Marie-France Sagot
- URL: http://gforge.inria.fr/projects/pitufo/

5.21. Sasita

KEYWORDS: Bioinformatics - Graph algorithmics - Systems Biology
FUNCTIONAL DESCRIPTION:SASITA is a software for the exhaustive enumeration of minimal precursor sets in metabolic networks.

- Contact: Marie-France Sagot
- URL: http://sasita.gforge.inria.fr/

5.22. Savage

KEYWORDS: Algorithm - Genomics
FUNCTIONAL DESCRIPTION: Reconstruction of viral quasi species without using a reference genome.

- Contact: Alexander Schonhuth
- URL: https://bitbucket.org/jbaaijens/savage

5.23. Smile

KEYWORDS: Bioinformatics - Genomic sequence
FUNCTIONAL DESCRIPTION: Motif inference algorithm taking as input a set of biological sequences.

- Participant: Marie-France Sagot
- Contact: Marie-France Sagot

5.24. Rime

KEYWORDS: Bioinformatics - Genomics - Sequence alignment
FUNCTIONAL DESCRIPTION: Detects long similar fragments occurring at least twice in a set of biological sequences.

- Contact: Nadia Pisanti

5.25. Totoro & Kotoura

KEYWORDS: Bioinformatics - Graph algorithmics - Systems Biology
FUNCTIONAL DESCRIPTION: Both TOTORO and KOTOURA decipher the reaction changes during a metabolic transient state, using measurements of metabolic concentrations. These are called metabolic hyperstories. TOTORO (for TOpological analysis of Transient metabOlic RespOnse) is based on a qualitative measurement of the concentrations in two steady-states to infer the reaction changes that lead to the observed differences in metabolite pools in both conditions. In the currently available release, a pre-processing and a post-processing steps are included. After the post-processing step, the solutions can be visualised using DINGHY (http://dinghy.gforge.inria.fr). KOTOURA (for Kantitative analysis Of Transient metabOlic and regUlatory Response And control) infers quantitative changes of the reactions using information on measurement of the metabolite concentrations in two steady-states.

- Contact: Marie-France Sagot
- URL: http://hyperstories.gforge.inria.fr/

5.26. WhatsHap

KEYWORDS: Bioinformatics - Genomics
FUNCTIONAL DESCRIPTION:WHATSHAP is a DP approach for haplotype assembly from long reads that works until 20x coverage and solves the minimum error correction problem exactly. PWHATSHAP is a parallelisation of the core dynamic programming algorithm of WHATSHAP.

- Contact: Nadia Pisanti
- URL: https://bitbucket.org/whatshap/whatshap

6. New Results

6.1. General comments

We present in this section the main results obtained in 2018.

We tried to organise these along the four axes as presented above. Clearly, in some cases, a result obtained overlaps more than one axis. In such case, we chose the one that could be seen as the main one concerned by such results.

We did not indicate here the results on more theoretical aspects of computer science if it did not seem for now that they could be relevant in contexts related to computational biology. Actually, those on string [32], [33], [36], [11] and graph algorithms in general [2], [35], [38], [37], [39], [41], [54], [42], [44], [43], [40], [47], [5], [45], [48], [49], [53], [23], [24], or on more general algorithmic problems notably related to data structures are already relevant for life sciences (biology or ecology) or in the future could become more specifically so. We do in particular believe that dynamic graph approaches could be of great interest in the future for some of the enumeration problems we constantly meet in biology.

A few other results of 2018 are not mentioned in this report, not because the corresponding work is not important, but because it was likewise more specialised [52], or the work represented a survey *e.g.* [21]). Likewise, also for space reasons, we do not detail the results presented in some biological papers of the team when these did not require a mathematical or algorithmical input or are surveys [1], [4], [7], [8], [12], [19], [20], [22], [25], [31].

On the other hand, we do mention a couple of works that were submitted towards the end of 2018.

6.2. Axis 1: Genomics

Genome hybrid assembly

Long read sequencing technologies are considered to be the solution for handling genome repeats, allowing near reference-level reconstructions of large genomes. However, long read de novo assembly pipelines are computationally intense and require a considerable amount of coverage, thereby hindering their broad application to the assembly of large genomes. Alternatively, hybrid assembly methods that combine short and long read sequencing technologies can reduce the time and cost required to produce de novo assemblies of large genomes. In [10], we proposed a new method, called FAST-SG, that uses a new ultrafast alignment-free algorithm specifically designed for constructing a scaffolding graph using lightweight data structures. FAST-SG can construct the graph from either short or long reads. This allows the reuse of efficient algorithms designed for short read data and permits the definition of novel modular hybrid assembly pipelines. Using comprehensive standard datasets and benchmarks, we showed how FAST-SG outperforms the state-of-the-art short read aligners when building the scaffolding graph and can be used to extract linking information from either raw or error-corrected long reads. We also showed how a hybrid assembly approach using FAST-SG with shallow long-read coverage (5X) and moderate computational resources can produce long-range and accurate reconstructions of the genomes of *Arabidopsis thaliana* (Ler-0) and human (NA12878). We are currently working on the assembly process itself, using the scaffolding graphs obtained with FAST-SG. The results obtained so far are extremely promising and a paper is currently in preparation. This is part of the work done by Alex di Genova, postdoc in ERABLE.

Variant annotation

Genome-wide analyses estimate that more than 90% of multi exonic human genes produce at least two transcripts through a genomic variant called alternative splicing (AS). Various bioinformatics methods are available to analyse AS from RNAseq data. Most methods start by mapping the reads to an annotated reference genome, but some start by a *de novo* assembly of the reads. In [3], we presented a systematic comparison of a mapping-first approach (FARLINE) and an assembly-first approach (scKisSplice). We applied these methods to two independent RNAseq datasets and found that the predictions of the two pipelines overlapped (70% of exon skipping events were common), but with noticeable differences. The assembly-first approach allowed to find more novel variants, including novel unannotated exons and splice sites. It also predicted AS in recently duplicated genes. The mapping-first approach allowed to find more lowly expressed splicing variants, and splice variants overlapping repeats. This work demonstrated that annotating AS with a single approach leads to missing out a large number of candidates, many of which are differentially regulated across conditions and can be validated experimentally. We therefore advocate for the combined use of both mapping-first and assembly-first approaches for the annotation and differential analysis of AS from RNAseq datasets. This was part of the work of Clara Benoît-Pilven, postdoc at Inserm and in ERABLE, to which also participated other current or ex-members of ERABLE, namely Camille Marchet (during her stay as ADT engineer with ERABLE), Emilie Chautard (when she was postdoc Inserm and in ERABLE), Gustavo Sacomoto (when he was PhD and then for one year postdoc in ERABLE), and Leandro Lima (current PhD student of ERABLE).

Another type of variant, namely SNPs was also considered in [51]. In this paper, mutations are detected by eBWT (extended Burrows-Wheeler Transform). Indeed, we notices that eBWT of a collection of DNA fragments tend to cluster together the copies of nucleotides sequenced from a genome. We showed that it is thus possible to accurately predict how many copies of any nucleotide are expected inside each such cluster, and that a precise LCP array based procedure can locate these clusters in the eBWT. These theoretical insights were validated in practice with SNPs being clustered in the eBWT of a reads collection. We developed a tool for finding SNPs with a simple scan of the eBWT and LCP arrays. Preliminary results show that our method requires much less coverage than the state-of-the-art tools while drastically improving precision and sensitivity.

Both types of variants correspond to special types of st-paths in graphs, a topic that was also explored from a more purely theoretical point of view in two papers, one already accepted [46] and on that is about to be submitted and extends the results obtained in 2017 on bubble (as st-paths are also called in bioinformatics) generators in directed graphs.

Full-length *de novo* viral quasispecies assembly through variation graph construction

Viruses populate their hosts as a viral quasispecies: a collection of genetically related mutant strains. Viral quasispecies assembly refers to reconstructing the strain-specific haplotypes from read data, and predicting their relative abundances within the mix of strains, an important step for various treatment-related reasons. Reference-genome-independent ("de novo") approaches have yielded benefits over reference-guided approaches, because reference-induced biases can become overwhelming when dealing with divergent strains. While being very accurate, extant *de novo* methods only yield rather short contigs. It remains to reconstruct full-length haplotypes together with their abundances from such contigs. In [34], we first constructed a variation graph, a recently popular, suitable structure for arranging and integrating several related genomes, from the short input contigs, without making use of a reference genome. To obtain paths through the variation graph that reflect the original haplotypes, we solved a minimisation problem that yields a selection of maximal-length paths that is optimal in terms of being compatible with the read coverages computed for the nodes of the variation graph. We output the resulting selection of maximal length paths as the haplotypes, together with their abundances. Benchmarking experiments on challenging simulated data sets showed significant improvements in assembly contiguity compared to the input contigs, while preserving low error rates. As a consequence, our method outperforms all state-of-the-art viral quasispecies assemblers that aim at the construction of full-length haplotypes, in terms of various relevant assembly measures. The tool, called VIRUS-VG, is availabe at https://bitbucket.org/jbaaijens/virus-vg.

A member of ERABLE was also involved in the Second Annual Meeting of the European Virus Bioinformatics Center (EVBC), held in Utrecht, Netherlands, and whose focus was on computational approaches in virology, with topics including (but not limited to) virus discovery, diagnostics, (meta-)genomics, modeling, epidemiology, molecular structure, evolution, and viral ecology. Approximately 120 researchers from around the world attended the meeting this year. An overview of new developments and novel research findings that emerged during the meeting was published in the journal *Viruses* [16].

Bacterial genome-wide association studies (GWAS)
Genome-wide association study (GWAS) methods applied to bacterial genomes have shown promising results for genetic marker discovery or detailed assessment of marker effect. Recently, alignment-free methods based on k-mer composition have proven their ability to explore the accessory genome. However, they lead to redundant descriptions and results which are sometimes hard to interpret. In [17], we introduced DBGWAS, an extended k-mer-based GWAS method producing interpretable genetic variants associated with distinct phenotypes. Relying on compacted de Bruijn graphs (cDBG), our method gathers cDBG nodes, identified by the association model, into subgraphs defined from their neighbourhood in the initial cDBG. DBGWAS is alignment-free and only requires a set of contigs and phenotypes. In particular, it does not require prior annotation or reference genomes. It produces subgraphs representing phenotype-associated genetic variants such as local polymorphisms and mobile genetic elements (MGE). It offers a graphical framework which helps interpret GWAS results. Importantly, it is also computationally efficient (the experiments took one hour and a half on average). We validated our method using antibiotic resistance phenotypes for three bacterial species. DBGWAS recovered known resistance determinants such as mutations in core genes in *Mycobacterium tuberculosis*, and genes acquired by horizontal transfer in *Staphylococcus aureus* and *Pseudomonas aeruginosa* along with their MGE context. It also enabled us to formulate new hypotheses involving genetic variants not yet described in the antibiotic resistance literature. This is part of the work of Magali Jaillard, PhD student of Laurent Jacob who is an external collaborator of ERABLE, and of Leandro I. S. de Lima, PhD student co-supervised by three members of ERABLE.

6.3. Axis 2: Metabolism and post-transcriptional regulation

Multi-objective metabolic mixed integer optimisation: with an application to yeast strain engineering
In a paper submitted and already available in bioRxiv (https://www.biorxiv.org/content/early/2018/11/22/476689), we explored the concept of multi-objective optimisation in the field of metabolic engineering when both continuous and integer decision variables are involved in the model. In particular, we proposed a multi-objective model which may be used to suggest reaction deletions that maximise and/or minimise several functions simultaneously. The applications may include, among others, the concurrent maximisation of a bioproduct and of biomass, or maximisation of a bioproduct while minimising the formation of a given by-product, two common requirements in microbial metabolic engineering. Production of ethanol by the widely used cell factory *Saccharomyces cerevisiae* was adopted as a case study to demonstrate the usefulness of the proposed approach in identifying genetic manipulations that improve productivity and yield of this economically highly relevant bioproduct. We did an *in vivo* validation and we could show that some of the predicted deletions exhibit increased ethanol levels in comparison with the wild-type strain. The multi-objective programming framework we developed, called MOMO, is open-source and uses POLYSCIP as underlying multi-objective solver. This is part of the work of Ricardo de Andrade, postdoc at Unversity of São Paulo with Roberto Marcondes, and in ERABLE. It is joint work with Susana Vinga, external collaborator of ERABLE and partner of the Inria Associated Team Compasso.

Metabolic shifts

With the increasing availability of so-called 'omics data – transcriptomics, proteomics, and metabolics – there has been growing interest in various ways of integrating them with the metabolic network. When the network is represented by a graph, 'omics data can guide the extraction of subnetworks of interest to find metabolic pathways or sets of related genes. Within the framework of constraint-based modelling, 'omics data can be used to improve the prediction of metabolic behaviour and to build context-specific metabolic models. One interesting application of metabolic reconstructions in conjunction with 'omics data is to use the two to understand metabolic shifts. When an organism encounters a change in environmental conditions, often a re-organisation of metabolism follows. Comparative measurements of gene expression and metabolite concentrations can be used to gain insight into these changes but this data is "structureless", meaning it lacks the information about how the metabolic components relate to each other. A metabolic network on the other hand contains this information, and can thus greatly benefit such an analysis. We developed a new method, called MOOMIN, that combines the results of a differential expression analysis comparing the gene expression levels in two different conditions with a metabolic network to produce a hypothesis of a metabolic shift. The idea is to use the network structure to define feasible global changes in metabolism. These changes are then scored based on the gene expression data with the goal of finding the change that best agrees with the observations. Finding the best-scoring change is formulated into an optimisation problem that can be solved using Mixed-Integer Linear Programming. This is part of the work of Henri Taneli Pusa, co-supervised by 3 members of ERABLE, whose manuscript was submitted to the reviewers and who should be defending his PhD in early February 2019. The paper on MOOMIN will be submitted soon, and the software then made available. Participated also in this work Mariana G. Ferrarini, postdoc at Insa and in ERABLE, and Ricardo Andrade, postdoc at University of São Paulo with Roberto Marcondes and in ERABLE.

Metabolic games

The PhD of Taneli also investigated game theory in the context of metabolism. Game theory is a branch of applied mathematics that deals with interacting rational agents with conflicting goals. When rationality is replaced with natural selection, *evolutionary* game theory can be used to explain the "decisions" taken by even microscopic organisms. The PhD manuscript presents the idea of a *metabolic game*, a game theoretical model for the prediction of metabolic behaviour. In contrast to Flux Balance Analysis, where the metabolic state is predicted using simple optimisation, a metabolic game takes into account the fact that optimality is influenced by the surrounding members of a microbial community. By changing the availability of nutrients, or secreting beneficial or harmful molecules, microbes essentially create their own environment and make optimal behaviour context-specific. A paper is submitted that reviews the literature that has applied game theory to the study of microbes, with a focus on metabolism and especially games derived using metabolic networks and constraint-based modelling. In the PhD manuscript, Taneli further explains the idea behind a metabolic game and discusses different aspects of defining such games: the choice of players, actions, and payoffs.

6.4. Axis 3: (Co)Evolution

Exploring the robustness of the parsimonious reconciliation method in host-symbiont cophylogeny

Following our previous work on reconciliation methods for cophylogeny, in [29], we explored the robustness of the parsimonious host-symbiont tree reconciliation method under editing or small perturbations of the input. The editing involved making different choices of unique symbiont mapping to a host in the case where multiple associations exist. This is made necessary by the fact that the tree reconciliation model is currently unable to handle such associations. The analysis performed could however also address the problem of errors. The perturbations were re-rootings of the symbiont tree to deal with a possibly wrong placement of the root specially in the case of fast-evolving species. In order to do this robustness analysis, we introduced a simulation scheme specifically designed for the host-symbiont cophylogeny context, as well as a measure to compare sets of tree reconciliations, both of which are of interest by themselves. This work was also part of the PhD of a previous student of ERABLE, Laura Urbini.

Geometric medians in reconciliation spaces

Recently, there has been much interest in studying spaces of tree reconciliations (as used in cophylogenetic studies), which arise by defining some metric d on the set $\mathcal{R}(P,H,\phi)$ of all possible reconciliations between two trees P and H where ϕ represents the map between the leaf-sets of P and H (corresponding to present-day associations). In [14], we studied the following question: how do we compute a *geometric median* for a given subset Ψ of $\mathcal{R}(P,H,\phi)$ relative to d, *i.e.* an element $\psi_{med} \in \mathcal{R}(P,H,\phi)$ such that

$$\sum_{\psi' \in \Psi} d(\psi_{med}, \psi') \leq \sum_{\psi' \in \Psi} d(\psi, \psi')$$

holds for all $\psi \in \mathcal{R}(P,H,\phi)$? For a model where so-called host-switches or transfers are not allowed, and for a commonly used metric d called the *edit-distance*, we showed that although the cardinality of $\mathcal{R}(P,H,\phi)$ can be super-exponential, it is still possible to compute a geometric median for a set Ψ in $\mathcal{R}(P,H,\phi)$ in polynomial time. We expect that this result could be useful for computing a summary or consensus for a set of reconciliations (*e.g.* for a set of suboptimal reconciliations). The collaboration with Katharina Huber and Vincent Moulton from the School of Computing Sciences at the University of New Anglia was made possible by a Royal Society Grant obtained by the two partners (UNA and ERABLE).

Exploring and Visualising Spaces of Tree Reconciliations

A common approach to tree reconciliation involves specifying a model that assigns costs to certain events, such as cospeciation, and then tries to find a mapping between two specified phylogenetic trees which minimises the total cost of the implied events. For such models, it has been shown, including by the ERABLE members in previous papers, that there may be a huge number of optimal solutions, or at least solutions that are close to optimal. It is therefore of interest to be able to systematically compare and visualise whole collections of reconciliations between a specified pair of trees. In [13] , we considered various metrics on the set of all possible reconciliations between a pair of trees, some that have been defined before but also new metrics that we proposed. We showed that the diameter for the resulting spaces of reconciliations can in some cases be determined theoretically, information that we used to normalise and compare properties of the metrics. We also implemented the metrics and compared their behaviour on several host parasite datasets, including the shapes of their distributions. In addition, we showed that in combination with multidimensional scaling, the metrics can be useful for visualising large collections of reconciliations, much in the same way as phylogenetic tree metrics can be used to explore collections of phylogenetic trees. Implementations of the metrics can be downloaded from https://team.inria.fr/erable/en/team-members/blerina-sinaimeri/reconciliation-distances/. This work was also funded by a Royal Society Grant obtained by the two partners (at University of New Anglia and ERABLE).

Variants of phylogenetic network problems

Although not falling within the general topic of coevolution, phylogenetic networks are of great interest as another way of representing the evolution of a set of species. In the context of such representations, unrooted and root-uncertain variants of several well-known phylogenetic network problems were explored. The hybridisation number problem requires to embed a set of binary rooted phylogenetic trees into a binary rooted phylogenetic network such that the number of nodes with indegree two is minimised. However, from a biological point of view accurately inferring the root location in a phylogenetic tree is notoriously difficult and poor root placement can artificially inflate the hybridisation number. To this end, we studied in [30] a number of relaxed variants of this problem. We started by showing that the fundamental problem of determining whether an unrooted phylogenetic network displays (*i.e.* embeds) an unrooted phylogenetic tree, is NP-hard. On the positive side, we show that this problem is FPT in the reticulation number. In the rooted case, the corresponding FPT result is trivial, but here we required more subtle argumentation. Next we showed that the hybridisation number problem for unrooted networks (when given two unrooted trees) is equivalent to the problem of computing the tree bisection and reconnect distance of the two unrooted trees. In the third part of the paper, we considered the “root uncertain” variant of hybridisation number. Here we were free to choose the root location in each of a set of unrooted input trees such that the hybridisation number of the resulting rooted trees is minimised. On the negative side, we showed that this problem is APX-hard. On the positive side, we showed that the problem is FPT in the hybridisation number, via kernelisation, for any number of input trees.

6.5. Axis 4: Human, animal and plant health

Hydrogen peroxide production and myo-inositol metabolism as important traits for virulence of Mycoplasma hyopneumoniae
Mycoplasma hyopneumoniae is the causative agent of enzootic pneumonia. In a previous work, we had reconstructed the metabolic models of this species along with two other mycoplasmas from the respiratory tract of swine: *Mycoplasma hyorhinis*, considered less pathogenic but which nonetheless causes disease and *Mycoplasma flocculare*, a commensal bacterium. We had identified metabolic differences that partially explained their different levels of pathogenicity. One important trait was the production of hydrogen peroxide from the glycerol metabolism only in the pathogenic species. Another important feature was a pathway for the metabolism of myo-inositol in *M. hyopneumoniae*. In the paper accepted this year [9], we tested these traits to understand their relation to the different levels of pathogenicity, comparing not only the species but also pathogenic and attenuated strains of *M. hyopneumoniae*. Regarding the myo-inositol metabolism, we showed that only *M. hyopneumoniae* assimilated this carbohydrate and remained viable when myo-inositol was the primary energy source. Strikingly, only the two pathogenic strains of *M. hyopneumoniae* produced hydrogen peroxide in complex medium. We also showed that this production was dependent on the presence of glycerol. Although further functional tests are needed, this work enabled to identify two interesting metabolic traits of *M. hyopneumoniae* that might be directly related to its enhanced virulence. This is part of the work of Mariana G. Ferrarini, currently postdoc at Insa and in ERABLE, and of Scheila G. Mucha whose PhD (defended in Sept. 2018) was co-supervised by Arnaldo Zaha and by a member of ERABLE.

Cancer
A member of ERABLE continues deeply involved with the Centre Léon Bérard in Lyon, and in that context, a number of works are running, all related to cancer genomics. In the first [28], an integrated genomic study was performed of 25 tumour tissues from radical prostatectomy of aggressive (defined by the International Society of Urological Pathology) prostate cancer patients (10 African Caribbean and 15 French Caucasian) using single nucleotide polymorphism arrays, whole-genome sequencing, and RNA sequencing. The results showed that African Caribbean tumours are characterised by a more frequent deletion at 1q41-43 encompassing the DNA repair gene PARP1, and a higher proportion of intra-chromosomal rearrangements including duplications associated with CDK12 truncating mutations. Transcriptome analyses showed an over-expression of genes related to androgen receptor activity in African Caribbean tumours, and of PVT1, a long non-coding RNA located at 8q24 that confirms the strong involvement of this region in prostate tumours from men of African ancestry. In a second study [15], gene-expression profiling data was used to build and validate a predictive model of outcome for patients with follicular lymphoma. A robust 23-gene expression-based predictor of progression-free survival that is applicable to routinely available formalin-fixed, paraffin-embedded tumour biopsies from such patients was thus developed and validated. Applying this score could allow individualised therapy for patients according to their risk category. In a third study, an integrated analysis highlighted APC11 protein expression as a likely new independent predictive marker for colorectal cancer [6].

In a parallel work by another member of ERABLE [27], it was proposed that cancer Is not (only) a senescence problem. Age is indeed one of the strongest predictors of cancer and risk of death from cancer. Cancer is therefore generally viewed as a senescence-related malady. However, cancer also exists at subclinical levels in humans and other animals, but its earlier effects on the body are poorly known by comparison. What was argued in [27] is that cancer is a significant but ignored burden on the body and is likely to be a strong selective force from early during the lifetime of an organism. It was thus proposed that time has come to adopt this novel view of malignant pathologies to improve our understanding of the ways in which oncogenic phenomena influence the ecology and evolution of animals long before their negative impacts become evident and fatal.

Xylella fastidiosa epidemiological model

Xylella fastidiosa is a notorious plant pathogenic bacterium that represents a threat to crops worldwide. Its subspecies, *Xylella fastidiosa* subsp. *fastidiosa* is the causal agent of Pierce's disease of grapevines. Pierce's disease has presented a serious challenge for the grapevine industry in the United States and turned into an epidemic in Southern California due to the invasion of the insect vector *Homalodisca vitripennis*. In an attempt to minimize the effects of *Xylella fastidiosa* subsp. *fastidiosa* in vineyards, various studies have been developing and testing strategies to prevent the occurrence of Pierce's disease, *i.e.*, prophylactic strategies. Research has also been undertaken to investigate therapeutic strategies to cure vines infected by *Xylella fastidiosa* subsp. *fastidiosa*. In [18], we explicitly review all the strategies published to date and specifies their current status. Furthermore, an epidemiological model of *Xylella fastidiosa* subsp. *fastidiosa* is proposed and key parameters for the spread of Pierce's disease deciphered in a sensitivity analysis of all model parameters. Based on these results, it is concluded that future studies should prioritise therapeutic strategies, while investments should only be made in prophylactic strategies that have demonstrated promising results in vineyards. This is part of the PhD of Henri Taneli Pusa in the context of the H2020 ITN MicroWine, together with another PhD student of the ITN, Ifigeneia Kyrkou. Ifigeneia was the first author of the paper [18] but the mathematical model is the work of Taneli.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants with Industry

7.1.1. Spock

- Title: characterization of hoSt-gut microbiota interactions and identification of key Players based on a unified reference for standardized quantitative metagenOmics and metaboliC analysis frameworK
- Industrial Partner: MaatPharma (Person responsible: Lilia Boucinha).
- ERABLE participants: Marie-France Sagot (ERABLE coordinator and PhD main supervisor with Susana Vinga from IST, Lisbon, Portugal, as PhD co-supervisor), Marianne Borderes (beneficiary of the PhD scholarship in MaatPharma).
- Type: ANR Technology (2018-2021).
- Web page: http://team.inria.fr/erable/en/projects/#anr-technology-spock.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. Muse

- Title: Multi-Omics and Metabolic models iNtegration to study growth Transition in *Escherichia coli*
- Coordinators: Delphine Ropers (EPI Ibis) and Marie-France Sagot
- ERABLE participants: Marie-France Sagot and Arnaud Mary.
- Type: IXXI Project (2018-2019).
- Web page: none for now.

8.2. National Initiatives

8.2.1. ANR

8.2.1.1. Aster

- Title: Algorithms and Software for Third gEneration Rna sequencing

- Coordinator: Hélène Touzet, University of Lille and CNRS.
- ERABLE participants: Vincent Lacroix (ERABLE coordinator), Clara Benoît-Pilven, Audric Cologne, Alex di Genova, Leandro I. S. de Lima, Arnaud Mary, Marie-France Sagot, Camille Sessegolo, Blerina Sinaimeri.
- Type: ANR (2016-2020).
- Web page: http://bioinfo.cristal.univ-lille.fr/aster/.

8.2.1.2. ExHyb

- Title: Exploring genomic stability in hybrids
- Coordinator: C. Vieira
- ERABLE participant(s): C. Vieira
- Type: ANR (2014-2018)
- Web page: Not available

8.2.1.3. GraphEn

- Title: Enumération dans les graphes et les hypergraphes : Algorithmes et complexité
- Coordinator: D. Kratsch
- ERABLE participant(s): A. Mary
- Type: ANR (2015-2019)
- Web page: http://graphen.isima.fr/

8.2.1.4. GrR

- Title: Graph Reconfiguration
- Coordinator: N. Bousquet
- ERABLE participant(s): A. Mary
- Type: ANR JCJC (2019-2021)
- Web page: Not available

8.2.1.5. Green

- Title: Deciphering host immune gene regulation and function to target symbiosis disturbance and endosymbiont control in insect pests
- Coordinator: A. Heddi
- ERABLE participant(s): M.-F. Sagot, C. Vieira
- Type: ANR (2018-2021)
- Web page: Not yet available

8.2.1.6. Hmicmac

- Title: Host-microbiota co-adaptations: mechanisms and consequences
- Coordinator: F. Vavre
- ERABLE participant(s): F. Vavre
- Type: ANR PRC (2017-2020)
- Web page: Not available

8.2.1.7. IMetSym

- Title: Immune and Metabolic Control in Intracellular Symbiosis of Insects
- Coordinator: A. Heddi
- ERABLE participant(s): H. Charles, S. Colella
- Type: ANR Blanc (2014-2017)

- Web page: Not available

8.2.1.8. Resist

- Title: Rapid Evolution of Symbiotic Interactions in response to STress: processes and mechanisms
- Coordinator: N. Kremer
- ERABLE participant(s): F. Vavre
- Type: ANR JCJC (2017-2020)
- Web page: Not available

8.2.1.9. Suzukill

- Title: Managing cold tolerance and quality of mass-produced *Drosophila suzukii* flies to facilitate the application of biocontrol through incompatible and sterile insect techniques
- Coordinator: H. Colinet
- ERABLE participant(s): F. Vavre
- Type: ANR PCRI (2015-2018)
- Web page: Not available

8.2.1.10. Swing

- Title: Worldwide invasion of the Spotted WING Drosophila: Genetics, plasticity and evolutionary potential
- Coordinator: P. Gibert
- ERABLE participant(s): C. Vieira
- Type: ANR PCR (2016-2020)
- Web page: Not available

8.2.1.11. U4atac-brain

- Title: Rôle de l'épissage mineur dans le développement cérébral
- Coordinator: Patrick Edery, Centre de Recherche en Neurosciences de Lyon.
- ERABLE participants: Vincent Lacroix (ERABLE coordinator), Clara Benoît-Pilven, Audric Cologne.
- Type: ANR (2018-2021).
- Web page: Not available.

8.2.2. Idex

8.2.2.1. Micro-be-have

- Title: Microbial Impact on insect behaviour: from niche and partner selection to the development of new control methods for pests and disease vectors
- Coordinator: F. Vavre
- ERABLE participant(s): F. Vavre
- Type: AO Scientific Breakthrough (2018-2021)
- Web page: Not available

8.2.3. ADT Inria

8.2.3.1. ADT Inria Kirikomics

- Main objective: Development of a portal to increase the visibility of the tools and resources elaborated by ERABLE around the analysis – using omics data – of metabolic networks modelled by hypergraphs, and enable to visualise the results. (the web page is for now private, it will be made public later in the project).
- Duration: 2016-2017, renewable one more year.
- Person responsible for ADT: Arnaud Mary with David Parsons (Inria).
- Beneficiary of ADT: Martin Wannagat.
- Funds received: Salary for engineer.

8.2.4. Others

Notice that were included here national projects of our members from Italy and the Netherlands when these have no other partners than researchers from the same country.

8.2.4.1. Advanced computational methodologies for the analysis of biomedical data

- Title: Advanced computational methodologies for the analysis of biomedical data
- Coordinator: P. Milazzo
- ERABLE participant(s): R. Grossi, N. Pisanti
- Type: PRA, MIUR PRIN, Italian Ministery of Research National Projects (2017-2018)
- Web page: Not available

8.2.4.2. Advanced Tools and Techniques for the analysis of criminal networks

- Title: Advanced Tools and Techniques for the analysis of criminal networks
- Coordinator: G. Italiano
- ERABLE participant(s): G. Italiano
- Type: LEONARDO SpA (2015-2018)
- Web page: Not available

8.2.4.3. Open Innovation: Digital Innovation for Driving

- Title: Open Innovation: Digital Innovation for Driving
- Coordinator: G. Italiano
- ERABLE participant(s): G. Italiano
- Type: Bridgestone (2018-2019)
- Web page: Not available

8.2.4.4. CMACBioSeq

- Title: Combinatorial Methods for analysis and compression of biological sequences
- Coordinator: G. Rosone
- ERABLE participant(s): N. Pisanti
- Type: SIR, MIUR PRIN, Italian Ministery of Research National Projects (2015-2019)
- Web page: http://pages.di.unipi.it/rosone/CMACBioSeq.html

8.2.4.5. Statistical Models for Structural Genetic Variants in the Genome of the Netherlands

- Title: Statistical Models for Structural Genetic Variants in the Genome of the Netherlands
- Coordinator: A. Schönhuth
- ERABLE participant(s): A. Schönhuth
- Type: Nederlandse Wetenschappelijke Organisatie (NWO) (2013-2018)
- Web page: Not available

8.2.4.6. TALS and splicing

- Title: Development of bioinformatic methods for the analysis of splicing events in patients with the Taybi-Linder Syndrome (TALS)
- Coordinator: P. Edery
- ERABLE participant(s): C. Benoît-Pilven, Audric Cologne, V. Lacroix
- Type: INSERM
- Web page: Not available

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

8.3.1.1. MicroWine

- Title: Microbial metagenomics and the modern wine industry

- Duration: January 2015 - January 2019
- Coordinator: Lars Hestbjerg Hansen, University of Copenhagen
- ERABLE participant(s): A. Marchetti-Spaccamela, A. Mary, H. T. Pusa, M.-F. Sagot, L. Stougie
- Type: H2020-MSCA-ETN-2014
- Web page: https://team.inria.fr/erable/en/microwine/ and http://www.microwine.eu/

8.3.2. Collaborations in European Programs, Except FP7 & H2020

8.3.2.1. Combinatorics of co-evolution

- Title: The combinatorics of co-evolution
- Duration: 2015 - 2018
- Coordinator: Katharina Huber, University of Warwick, UK
- ERABLE participant(s): M.-F. Sagot, B. Sinaimeri
- Type: The Royal Society
- Web page: not available

8.3.3. Collaborations with Major European Organizations

By itself, ERABLE is built from what initially were collaborations with some major European Organisations (CWI, Sapienza University of Rome, Universities of Florence and Pisa, Free University of Amsterdam) and now has become a European Inria Team.

8.4. International Initiatives

8.4.1. Inria Associate Teams Not Involved in an Inria International Labs

Compasso

- Title: COMmunity Perspective in the health sciences: Algorithms and Statistical approacheS for explOring it
- Duration: 2018, renewable from 2 to 5 years more
- Coordinator: On the Portuguese side, Susana Vinga, IST, Lisbon, Portugal; on the French side, Marie-France Sagot
- ERABLE participant(s): R. Andrade, M. Ferrarini, G. Italiano, A. Marchetti-Spaccamela, A. Mary, H. T. Pusa, M.-F. Sagot, B. Sinaimeri, L. Stougie, A. Viari, I. Ziska
- Web page: http://team.inria.fr/erable/en/projects/inria-associated-team-compasso/

8.4.2. Participation in International Programs

ERABLE is coordinator of a CNRS-UCBL-Inria Laboratoire International Associé (LIA) with the Laboratório Nacional de Computação Científica (LNCC), Petrópolis, Brazil. The LIA has for acronym LIRIO ("Laboratoire International de Recherche en bIOinformatique") and is coordinated by Ana Tereza Vasconcelos from the LNCC and Marie-France Sagot from BAOBAB-ERABLE. The LIA was created in January 2012 for 4 years, renewable once for 4 more years. A web page for the LIA LIRIO is available at this address: http://team.inria.fr/erable/en/cnrs-lia-laboratoire-international-associe-lirio/.

ERABLE also participated to the BASIS project. This was funded by the European Community Seventh Framework Programme (Grant 242006 - 2010-2015). It was led by Dr. Mike Stratton and involved six European countries. It was primarily focused on ER+/HER2- breast cancers, but during the course of the project, was merged with the HER2+ French-ICGC and triple negative UK-ICGC projects, resulting in the analysis of the whole spectrum of breast cancers. The French group was initiated by Dr. Gilles Thomas and was pursued by Alain Viari after the loss of Dr. Thomas in 2014. The project resulted in the sequencing and thorough analysis of 560 breast cancer whole genomes (Nik-Zainai *et al.*, *Nature*, 534:47-54, 2016), including 75 HER2+ performed by the French working group (Ferrari *et al.*, *Nature Communications*, 7, 2016) and funded by the Institut National du Cancer and by Inserm.

Finally, Marie-France Sagot participates in a Portuguese FCT project, Perseids for “Personalizing cancer therapy through integrated modeling and decision” (2016-2019), with Susana Vinga and a number of other Portuguese researchers. The budget of Perseids is managed exclusively by the other Portuguese partner.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

In 2018, ERABLE greeted the following International scientists:

- In France: Alexander Stuart Ralph (University of Melbourne), Katharina Huber and Vincent Moulton (University of Warwick, UK), Ifigeneia Kyrkou (Aarhus University, Danemark), Ana Tereza Vasconcelos from the LNCC, Brazil), Nuno Mira (IST Portugal), May Alzamel and Costas Iliopoulos (King’s College, London, UK), Simona Rombo (University of Palermo, Italy).
- In Italy: Loukas Georgiadis (University of Ioannina, Greece), Matthias Mnich (University of Bonn, Germany), Adam Karczmarcz (University of Warsaw, Poland).
- In the Netherlands: Solon Pissis (King’s College, UK).

8.5.1.1. Internships

In 2018, ERABLE in France greeted the following Internships:

- Gabriela Paludo, Federal University of Mato Grosso do Sul, Brazil, from Dec 1st 2017 to June 30 2018, funds for 6 months from Capes, Brazil, and for the last month from ERABLE;
- Rafael Nahat, University of São Paulo, Brazil, from June 1st 2018 to Dec 15, 2018.

In the Netherlands, ERABLE greeted the following Internship: Luca Denti, University Bicocca of Milano, Italy, from October 1 to January 2019.

8.5.2. Visits to International Teams

8.5.2.1. Research Stays Abroad

In 2018, two members of ERABLE from France did research stays at Sapienza University of Rome. These were Marie-France Sagot and Blerina Sinaimeri, both funded by Sapienza, M.-F. Sagot as senior scientist for a visit of one month, and B. Sinaimeri as a junior scientist for a visit of three months. The visits took place at the beginning of 2018, Jan-Feb for M.-F. Sagot, and Jan-Apr for B. Sinaimeri. In the context of her visit to Sapienza, Blerina furthermore gave a mini-course (9h) at the University.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- Giuseppe Italiano is member of the Steering Committee of the *Workshop on Algorithm Engineering and Experimentation (ALENEX)*, of the *International Colloquium on Automata, Languages and Programming (ICALP)*, and the Workshop/Symposium on Experimental Algorithms (SEA).
- Alberto Marchetti-Spaccamela is a member of the Steering committee of *Workshop on Graph Theoretic Concepts in Computer Science (WG)*), and of *Workshop on Algorithmic Approaches for Transportation Modeling, Optimization, and Systems (ATMOS)*.
- Arnaud Mary is member of the Steering Committee of *Workshop on Enumeration Problems and Applications (WEPA)*.

- Marie-France Sagot is member of the Steering Committee of *European Conference on Computational Biology (ECCB)*, *International Symposium on Bioinformatics Research and Applications (ISBRA)*, and *Workshop on Enumeration Problems and Applications (WEPA)*.
- Alexander Schönhuth is member of the Steering committee of the *Research in Computational Molecular Biology, satellite conference on massively parallel sequencing (RECOMB-seq)*.

9.1.1.2. Member of the Organizing Committees

- Marie-France Sagot was co-organiser of the Dagstuhl Seminar on "Algorithmic Enumeration: Output-sensitive, Input-Sensitive, Parameterized, Approximative", Dagstuhl Schloss, Oct 14 to 19, 2018.
- Alexander Schönhuth was main organiser of main organizer of the Lorentz workshop on "Single Cell Data Science", Lorentz center, Leiden, June 4-8, 2018.
- Leen Stougie was member of the Organizing Committee of the *3rd Highlights of Algorithms (HALG 2018) conference* in the Free University of Amsterdam, June 4-6, 2018.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Giuseppe Italiano was a member of the Program Committee of *ISAAC*,*STOC*, *ICALP*, and *WALCOM* in 2018.
- Arnaud Mary was a member of the Program Committee of *CIAC* in 2018.
- Nadia Pisanti was a member of the Program Committee of *BIOINFORMATICS*, *FUN*, *ISBRA*, *BIOKDD*, *CIBB*, *Hi BI BI* in 2018.
- Marie-France Sagot was a member of the Program Committee of *SEA*, *CIAC*, *WABI*, *WEPA*, in 2018.
- Alexander Schönhuth was a member of the Program Committee of *RECOMB*, *ISMB*, *GCB* (German bioinformatics conference), and *ISMCO* (International Symposium on Mathematical and Computational Oncology) in 2018.

9.1.2.2. Reviewer

Members of ERABLE have reviewed papers for a number of workshops and conferences including: *CPM*, *ISMB*, *RECOMB*, *WEPA*, *WABI*.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Pierluigi Crescenzi is member of the Editorial Board of *Journal of Computer and Systems Science and Electronic Notes on Theoretical Computer Science*.
- Roberto Grossi is member of the Editorial Board of *Theory of Computing Systems (TOCS)* and pf *RAIRO – Theoretical Informatics and Applications*.
- Giuseppe Italiano is member of the Editorial Board of *Algorithmica* and *Theoretical Computer Science*.
- Alberto Marchetti-Spaccamela is member of the Editorial Board of *Theoretical Computer Science*.
- Arnaud Mary is Editor-in-Chief of a special issue of *Discrete Applied Mathematics* dedicated to *WEPA* 2016.
- Nadia Pisanti is since 2012 member of Editorial Board of *International Journal of Computer Science and Application (IJCSA)* and since 2017 of *Network Modeling Analysis in Health Informatics and Bioinformatics*.
- Marie-France Sagot is member of the Editorial Board of *BMC Bioinformatics*, *Algorithms for Molecular Biology*, and *Lecture Notes in BioInformatics*.

- Leen Stougie is member of the Editorial Board of *AIMS Journal of Industrial and Management Optimization*.
- Cristina Vieira is Executive Editor of *Gene*, and since 2014 member of the Editorial Board of *Mobile DNA*.

9.1.3.2. Reviewer - Reviewing Activities

Members of ERABLE have reviewed papers for a number of journals including: *Theoretical Computer Science*, *Algorithmica*, *Algorithms for Molecular Biology*, *Bioinformatics*, *BMC Bioinformatics*, *Genome Biology*, *Genome Research*, *IEEE/ACM Transactions in Computational Biology and Bioinformatics (TCBB)*, *Molecular Biology and Evolution*, *Nucleic Acid Research*.

9.1.4. Invited Talks

Giuseppe Italiano gave an Invited Talk on “2-Connectivity in Directed Graphs” at the *17th International Symposium on Experimental Algorithms (SEA 2018)* (http://cs.gssi.it/sea2018/, in L'Aquila, Italy, June 27–29, 2018.

Vincent Lacroix gave an Invited Presentation on “Traitement des données -omics” at the “9ème réunion annuelle Institut Thématique Multi-Organismes Technologies pour la Santé”, https://its.aviesan.fr/index.php?pagendx=1046.

Alexander Schönhuth gave Invited Talks at the CONTRA workshop, Warszaw (organisers: Jens Lagergren, Ewa Szczurek, Niko Beerenwinkel), in September and at Harvard Medical School, Boston, in October.

Leen Stougie gave an Invited Talk on “A decomposition theory for vertex enumeration of convex polyhedra” at the *Martin Dyer Day and Queen Mary Algorithms Day*, Queen Mary University, London, UK, July 16-17, 2018. He gave Invited Lectures on “Full-length de novo viral quasispecies assembly through variation graph construction” at the *4th Dutch Bioinformatics & Systems Biology Conference*, De Werelt, Lunteren, Netherlands, May 15; on “Polynomial time vertex enumeration of convex polytopes of bounded branch-width” at the Department of Mathematics, University of Bremen, Germany, Sept 11; and on “Algorithmic Problems in Biological Networks” at the *9th Networks Day*, Hortus Botanicus, Leiden, Netherlands, Sept 19, 2018.

Fabrice Vavre gave Invited Talks on “Impact du microbiote sur l'évolution et la diversité des insectes” at the *Colloque de la Société Francophone de Microbiologie*, Paris, Oct 1-3, 2018; and on “Evolution of host-microbiota interactions: what insects tell us about selection levels?” at the Journées de l'IXXI, Lyon, Oct 16, 2018.

9.1.5. Scientific Expertise

Giuseppe F. Italiano is member of the Council of the European Association for Theoretical Computer Science. Leen Stougie is member of the General Board of the Dutch Network on the Mathematics of Operations Research (Landelijk Netwerk Mathematische Besliskunde (LNMB)).

9.1.6. Research Administration

Hubert Charles is director of the Biosciences Department of the Insa-Lyon and co-director of studies of the “Bioinformatique et Modélisation (BIM)” track.

Giuseppe Italiano is member of the Advisory Board of MADALGO - Center for MAssive Data ALGOrithmics, Aarhus, Denmark.

Alberto Marchetti-Spaccamela was Director of the Department of Computer Engineering and Management Antonio Ruberti at Sapienza University from 2013 until end of May 2018.

Nadia Pisanti is since November 1st 2017 member of the Board of the PhD School in Data Science (University of Pisa jointly with Scuola Normale Superiore Pisa, Scuola S. Anna Pisa, IMT Lucca).

Marie-France Sagot is member of the Advisory Board of CWI, Amsterdam, the Netherlands, and chair of the CSS for MBIO at Inra.

Alexander Schönhuth is member of the Scientific Board of BioSB (the Dutch organisation for bioinformatics) since May 2017.

Leen Stougie is since April 2017 Leader of the Life Science Group at CWI.

Alain Viari was until Feb 2018 Deputy Scientific Director at Inria responsible for the ICST for Life and Environmental Sciences. He thus represented Inria at several national instances related to Life Sciences and Health (Allenvi, Aviesan, Ibisa, etc.). He is member of a number of scientific advisory boards (IRT (Institut de Recherche Technologique) BioAster; Centre Léon Bérard). He also coordinates together with J.-F. Deleuze (CNRGH-Evry) the Research & Development part (CRefIX) of the "Plan France Médecine Génomique 2025".

Fabrice Vavre is President of the Section 29 of the CoNRS and participated to the HCERES Evaluation Committee of the IAME Laboratory (UMR1137, Dir. E. Denamur), Paris, Mar 1-2, 2018.

Cristina Vieira is member of the "Conseil National des Universités" (CNU) 67 ("Biologie des Populations et Écologie"), and since 2017 member of the "Conseil de la Faculté des Sciences et Technologies (FST)" of the University Lyon 1.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

9.2.1.1. France

The members of ERABLE teach both at the Department of Biology of the University of Lyon (in particular within the BISM (BioInformatics, Statistics and Modelling) specialty, and at the department of Bioinformatics of the Insa (National Institute of Applied Sciences). Cristina Vieira is responsible for the Master Biodiversity, Ecology and Evolution (https://www.bee-lyon-univ.fr/). She teaches genetics 192 hours per year at the University and at the ENS-Lyon. Hubert Charles is responsible for the Master of Modelling and Bioinformatics (BIM) at the Insa of Lyon (http://biosciences.insa-lyon.fr/). He teaches 192 hours per year in statistics and biology. Vincent Lacroix is responsible for several courses of the Master in Bioinformatics (https://www.bioinfo-lyon.fr/) (L3: Advanced Bioinformatics, M1: Methods for Data Analysis in Genomics, M1: Methods for Data Analysis in Transcriptomics, M1: Bioinformatics Project, M2: Ethics). He teaches 192 hours per year in bioinformatics. Arnaud Mary is responsible for two courses of the Bioinformatics Curriculum at the University (L2: Introduction to Bioinformatics and Biostatistics, M1: Object Oriented Programming) and one at Insa (Discrete Mathematics). He taught 198 hours in 2018. Blerina Sinaimeri taught 36 hours in 2018 on graph algorithms for the M1 students of the Master in Bioinformatics, and on Discrete Mathematics at Insa. Fabrice Vavre taught 20h at the Master level.

The ERABLE team regularly welcomes M1 and M2 interns from the bioinformatics Master.

Vincent Lacroix was an instructor in NGS data analysis training for the CNRS Formation in the last 4 years since 2015, a course coordinated by Eric Rivals from the LIRMM, Montpellier (https://www.france-bioinformatique.fr/fr/formations/bioinformatique-pour-traitement-donn%C3%A9es-s%C3%A9quen%C3%A7age-ngs.

All French members of the ERABLE team are affiliated to the doctoral school E2M2 (Ecology-Evolution-Microbiology-Modelling, http://e2m2.universite-lyon.fr/).

9.2.1.2. Italy & The Netherlands

Italian researchers teach between 90 and 140 hours per year, at both the undergraduate and at the Master levels. The teaching involves pure computer science courses (such as Programming foundations, Programming in C or in Java, Computing Models, Distributed Algorithms) and computational biology (such as Algorithms for Bioinformatics).

Dutch researchers teach between 40 and 270 hours per year, again at the undergraduate and Master levels, in pure computer science (*e.g.* Algorithm Engineering, Randomised Algorithms), applied mathematics (*e.g.* Operational Research, Advanced Linear Programming) and computational biology (*e.g.* Biological Network Analysis).

9.2.2. Supervision

The following PhDs were defended in ERABLE in 2018:

- Damiano di Francesco Maesa, University of Pisa, co-supervisors: Laura Ricci and Andrea Marino, Oct 2018.
- Mattia Gastaldello, Sapienza University of Rome and University of Lyon 1 (funded by "Vinci Program-Université Franco-Italienne", co-supervisors: Tiziana Calamoneri, Sapienza University of Rome; Marie-France Sagot), Feb 2018.
- Scheila Gabriele Mucha, Federal University of Rio Grande do Sul, Brazil (PhD scholarship from Brazil with one year spent as "sandwich PhD" in France funded by Capes, and 6 more months funded by a grant from ERABLE, co-supervisors: Arnaldo Zaha, Federal University of Rio Grande do Sul; Marie-France Sagot), Sept 2018.

The following are the PhDs in progress:

- Marianne Borderes, University Lyon 1 (funded by ANR Technology Spock, co-supervisors: Susana Vinga – Instituto Superior Técnico at Lisbon; Marie-France Sagot)
- Audric Cologne, University of Lyon 1 (funded by Inserm and Inria, co-supervisors: Patrick Edery – Federation of Health Research of Lyon-Est, Vincent Lacroix)
- Leandro Ishi Soares de Lima, University of Lyon 1 (funded by the Brazilian "Science without Borders" program, co-supervisors: Giuseppe Italiano, Vincent Lacroix, Marie-France Sagot)
- Carol Moraga Quinteros, University of Lyon 1 (funded by Conicyt Chile, co-supervisors: Rodrigo Gutierrez – Catholic University of Chile, Marie-France Sagot)
- Henri Taneli Pusa, University of Lyon 1 (funded by H2020-MSCA-ETN-2014 project MicroWine, co-supervisors: Alberto Marchetti-Spaccamela, Arnaud Mary, Marie-France Sagot)
- Camille Sessegolo, University of Lyon 1 (funded by ANR Aster; co-supervisors: Vincent Lacroix, Arnaud Mary)
- Yishu Wang, University Lyon 1 (funded by Ministère de l?Enseignement supérieur, de la Recherche et de l?Innovation, co-supervisors: Mário Figueiredo – Instituto Superior Técnico at Lisbon; Marie-France Sagot; Blerina Sinaimeri)
- Irene Ziska, University Lyon 1 (funded by Inria Cordi-S, co-supervisors: Susana Vinga – Instituto Superior Técnico at Lisbon; Marie-France Sagot)

Besides the PhD students indicated above, who are physically located within one of the premises of ERABLE, the project-team has PhD students in co-supervision who spend the majority or the whole of their time in the premises of other teams. These include: Rita Ramos (funded by Portuguese FCT, co-supervisors: Claudia Nunes dos Santos – ITQB Lisbon, Marie-France Sagot), and André Veríssimo (funded by Portuguese FCT, co-supervisors: Susana Vinga – Instituto Superior Técnico in Lisbon, Marie-France Sagot).

9.2.3. Juries

The following are the PhD or HDR juries to which members of ERABLE participated in 2018.

- Giuseppe Italiano: External Examiner of the PhD of Daniel Wolleb-Graf, ETH, Zurich, Switzerland, Dec 3, 2018.
- Marie-France Sagot: External Reviewer of the PhD of Clémence Frioux, University of Rennes 1, France, Nov 19, 2018.
- Leen Stougie: Member of the Committee of the PhD of Bart Kamphorst, Technical University Eindhoven, Netherlands, May 2018; Member of the Committee of Gregorios Koumoutsos, Technical University Eindhoven, Netherlands, Sept 2018.
- Fabrice Vavre: External Reviewer of the HDR Committee of Thierry Lefevre, Univ of Montpellier, France, March 3, 2018.

9.3. Popularization

9.3.1. Interventions

Blerina Sinaimeri and Ricardo Andrade participated to the Inria Fête de la Science in October 2018. Blerina Sinaimeri and Mariana G. Ferrarini presented the research topics of the team in an event for welcoming international scientists at the Espace Ulys of the University of Lyon. The presentation was entitled “ERABLE : A bio-info symbiosis” and discussed mainly about the multidisciplinarity of the ERABLE Team. Finally, Mariana Galvao Ferrarini participated in the Fête de la Science INRA at the Mediathéque de Bron in October 2018.

Fabrice Vavre participated to the Forum des métiers, Collège Clément Marot, on “La recherche au CNRS”.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] U. Ashraf, C. Benoit-Pilven, V. Lacroix, V. Navratil, N. Naffakh.*Advances in Analyzing Virus-Induced Alterations of Host Cell Splicing*, in "Trends in Microbiology", December 2018, https://hal.inria.fr/hal-01964983

[2] L. Beaudou, P. Dankelmann, F. Foucaud, M. A. Henning, A. Mary, A. Parreau.*Bounding the Order of a Graph Using Its Diameter and Metric Dimension: A Study Through Tree Decompositions and VC Dimension*, in "Siam Journal on Discrete Mathematics", April 2018, vol. 32, n^{o} 2, p. 902 - 918 [*DOI :* 10.1137/16M1097833], https://hal.archives-ouvertes.fr/hal-01848648

[3] C. Benoit-Pilven, C. Marchet, E. Chautard, L. Lima, M.-P. Lambert, G. Sacomoto, A. Rey, A. Cologne, S. Terrone, L. Dulaurier, J.-B. Claude, C. Bourgeois, D. Auboeuf, V. Lacroix.*Complementarity of assembly-first and mapping-first approaches for alternative splicing annotation and differential analysis from RNAseq data*, in "Scientific Reports", December 2018, vol. 8, n^{o} 1 [*DOI :* 10.1038/s41598-018-21770-7], https://hal.inria.fr/hal-01924204

[4] S. Colella, N. Parisot, P. Simonet, K. Gaget, G. Duport, P. Baa-Puyoulet, Y. Rahbé, H. Charles, G. Febvay, P. Callaerts, F. Calevro.*Bacteriocyte Reprogramming to Cope With Nutritional Stress in a Phloem Sap Feeding Hemipteran, the Pea Aphid Acyrthosiphon pisum*, in "Frontiers in Physiology", 2018, vol. 9, 1498 [*DOI :* 10.3389/fphys.2018.01498], https://hal.archives-ouvertes.fr/hal-01905508

[5] D. Di Francesco Maesa, A. Marino, L. Ricci.*Data-driven analysis of Bitcoin properties: exploiting the users graph*, in "International Journal of Data Science and Analytics", August 2018, vol. 6, n^{o} 1, p. 63-80, https://hal.inria.fr/hal-01964718

[6] Y. Drouet, I. Treilleux, A. Viari, S. Léon, M. Devouassoux-Shisheboran, N. Voirin, C. de la Fouchardiere, B. Manship, A. Puisieux, C. Lasset, C. Moyret-Lalle.*Integrated analysis highlights APC11 protein expression as a likely new independent predictive marker for colorectal cancer*, in "Scientific Reports", December 2018, vol. 8, n^{o} 1 [*DOI :* 10.1038/s41598-018-25631-1], https://hal.inria.fr/hal-01920572

[7] O. DURON, P. DOUBLET, F. VAVRE, D. BOUCHON.*The Importance of Revisiting Legionellales Diversity*, in "Trends in Parasitology", December 2018, vol. 34, n^{o} 12, p. 1027 - 1037 [*DOI :* 10.1016/J.PT.2018.09.008], https://hal.archives-ouvertes.fr/hal-01933206

[8] O. DURON, O. MOREL, V. NOËL, M. BUYSSE, F. BINETRUY, R. LANCELOT, E. LOIRE, C. MÉNARD, O. BOUCHEZ, F. VAVRE, L. VIAL.*Tick-Bacteria Mutualism Depends on B Vitamin Synthesis Pathways*, in "Current Biology - CB", June 2018, vol. 28, n^{o} 12, p. 1896 - 1902.e5 [*DOI :* 10.1016/J.CUB.2018.04.038], https://hal.archives-ouvertes.fr/hal-01933362

[9] M. G. FERRARINI, S. G. MUCHA, D. PARROT, G. MEIFFREN, J. F. R. BACHEGA, G. COMTE, A. ZAHA, M.-F. SAGOT.*Hydrogen peroxide production and myo-inositol metabolism as important traits for virulence of Mycoplasma hyopneumoniae*, in "Molecular Microbiology", 2018, p. 1-45 [*DOI :* 10.1111/MMI.13957], https://hal.inria.fr/hal-01765751

[10] A. D. GENOVA, G. A. RUZ, M.-F. SAGOT, A. MAASS.*Fast-SG: an alignment-free algorithm for hybrid assembly*, in "GigaScience", May 2018, vol. 7, n^{o} 5, p. 1 - 15 [*DOI :* 10.1093/GIGASCIENCE/GIY048], https://hal.inria.fr/hal-01842462

[11] R. P. GROSSI, G. MENCONI, N. P. PISANTI, R. TRANI, S. VIND.*Motif trie: An efficient text index for pattern discovery with don't cares*, in "Theoretical Computer Science", February 2018, vol. 710, p. 74-87, https://hal.inria.fr/hal-01964717

[12] L. GUIO, C. VIEIRA, J. GONZÁLEZ.*Stress affects the epigenetic marks added by natural transposable element insertions in Drosophila melanogaster*, in "Scientific Reports", December 2018, vol. 8, n^{o} 1, https://hal.inria.fr/hal-01965035

[13] K. HUBER, V. MOULTON, M.-F. SAGOT, B. SINAIMERI.*Exploring and Visualising Spaces of Tree Reconciliations*, in "Systematic Biology", December 2018, https://hal.inria.fr/hal-01964682

[14] K. HUBER, V. MOULTON, M.-F. SAGOT, B. SINAIMERI.*Geometric medians in reconciliation spaces*, in "Information Processing Letters", 2018, vol. 136, p. 96 - 101 [*DOI :* 10.1016/J.IPL.2018.04.001], https://hal.inria.fr/hal-01842439

[15] S. HUET, B. TESSON, J.-P. JAIS, A. FELDMAN, L. MAGNANO, E. THOMAS, A. TRAVERSE-GLEHEN, B. ALBAUD, M. CARRÈRE, L. XERRI, S. ANSELL, L. BASEGGIO, C. REYES, K. TARTE, S. BOYAULT, C. HAIOUN, B. LINK, P. FEUGIER, A. LOPEZ-GUILLERMO, H. TILLY, P. BRICE, S. HAYETTE, F. JARDIN, F. OFFNER, P. SUJOBERT, D. GENTIEN, A. VIARI, E. CAMPO, J. CERHAN, G. SALLES.*A gene-expression profiling score for prediction of outcome in patients with follicular lymphoma: a retrospective training and validation analysis in three international cohorts*, in "Lancet Oncology", April 2018, vol. 19, n^{o} 4, p. 549 - 561, Erratum in: Correction to Lancet Oncol 2018; 19: 549-61. [Lancet Oncol. 2018] [*DOI :* 10.1016/S1470-2045(18)30102-5], https://hal.archives-ouvertes.fr/hal-01788996

[16] B. IBRAHIM, K. ARKHIPOVA, A. C. ANDEWEG, S. POSADA-CÉSPEDES, F. ENAULT, A. GRUBER, E. KOONIN, A. KUPCZOK, P. LEMEY, A. C. MCHARDY, D. P. MCMAHON, B. E. PICKETT, D. ROBERTSON, R. H. SCHEUERMANN, A. ZHERNAKOVA, M. P. ZWART, A. SCHÖNHUTH, B. E. DUTILH, M. MARZ.*Bioinformatics Meets Virology: The European Virus Bioinformatics Center's Second Annual Meeting*, in "Viruses", May 2018, vol. 10, n^{o} 5, 256 [*DOI :* 10.3390/V10050256], https://hal.inria.fr/hal-01964743

[17] M. JAILLARD, L. LIMA, M. TOURNOUD, P. MAHÉ, A. VAN BELKUM, V. LACROIX, L. JACOB.*A fast and agnostic method for bacterial genome-wide association studies: bridging the gap between k-mers and genetic events*, in "PLoS Genetics", 2018, p. 1-28 [*DOI :* 10.1371/JOURNAL.PGEN.1007758], https://hal.inria.fr/hal-01920359

[18] I. KYRKOU, T. PUSA, L. ELLEGAARD-JENSEN, M.-F. SAGOT, L. H. HANSEN.*Pierce's Disease of Grapevines: A Review of Control Strategies and an Outline of an Epidemiological Model*, in "Frontiers in Microbiology", September 2018, vol. 9, p. 1-23 [*DOI :* 10.3389/FMICB.2018.02141], https://hal.inria.fr/hal-01920527

[19] E. LERAT, C. GOUBERT, S. GUIRAO-RICO, M. MERENCIANO, A.-B. DUFOUR, C. VIEIRA, J. GONZÁLEZ.*Population specific dynamics and selection patterns of transposable element insertions in European natural populations*, in "Molecular Ecology", 2018, https://hal.inria.fr/hal-01965027

[20] G. A. MARAIS, J.-M. GAILLARD, C. VIEIRA, I. PLOTTON, D. SANLAVILLE, F. GUEYFFIER, J.-F. LEMAÎTRE.*Sex gap in aging and longevity: can sex chromosomes play a role?*, in "Biology of Sex Differences", December 2018, vol. 9, n^{o} 1, https://hal.inria.fr/hal-01965036

[21] T. MARSCHALL, M. MARZ, T. ABEEL, L. DIJKSTRA, B. E. DUTILH, A. GHAFFAARI, P. KERSEY, W. P. KLOOSTERMAN, V. MAKINEN, A. M. NOVAK, B. PATEN, D. PORUBSKY, E. RIVALS, C. ALKAN, J. A. BAAIJENS, P. I. W. DE BAKKER, V. BOEVA, R. J. P. BONNAL, F. CHIAROMONTE, R. CHIKHI, F. D. CICCARELLI, R. CIJVAT, E. DATEMA, C. M. V. DUIJN, E. E. EICHLER, C. ERNST, E. ESKIN, E. GARRISON, M. EL-KEBIR, G. W. KLAU, J. O. KORBEL, E.-W. LAMEIJER, B. LANGMEAD, M. MARTIN, P. MEDVEDEV, J. C. MU, P. NEERINCX, K. OUWENS, P. PETERLONGO, N. PISANTI, S. RAHMANN, B. RAPHAEL, K. REINERT, D. DE RIDDER, J. DE RIDDER, M. SCHLESNER, O. SCHULZ-TRIEGLAFF, A. D. SANDERS, S. SHEIKHIZADEH, C. SHNEIDER, S. SMIT, D. VALENZUELA, J. WANG, L. WESSELS, Y. ZHANG, V. GURYEV, F. VANDIN, K. YE, A. SCHÖNHUTH.*Computational pan-genomics: status, promises and challenges*, in "Briefings in Bioinformatics", 2018, vol. 19, n^{o} 1, p. 118-135 [*DOI :* 10.1093/BIB/BBW089], https://hal.inria.fr/hal-01390478

[22] B. MENEZES, J. SALCES-ORTIZ, H. MULLER, N. BURLET, S. MARTINEZ, M. FABLET, C. VIEIRA.*An attempt to select non-genetic variation in resistance to starvation and reduced chill coma recovery time in Drosophila melanogaster*, in "The Journal of Experimental Biology", December 2018, vol. 221, n^{o} 23, jeb186254, https://hal.inria.fr/hal-01965032

[23] A. MONTI, B. SINAIMERI.*On variants of Vertex Geography on undirected graphs*, in "Discrete Applied Mathematics", December 2018, vol. 251, p. 268-275, https://hal.inria.fr/hal-01965197

[24] D. P. MORTON, W. W. ROMEIJNDERS, R. R. SCHULTZ, L. STOUGIE.*The stochastic programming heritage of Maarten van der Vlerk*, in "Computational Management Science", October 2018, vol. 15, n^{o} 3-4, p. 319 - 323 [*DOI :* 10.1007/S10287-018-0335-8], https://hal.inria.fr/hal-01925979

[25] R. ROMBA, O. GNANKINE, S. F. DRABO, F. TIENDRÉBÉOGO, H. HENRI, L. MOUTON, F. VAVRE.*Abundance of Bemisia tabaci Gennadius (Hemiptera: Aleyrodidae) and its parasitoids on vegetables and cassava plants in Burkina Faso (West Africa)*, in "Ecology and Evolution", 2018, vol. 8, n^{o} 12, p. 6091-6103 [*DOI :* 10.1002/ECE3.4078], https://hal.archives-ouvertes.fr/hal-01916799

[26] P. SIMONET, K. GAGET, S. BALMAND, M. RIBEIRO LOPES, N. PARISOT, K. BUHLER, G. DUPORT, V. VULSTEKE, G. FEBVAY, A. HEDDI, H. CHARLES, P. CALLAERTS, F. CALEVRO.*Bacteriocyte cell death in*

the pea aphid/ Buchnera symbiotic system, in "Proceedings of the National Academy of Sciences of the United States of America ", February 2018, vol. 115, n^o 8, p. E1819 - E1828 [*DOI :* 10.1073/PNAS.1720237115], https://hal.archives-ouvertes.fr/hal-01768487

[27] F. THOMAS, F. VAVRE, T. TISSOT, M. VITTECOQ, M. GIRAUDEAU, F. F. BERNEX, D. MISSÉ, F. RENAUD, N. RAVEN, C. BECKMANN, R. HAMEDE, P. BIRO, B. UJVARI.*Cancer Is Not (Only) a Senescence Problem*, in "Trends in Cancer", March 2018, vol. 4, n^o 3, p. 169 - 172 [*DOI :* 10.1016/J.TRECAN.2018.01.002], https://hal.archives-ouvertes.fr/hal-01933408

[28] L. TONON, G. FROMONT, S. BOYAULT, E. THOMAS, A. FERRARI, A.-S. SERTIER, J. KIELBASSA, V. LE TEXIER, A. KAMOUN, N. ELAROUCI, J. IRANI, L. MULTIGNER, I. I. GUT, M. GUT, P. BLANCHET, A. DE REYNIÈS, G. CANCEL-TASSIN, A. VIARI, O. CUSSENOT.*Mutational Profile of Aggressive, Localised Prostate Cancer from African Caribbean Men Versus European Ancestry Men*, in "European Urology", September 2018, vol. 75, n^o 1, p. 11-15 [*DOI :* 10.1016/J.EURURO.2018.08.026], https://hal.inria.fr/hal-01921597

[29] L. URBINI, B. SINAIMERI, C. MATIAS, M.-F. SAGOT.*Exploring the Robustness of the Parsimonious Reconciliation Method in Host-Symbiont Cophylogeny*, in "IEEE/ACM Transactions on Computational Biology and Bioinformatics", May 2018, p. 1-11 [*DOI :* 10.1109/TCBB.2018.2838667], https://hal.inria.fr/hal-01842451

[30] L. VAN IERSEL, S. KELK, G. STAMOULIS, L. STOUGIE, O. BOES.*On Unrooted and Root-Uncertain Variants of Several Well-Known Phylogenetic Network Problems*, in "Algorithmica", 2018, vol. 64, n^o 4, p. 621 - 637 [*DOI :* 10.1093/SYSBIO/SYV020], https://hal.inria.fr/hal-01599716

[31] G. L. WALLAU, C. VIEIRA, É. L. S. LORETO.*Genetic exchange in eukaryotes through horizontal transfer: connected by the mobilome*, in "Mobile DNA", December 2018, vol. 9, n^o 1, https://hal.inria.fr/hal-01965038

International Conferences with Proceedings

[32] M. ALZAMEL, L. A. K. AYAD, G. BERNARDINI, R. P. GROSSI, C. S. ILIOPOULOS, N. P. PISANTI, S. P. PISSIS, G. ROSONE.*Degenerate String Comparison and Applications*, in "Workshop on Algorithms in Bioinformatics (WABI)", Helsinki, Finland, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018, vol. 113 [*DOI :* 10.4230/LIPICS.WABI.2018.21], https://hal.inria.fr/hal-01964706

[33] L. A. K. AYAD, G. BERNARDINI, R. P. GROSSI, C. S. ILIOPOULOS, N. PISANTI, S. P. PISSIS, G. ROSONE.*Longest Property-Preserved Common Factor*, in "International Symposium on String Processing and Information Retrieval", Lima, Peru, 2018, https://hal.inria.fr/hal-01921603

[34] J. A. BAAIJENS, B. VAN DER ROEST, J. KOESTER, L. STOUGIE, A. SCHOENHUTH.*Full-length de novo viral quasispecies assembly through variation graph construction*, in "RECOMB Satellite on Massively Parallel Sequencing (RECOMB-Seq)", Paris, France, 2018 [*DOI :* 10.1101/287177], https://hal.inria.fr/hal-01964730

[35] S. BHATTACHARYA, M. HENZINGER, G. F. ITALIANO.*Design of Dynamic Algorithms via Primal-Dual Method*, in "International Colloquium on Automata, Languages, and Programming (ICALP)", Kyoto, Japan, Lecture Notes in Computer Science, 2018, vol. 9134, p. 206-218, https://arxiv.org/abs/1604.05337 , https://hal.inria.fr/hal-01964700

[36] P. BILLE, M. BERGGREEN ETTIENNE, R. P. GROSSI, I. L. GØRTZ, E. ROTENBERG.*Compressed Communication Complexity of Longest Common Prefixes*, in "International Symposium on String Processing and Information Retrieval (SPIRE)", Lima, Peru, Lecture Notes in Computer Science, 2018, vol. 11147, p. 74-87, https://hal.inria.fr/hal-01964719

[37] T. BOSMAN, M. VAN EE, Y. JIAO, A. MARCHETTI-SPACCAMELA, R. RAVI, L. STOUGIE.*Approximation Algorithms for Replenishment Problems with Fixed Turnover Times*, in "Latin American Symposium on Theoretical Informatics (LATIN)", Buenos Aires, Argentina, Lecture Notes in Computer Science, 2018, vol. 10807, p. 217-230, https://arxiv.org/abs/1712.05218 , https://hal.inria.fr/hal-01964685

[38] G. BOTTAZZI, G. F. ITALIANO, D. SPERA.*Preventing Ransomware Attacks Through File System Filter Drivers*, in "Second Italian Conference on Cyber Security", Milan, Italy, 2018, https://hal.inria.fr/hal-01925958

[39] N. BOUSQUET, A. MARY.*Reconfiguration of graphs with connectivity constraints*, in "International Workshop on Approximation and Online Algorithms (WAOA)", Helsinki, Finland, Lecture Notes in Computer Science, 2018, vol. 11312, p. 295-309, https://arxiv.org/abs/1809.05443 , https://hal.inria.fr/hal-01964723

[40] A. CONTE, T. DE MATTEIS, D. DE SENSI, R. P. GROSSI, A. MARINO, L. P. VERSARI.*D2K: Scalable Community Detection in Massive Networks via Small-Diameter k-Plexes*, in "ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)", London, United Kingdom, 2018, p. 1272-1281, https://hal.inria.fr/hal-01964688

[41] A. CONTE, G. FERRARO, R. P. GROSSI, A. MARINO, K. SADAKANE, T. UNO.*Node Similarity with q -Grams for Real-World Labeled Networks*, in "ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)", London, United Kingdom, Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2018, p. 1282-1291 [*DOI :* 10.1145/3219819.3220085], https://hal.inria.fr/hal-01964716

[42] A. CONTE, R. P. GROSSI, A. MARINO, R. RIZZI, T. UNO, L. P. VERSARI.*Tight Lower Bounds for the Number of Inclusion-Minimal st-Cuts*, in "International Workshop on Graph-Theoretic Concepts in Computer Science (WG)", Cottbus, Germany, Lecture Notes in Computer Science, 2018, vol. 11159, p. 100-110, https://hal.inria.fr/hal-01964710

[43] A. CONTE, R. P. GROSSI, A. MARINO, R. RIZZI, L. P. VERSARI.*Listing Subgraphs by Cartesian Decomposition*, in "International Symposium on Mathematical Foundations of Computer Science (MFCS)", Dagstuhl, Germany, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018, vol. 117, p. 1868-8969 [*DOI :* 10.4230/LIPIcs.MFCS.2018.84], https://hal.inria.fr/hal-01964703

[44] A. CONTE, R. P. GROSSI, A. MARINO, L. P. VERSARI.*Finding Maximal Common Subgraphs via Time-Space Efficient Reverse Search*, in "International Computing and Combinatorics Conference (COCOON)", Qing Dao, China, Lecture Notes in Computer Science, 2018, vol. 10976, p. 328-340, https://hal.inria.fr/hal-01964709

[45] L. GEORGIADIS, G. F. ITALIANO, N. PAROTSIDIS.*Incremental Strong Connectivity and 2-Connectivity in Directed Graphs*, in "Latin American Symposium on Theoretical Informatics", Buenos Aires, Argentina, 2018, https://arxiv.org/abs/1802.10189 , https://hal.inria.fr/hal-01925953

[46] R. P. GROSSI, A. MARINO, L. P. VERSARI.*Efficient Algorithms for Listing k Disjoint st-Paths in Graphs*, in "Latin American Symposium on Theoretical Informatics (LATIN)", Buenos Aires, Argentina, Lecture Notes in Computer Science, 2018, vol. 10807, p. 544-557, https://hal.inria.fr/hal-01964712

[47] R. P. GROSSI, L. P. VERSARI.*Round-Hashing for Data Storage: Distributed Servers and External-Memory Tables*, in "European Symposium on Algorithms (ESA)", Helsinki, Finland, Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018, vol. 43, p. 1 - 43 [*DOI :* 10.4230/LIPIcs.ESA.2018.43], https://hal.inria.fr/hal-01964721

[48] J. HOLM, G. F. ITALIANO, A. KARCZMARZ, J. ŁĄCKI, E. ROTENBERG.*Decremental SPQR-trees for Planar Graphs*, in "European Symposium on Algorithms", Helsinki, Finland, 2018, https://arxiv.org/abs/1806.10772 , https://hal.inria.fr/hal-01925961

[49] G. F. ITALIANO, Y. MANOUSSAKIS, K. T. NGUYEN, H.-P. PHAM.*Maximum colorful cliques in vertex-colored graphs*, in "COCOON 2018 - 24th International Computing and Combinatorics Conference", Qingdao, China, LNCS, Springer, July 2018, vol. 10976, p. 480-491 [*DOI :* 10.1007/978-3-319-94776-1_40], https://hal.archives-ouvertes.fr/hal-01827893

[50] N. OLVER, K. PRUHS, K. SCHEWIOR, R. SITTERS, L. STOUGIE.*The Itinerant List Update Problem*, in "Workshop on Approximation and Online Algorithms (WAOA)", Helsinki, Finland, 2018, p. 310-326, https://hal.inria.fr/hal-01972562

[51] N. PREZZA, N. P. PISANTI, M. SCIORTINO, G. ROSONE.*Detecting Mutations by eBWT*, in "Workshop on Algorithms in Bioinformatics", Helsinki, Finland, 2018, https://arxiv.org/abs/1805.01876 [*DOI :* 10.4230/LIPIcs.CVIT.2016.23], https://hal.inria.fr/hal-01925950

[52] B. YIN, M. BALVERT, D. ZAMBRANO, A. SCHÖNHUTH, S. BOHTE.*An image representation based convolutional network for DNA classification*, in "International Conference on Learning Representations (ICLR)", Vancouver, Canada, 2018, https://arxiv.org/abs/1806.04931 - Published at ICLR 2018, https://openreview.net/pdf?id=HJvvRoe0W, https://hal.inria.fr/hal-01964736

Conferences without Proceedings

[53] G. F. ITALIANO, Y. MANOUSSAKIS, N. KIM THANG, H.-P. PHAM.*Maximum Colorful Cycles in Vertex-Colored Graphs*, in "CSR 2018 - 13th International Computer Science Symposium in Russia", Moscow, Russia, Lecture Notes in Computer Science, Springer, June 2018, vol. 10846, p. 106-117 [*DOI :* 10.1007/978-3-319-90530-3_10], https://hal.archives-ouvertes.fr/hal-01762581

Other Publications

[54] R. CHIKHI, A. SCHÖNHUTH.*Dualities in Tree Representations*, November 2018, https://arxiv.org/abs/1804.04263 - CPM 2018, extended version, https://hal.archives-ouvertes.fr/hal-01935566

[55] C. MARCHET, L. LIMA.*Simulation of RNA sequencIng with Oxford Nanopore Technologies*, July 2018, p. 1-2, JOBIM 2018 - Journées Ouvertes Biologie, Informatique et Mathématiques, Poster, https://hal.archives-ouvertes.fr/hal-01929917

Project-Team IBIS

Modeling, simulation, measurement, and control of bacterial regulatory networks

IN PARTNERSHIP WITH:
Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Computational Biology

Table of contents

Project-Team IBIS

Creation of the Project-Team: 2009 January 01

Keywords:

Computer Science and Digital Science:

A3.1.1. - Modeling, representation
A3.4.5. - Bayesian methods
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.4. - Statistical methods
A6.3.1. - Inverse problems
A6.3.2. - Data assimilation
A6.3.3. - Data processing
A6.4.1. - Deterministic control

Other Research Topics and Application Domains:

B1. - Life sciences
B1.1.2. - Molecular and cellular biology
B1.1.4. - Genetics and genomics
B1.1.7. - Bioinformatics
B1.1.8. - Mathematical biology
B1.1.10. - Systems and synthetic biology
B4.3.1. - Biofuels

1. Team, Visitors, External Collaborators

Research Scientists

Hidde de Jong [Team leader, Inria, Senior Researcher, HDR]
Eugenio Cinquemani [Inria, Researcher]
Delphine Ropers [Inria, Researcher]

Faculty Members

Johannes Geiselmann [Univ Grenoble Alpes, Professor]
Stephan Lacour [Univ Grenoble Alpes, Associate Professor]
Yves Markowicz [Univ Grenoble Alpes, Associate Professor]
Michel Page [Univ Grenoble Alpes, Associate Professor]

Technical Staff

Tamas Muszbek [Inria]
Célia Boyat [Univ Grenoble Alpes]
Corinne Pinel [CNRS]

PhD Students

Thibault Etienne [INRA]
Antrea Pavlou [Inria, from Nov 2018]

Post-Doctoral Fellows

Aline Marguet [Inria]

Marco Mauri [Inria]

Administrative Assistant

Alexandra Fitzgerald [Inria]

2. Overall Objectives

2.1. Overview

When confronted with changing environmental conditions, bacteria and other microorganisms have a remarkable capacity to adapt their functioning. The responses of bacteria to changes in their environment are controlled on the molecular level by large and complex networks of biochemical interactions involving genes, mRNAs, proteins, and metabolites. The study of bacterial regulatory networks requires experimental tools for mapping the interaction structure of the networks and measuring the dynamics of cellular processes. In addition, when dealing with such large and complex systems, we need mathematical modeling and computer simulation to integrate available biological data, and understand and predict the dynamics of the system under various physiological and genetic perturbations. The analysis of living systems through the combined application of experimental and computational methods has gathered momentum in recent years under the name of systems biology.

The first aim of the IBIS project-team is to apply such a systems-biology approach to gain a deeper understanding, on the mechanistic level, of the strategies that bacteria have developed to respond to changes in their environment. [0] In particular, we focus on the enterobacterium *Escherichia coli*, for which enormous amounts of genomic, genetic, biochemical and physiological data have accumulated over the past decades. A better understanding of the adaptive capabilities of *E. coli* to nutritional limitations or other environmental changes is an aim in itself, but also a necessary prerequisite for the second and most ambitious aim of the project: interfering with the cellular responses by specific perturbations or by rewiring the underlying regulatory networks. This does not only spawn fundamental research on the control of living matter, but may ultimately also lead to practical applications. Because *E. coli* is easy to manipulate in the laboratory, it serves as a model for many pathogenic bacteria and is widely used in biotechnology, for such diverse applications as the development of vaccines, the mass production of enzymes and other (heterologous) proteins, and the production of biofuels.

The aims of IBIS raise new questions on the interface of biology, applied mathematics, and computer science. In particular, the following objectives have structured the work of the project-team: (1) the analysis of the qualitative dynamics of gene regulatory networks, (2) the inference of gene regulatory networks from time-series data, (3) the analysis of integrated metabolic and regulatory networks, and (4) natural and engineered control of regulatory networks. Although these axes cover most of the work carried out in IBIS, some members have also made contributions to research projects on different topics. Since this usually represents a minor proportion of the overall research effort of the project-team, we will not describe this work in detail in the activity report. The publications resulting from these side-tracks have been included in the bibliography.

The challenges of the research programme of the IBIS team require a wide range of competences on the interface of (experimental) biology, applied mathematics, and computer science (Figure 1). Since no single person can be expected to possess all of these competences, the international trend in systems biology is to join researchers from different disciplines into a single group. In line with this development, the IBIS team is a merger of a microbiology and molecular genetics group on the one hand, and a bioinformatics and mathematical biology group on the other hand. In particular, the IBIS team is composed of members of the group of Johannes Geiselmann, formerly at the Laboratoire Adaptation et Pathogénicité des Microorganismes of the Univ Joseph Fourier (UJF, CNRS UMR 5163), and since September 2014 at the Laboratoire Interdisciplinaire de Physique (CNRS UMR 5588), and the members of the network modeling and simulation group formerly

[0] The ibis was an object of religious veneration in ancient Egypt, particularly associated with the god Thoth. Thoth was seen, among other things, as a god of the measurement and regulation of events.

part of the HELIX project-team at Inria Grenoble - Rhône-Alpes, a group coordinated by Hidde de Jong. The two groups have established a fruitful collaboration, which has resulted in more than 60 peer-reviewed publications in journals, conferences, and books since 2000. [0]

Hidde de Jong is the head of the IBIS project-team and Johannes Geiselmann its co-director. The experimental component of IBIS is also part of the Laboratoire Interdisciplinaire de Physique, and Johannes Geiselmann continues to represent this group in the interactions with the laboratory and university administration.

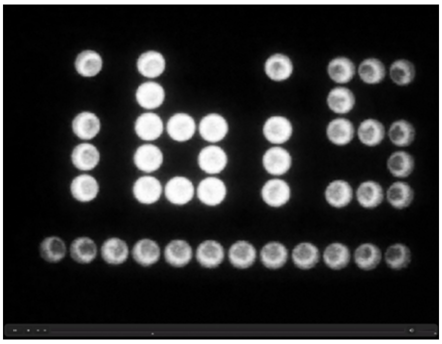


Figure 1. Display of the project-team name on a "bacterial billboard" (see http://team.inria.fr/ibis for the corresponding movie). A microplate containing a minimal medium (with glucose and acetate) is filmed during 36 hours. Wells contain E. coli bacteria which are transformed with a reporter plasmid carrying the luciferase operon (luxCDABE) under control of the acs promoter. This promoter is positively regulated by the CRP-cAMP complex. When bacteria have metabolized all the glucose, the cAMP concentration increases quickly and activates the global regulator CRP which turns on the transcription of the luciferase operon producing the light. The glucose concentration increases from left to right on the microplate, so its consumption takes more time when going up the gradient and the letters appear one after the other. The luciferase protein needs reductive power ($FMNH_2$) to produce light. At the end, when acetate has been depleted, there is no carbon source left in the medium. As a consequence, the reductive power falls and the bacterial billboard switches off. Source: Guillaume Baptist.

3. Research Program

3.1. Analysis of qualitative dynamics of gene regulatory networks

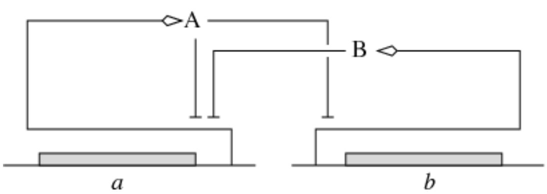
Participants: Hidde de Jong [Correspondent], Michel Page, Delphine Ropers.

The dynamics of gene regulatory networks can be modeled by means of ordinary differential equations (ODEs), describing the rate of synthesis and degradation of the gene products as well as regulatory interactions between gene products and metabolites. In practice, such models are not easy to construct though, as the parameters are often only constrained to within a range spanning several orders of magnitude for most systems of biological interest. Moreover, the models usually consist of a large number of variables, are strongly nonlinear, and include different time-scales, which makes them difficult to handle both mathematically and computationally. This has motivated the interest in qualitative models which, from incomplete knowledge of the system, are able to provide a coarse-grained picture of its dynamics.

[0] See http://team.inria.fr/ibis for a complete list.

A variety of qualitative modeling formalisms have been introduced over the past decades. Boolean or logical models, which describe gene regulatory and signalling networks as discrete-time finite-state transition systems, are probably most widely used. The dynamics of these systems are governed by logical functions representing the regulatory interactions between the genes and other components of the system. IBIS has focused on a related, hybrid formalism that embeds the logical functions describing regulatory interactions into an ODE formalism, giving rise to so-called piecewise-linear differential equations (PLDEs, Figure 2). The use of logical functions allows the qualitative dynamics of the PLDE models to be analyzed, even in high-dimensional systems. In particular, the qualitative dynamics can be represented by means of a so-called state transition graph, where the states correspond to (hyperrectangular) regions in the state space and transitions between states arise from solutions entering one region from another.

First proposed by Leon Glass and Stuart Kauffman in the early seventies, the mathematical analysis of PLDE models has been the subject of active research for more than four decades. IBIS has made contributions on the mathematical level, in collaboration with the BIOCORE and BIPOP project-teams, notably for solving problems induced by discontinuities in the dynamics of the system at the boundaries between regions, where the logical functions may abruptly switch from one discrete value to another, corresponding to the (in)activation of a gene. In addition, many efforts have gone into the development of the computer tool GENETIC NETWORK ANALYZER (GNA) and its applications to the analysis of the qualitative dynamics of a variety of regulatory networks in microorganisms. Some of the methodological work underlying GNA, notably the development of analysis tools based on temporal logics and model checking, which was carried out with the Inria project-teams CONVEX (ex-VASY) and POP-ART, has implications beyond PLDE models as they apply to logical and other qualitative models as well.



$$\dot{x}_a = \kappa_a\, s^-(x_a, \theta_a^2)\, s^-(x_b, \theta_b) - \gamma_a\, x_a$$

$$\dot{x}_b = \kappa_b\, s^-(x_a, \theta_a^1) - \gamma_b\, x_b$$

$$s^+(x, \theta) = \begin{cases} 1, & \text{if } x > \theta \\ 0, & \text{if } x < \theta \end{cases}$$

$$s^-(x, \theta) = 1 - s^+(x, \theta)$$

Figure 2. (Left) Example of a gene regulatory network of two genes (a and b), each of which codes for a regulatory protein (A and B). Protein B inhibits the expression of gene a, while protein A inhibits the expression of gene b and its own gene. (Right) PLDE model corresponding to the network in (a). Protein A is synthesized at a rate κ_a, if and only if the concentration of protein A is below its threshold θ_a^2 ($x_a < \theta_a^2$) and the concentration of protein B below its threshold θ_b ($x_b < \theta_b$). The degradation of protein A occurs at a rate proportional to the concentration of the protein itself ($\gamma_a\, x_a$).

3.2. Inference of gene regulatory networks from time-series data

Participants: Eugenio Cinquemani [Correspondent], Johannes Geiselmann, Hidde de Jong, Stephan Lacour, Aline Marguet, Michel Page, Corinne Pinel, Delphine Ropers.

Measurements of the transcriptome of a bacterial cell by means of DNA microarrays, RNA sequencing, and other technologies have yielded huge amounts of data on the state of the transcriptional program in different growth conditions and genetic backgrounds, across different time-points in an experiment. The information on the time-varying state of the cell thus obtained has fueled the development of methods for inferring regulatory interactions between genes. In essence, these methods try to explain the observed variation in the activity of one gene in terms of the variation in activity of other genes. A large number of inference methods have been proposed in the literature and have been successful in a variety of applications, although a number of difficult problems remain.

Current reporter gene technologies, based on Green Fluorescent Proteins (GFPs) and other fluorescent and luminescent reporter proteins, provide an excellent means to measure the activity of a gene *in vivo* and in real time (Figure 3). The underlying principle of the technology is to fuse the promoter region and possibly (part of) the coding region of a gene of interest to a reporter gene. The expression of the reporter gene generates a visible signal (fluorescence or luminescence) that is easy to capture and reflects the expression of a gene of interest. The interest of the reporter systems is further enhanced when they are applied in mutant strains or combined with expression vectors that allow the controlled induction of any particular gene, or the degradation of its product, at a precise moment during the time-course of the experiment. This makes it possible to perturb the network dynamics in a variety of ways, thus obtaining precious information for network inference.

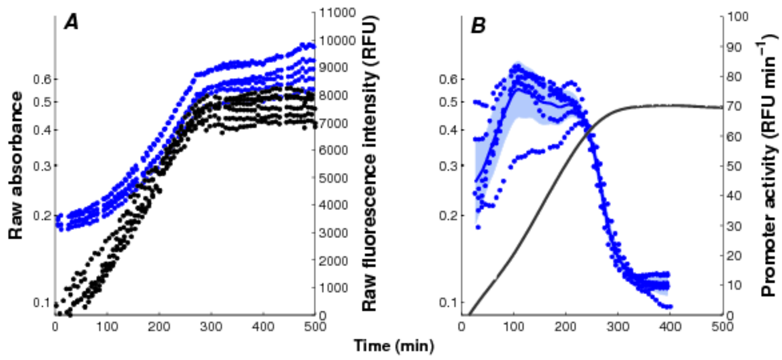


Figure 3. Monitoring of bacterial gene expression in vivo using fluorescent reporter genes (Stefan et al., PLoS Computational Biology, 11(1):e1004028, 2015). The plots show the primary data obtained in a kinetic experiment with E. coli cells, focusing on the expression of the motility gene tar in a mutant background. A: Absorbance (•, black) and fluorescence (•, blue) data, corrected for background intensities, obtained with the ΔcpxR strain transformed with the ptar-gfp reporter plasmid and grown in M9 with glucose. B: Activity of the tar promoter, computed from the primary data. The solid black line corresponds to the mean of 6 replicate absorbance measurements and the shaded blue region to the mean of the promoter activities $\pm$ twice the standard error of the mean.

The specific niche of IBIS in the field of network inference has been the development and application of genome engineering techniques for constructing the reporter and perturbation systems described above, as well as the use of reporter gene data for the reconstruction of gene regulation functions. We have developed an experimental pipeline that resolves most technical difficulties in the generation of reproducible time-series measurements on the population level. The pipeline comes with data analysis software that converts the primary data into measurements of time-varying promoter activities. In addition, for measuring gene expression on the single-cell level by means of microfluidics and time-lapse fluorescence microscopy, we have established collaborations with groups in Grenoble and Paris. The data thus obtained can be exploited for the structural and parametric identification of gene regulatory networks, for which methods with a solid mathematical foundation are developed, in collaboration with colleagues at ETH Zürich and EPF Lausanne (Switzerland). The vertical integration of the network inference process, from the construction of the biological material to the data analysis and inference methods, has the advantage that it allows the experimental design to be precisely tuned to the identification requirements.

3.3. Analysis of integrated metabolic and gene regulatory networks

Participants: Eugenio Cinquemani, Hidde de Jong, Thibault Etienne, Johannes Geiselmann, Stephan Lacour, Yves Markowicz, Marco Mauri, Michel Page, Corinne Pinel, Delphine Ropers [Correspondent].

The response of bacteria to changes in their environment involves responses on several different levels, from the redistribution of metabolic fluxes and the adjustment of metabolic pools to changes in gene expression. In order to fully understand the mechanisms driving the adaptive response of bacteria, as mentioned above, we need to analyze the interactions between metabolism and gene expression. While often studied in isolation, gene regulatory networks and metabolic networks are closely intertwined. Genes code for enzymes which control metabolic fluxes, while the accumulation or depletion of metabolites may affect the activity of transcription factors and thus the expression of enzyme-encoding genes.

The fundamental principles underlying the interactions between gene expressions and metabolism are far from being understood today. From a biological point of view, the problem is quite challenging, as metabolism and gene expression are dynamic processes evolving on different time-scales and governed by different types of kinetics. Moreover, gene expression and metabolism are measured by different experimental methods generating heterogeneous, and often noisy and incomplete data sets. From a modeling point of view, difficult methodological problems concerned with the reduction and calibration of complex nonlinear models need to be addressed.

Most of the work carried out within the IBIS project-team specifically addressed the analysis of integrated metabolic and gene regulatory networks in the context of *E. coli* carbon metabolism (Figure 4). While an enormous amount of data has accumulated on this model system, the complexity of the regulatory mechanisms and the difficulty to precisely control experimental conditions during growth transitions leave many essential questions open, such as the physiological role and the relative importance of mechanisms on different levels of regulation (transcription factors, metabolic effectors, global physiological parameters, ...). We are interested in the elaboration of novel biological concepts and accompanying mathematical methods to grasp the nature of the interactions between metabolism and gene expression, and thus better understand the overall functioning of the system. Moreover, we have worked on the development of methods for solving what is probably the hardest problem when quantifying the interactions between metabolism and gene expression: the estimation of parameters from hetereogeneous and noisy high-throughput data. These problems are tackled in collaboration with experimental groups at Inra/INSA Toulouse and CEA Grenoble, which have complementary experimental competences (proteomics, metabolomics) and biological expertise.

3.4. Natural and engineered control of growth and gene expression

Participants: Célia Boyat, Eugenio Cinquemani, Johannes Geiselmann [Correspondent], Hidde de Jong [Correspondent], Stephan Lacour, Marco Mauri, Tamas Muszbek, Michel Page, Antrea Pavlou, Delphine Ropers.

The adaptation of bacterial physiology to changes in the environment, involving changes in the growth rate and a reorganization of gene expression, is fundamentally a resource allocation problem. It notably poses the question how microorganisms redistribute their protein synthesis capacity over different cellular functions when confronted with an environmental challenge. Assuming that resource allocation in microorganisms has been optimized through evolution, for example to allow maximal growth in a variety of environments, this question can be fruitfully formulated as an optimal control problem. We have developed such an optimal control perspective, focusing on the dynamical adaptation of growth and gene expression in response to envrionmental changes, in close collaboration with the BIOCORE project-team.

A complementary perspective consists in the use of control-theoretical approaches to modify the functioning of a bacterial cell towards a user-defined objective, by rewiring and selectively perturbing its regulatory networks. The question how regulatory networks in microorganisms can be externally controlled using engineering approaches has a long history in biotechnology and is receiving much attention in the emerging field of synthetic biology. Within a number of on-going projects, IBIS is focusing on two different questions. The first concerns the development of open-loop and closed-loop growth-rate controllers of bacterial cells for both fundamental research and biotechnological applications (Figure 5). Second, we are working on the development of methods for the real-time control of the expression of heterologous proteins in communities of interacting bacterial populations. The above projects involve collaborations with, among others, the Inria

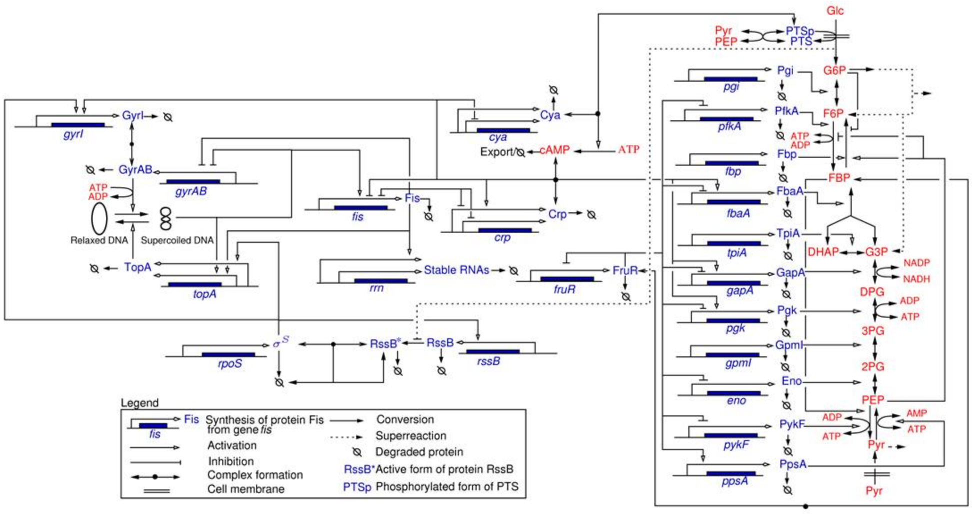


Figure 4. Network of key genes, proteins, and regulatory interactions involved in the carbon assimilation network in E. coli (Baldazzi et al., PLoS Computational Biology, 6(6):e1000812, 2010). The metabolic part includes the glycolysis/gluconeogenesis pathways as well as a simplified description of the PTS system, via the phosphorylated and non-phosphorylated form of its enzymes (represented by PTSp and PTS, respectively). The pentose-phosphate pathway (PPP) is not explicitly described but we take into account that a small pool of G6P escapes the upper part of glycolysis. At the level of the global regulators the network includes the control of the DNA supercoiling level, the accumulation of the sigma factor RpoS and the Crp·cAMP complex, and the regulatory role exerted by the fructose repressor FruR.

project-teams LIFEWARE (INBIO), BIOCORE, and McTAO as well as with a biophysics group at Univ Paris Descartes and a mathematical modeling group at INRA Jouy-en-Josas.

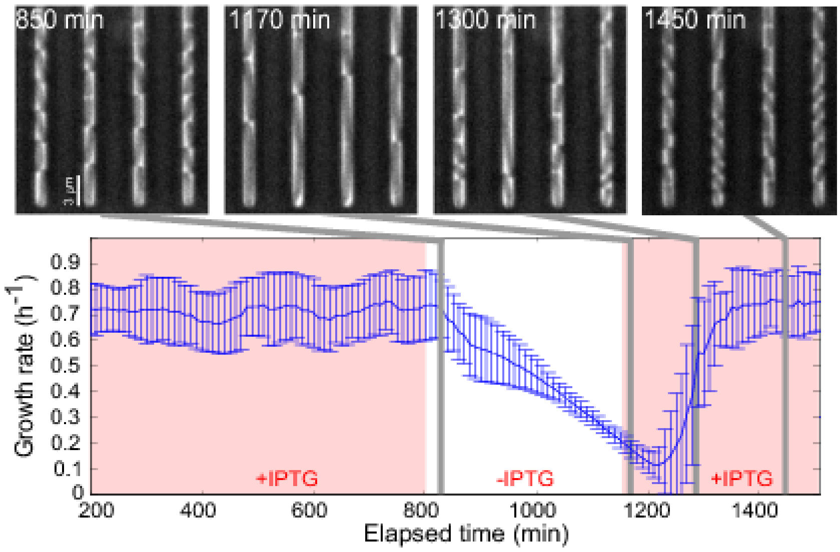


Figure 5. Growth arrest by external control of the gene expression machinery (Izard, Gomez Balderas et al., Molecular Systems Biology, 11:840, 2015). An E. coli strain in which an essential component of the gene expression machinery, the $\beta\beta'$ subunits of RNA polymerase, was put under the control of an externally-supplied inducer (IPTG), was grown in a microfluidics device and phase-contrast images were acquired every 10 min. The cells were grown in minimal medium with glucose, initially in the presence of 1 mM IPTG. 6 h after removing IPTG from the medium, the growth rate slows down and cells are elongated. About 100 min after adding back 1 mM IPTG into the medium, the elongated cells divide and resume normal growth. The growth rates in the plot are the (weighted) mean of the growth rates of 100 individual cells. The error bars correspond to $\pm$ one standard deviation. The results of the experiment show that the growth rate of a bacterial can be switched off in a reversible manner by an external inducer, based on the reengineering of the natural control of the expression of RNA polymerase.

4. Highlights of the Year

4.1. Highlights of the Year

Two new projects with participation from IBIS started this year: the ANR project RIBECO and the IXXI project MuSE (Section 7.2). The web application WellInverter was made available through the new cloud of the French Institute of Bioinformatics (IFB) (Section 5.2). A publication summarizing several conference contributions on the stochastic modeling and inference of gene regulatory networks was published in the main control journal *Automatica* this year [16].

5. New Software and Platforms

5.1. WellFARE

KEYWORDS: Bioinformatics - Statistics - Data visualization - Data modeling

SCIENTIFIC DESCRIPTION: WellFARE is a Python library implementing linear inversion methods for the reconstruction of gene expression profiles from fluorescent or luminescent reporter gene data. WellFARE form the computational core of the WellInverter web application.
FUNCTIONAL DESCRIPTION: As input, WellFARE reads the primary data file produced by a 96-well microplate reader, containing time-series measurements of the absorbance (optical density) as well as the fluorescence and luminescence intensities in each well (if available). Various functions exist to analyze the data, in particular for detecting outliers, subtracting background, estimating growth rates, promoter activities and protein concentrations, visualizing expression profiles, synchronizing replicate profiles, etc. WellFARE is the computational core of the web application WellInverter.
NEWS OF THE YEAR: Submission of a journal publication describing the new version of WellInverter and WellFARE

- Participants: Delphine Ropers, Hans Geiselmann, Hidde De Jong, Michel Page, Valentin Zulkower and Yannick Martin
- Partner: UGA
- Contact: Hidde De Jong
- Publication: Robust reconstruction of gene expression profiles from reporter gene data using linear inversion
- URL: https://github.com/ibis-inria/wellfare

5.2. WellInverter

KEYWORDS: Bioinformatics - Statistics - Data visualization - Data modeling
SCIENTIFIC DESCRIPTION: WellInverter is a web application that implements linear inversion methods for the reconstruction of gene expression profiles from fluorescent or luminescent reporter gene data. WellInverter makes the methods available to a broad audience of biologists and bioinformaticians. In particular, we have put in place a parallel computing architecture with a load balancer to distribute the analysis queries over several back-end servers, redesigned the graphical user interface, and developed a plug-in system for defining high-level routines for parsing data files produced by microplate readers from different manufacturers.
FUNCTIONAL DESCRIPTION: As input, WellInverter reads the primary data file produced by a 96-well microplate reader, containing time-series measurements of the absorbance (optical density) as well as the fluorescence and luminescence intensities in each well (if available). Various modules exist to analyze the data, in particular for detecting outliers, subtracting background, estimating growth rates, promoter activities and protein concentrations, visualizing expression profiles, synchronizing replicate profiles, etc. The computational core of the web application consists of the Python library WellFARE.
NEWS OF THE YEAR: Deployment of WellInverter on an Inria server and on the new cloud of the French Institute for Bioinformatics (see the web page for details). Submission of a journal article describing the new version of the application.

- Participants: Delphine Ropers, Hans Geiselmann, Hidde De Jong, Johannes Geiselmann, Michel Page, Valentin Zulkower and Yannick Martin
- Partner: UGA
- Contact: Hidde De Jong
- Publication: Robust reconstruction of gene expression profiles from reporter gene data using linear inversion
- URL: https://team.inria.fr/ibis/wellinverter/

5.3. FluoBacTracker

KEYWORDS: Bioinformatics - Biology - Biomedical imaging

SCIENTIFIC DESCRIPTION: FluoBacTracker is an ImageJ plugin allowing the segementation and tracking of growing bacterial cells from time-lapse microscopy movies. The segmentation and tracking algorithms used by FluoBacTracker have been developed by Lionel Moisan and colleagues at Université Paris Descartes.
FUNCTIONAL DESCRIPTION: FluoBacTracker has the following functionalities: 1) Select regions of interest in images of microcolonies 2) Denoise and renormalize the images 3) Identify each cells in each image (segmentation) 4) Follow cells through the whole movie (tracking), including the detection of cells washed out from a microfluidics channel 5) Detect divisions and construct cell lineage of the population
NEWS OF THE YEAR: Version 2 of FluoBacTracker also allows the analysis of microscopy of bacteria growing in a microfluidics device called "mother machine".

- Participants: Hugues Berry, Cyril Dutrieux, Hidde De Jong, Charles Kervrann, David Parsons and Magali Vangkeosay
- Partners: Université Descartes - UGA
- Contact: Hugues Berry
- URL: http://fluobactracker.inrialpes.fr

5.4. GNA

Genetic Network Analyzer
KEYWORDS: Model Checking - Bioinformatics - Gene regulatory networks - Qualitative simulation
SCIENTIFIC DESCRIPTION: Genetic Network Analyzer (GNA) is the implementation of methods for the qualitative modeling and simulation of gene regulatory networks developed in the IBIS project-team.
FUNCTIONAL DESCRIPTION: The input of GNA consists of a model of the regulatory network in the form of a system of piecewise-linear differential equations (PLDEs), supplemented by inequality constraints on the parameters and initial conditions. From this information, GNA generates a state transition graph summarizing the qualitative dynamics of the system. In order to analyze large graphs, GNA allows the user to specify properties of the qualitative dynamics of a network in temporal logic, using high-level query templates, and to verify these properties on the state transition graph by means of standard model-checking tools, either locally installed or accessible through a remote web server.

RELEASE FUNCTIONAL DESCRIPTION: (1) it supports the editing and visualization of regulatory networks, in an SBGN-compatible format, (2) it semi-automatically generates a prototype model from the network structure, thus accelerating the modeling process, and (3) it allows models to be exported in the SBML Qual standard.
NEWS OF THE YEAR: New mode of distribution from the IBIS web site.

- Participants: Hidde De Jong, Michel Page and Delphine Ropers
- Partner: UGA
- Contact: Hidde De Jong
- Publications: Genetic Network Analyzer: A Tool for the Qualitative Modeling and Simulation of Bacterial Regulatory Networks - Piecewise linear approximations to model the dynamics of adaptation to osmotic stress by food-borne pathogens
- URL: http://www-helix.inrialpes.fr/gna

6. New Results

6.1. Analysis of fluorescent reporter gene data

The use of fluorescent and luminescent reporter genes allows real-time monitoring of gene expression, both at the level of individual cells and cell populations (Section 3.2). Over the years, many useful resources have appeared, such as libraries of reporter strains for model organisms and computer tools for designing reporter plasmids. Moreover, the widespread adoption of thermostated microplate readers in experimental laboratories has made it possible to automate and multiplex reporter gene assays on the population level. This has resulted in large time-series data sets, typically comprising $10^5 - 10^6$ measurements of absorbance, fluorescence, and luminescence for 10^3 wells on the microplate. In order to fully exploit these data sets, we need sound mathematical methods to infer biologically relevant quantities from the primary data and computer tools to apply the methods in an efficient and user-friendly manner.

In the past few years we developed novel methods for the analysis of reporter gene data obtained in microplate experiments, based on the use of regularized linear inversion. This allows a range of estimation problems to be solved, notably the inference of growth rate, promoter activity, and protein concentration profiles. The linear inversion methods, published in *Bioinformatics* in 2015 [12], have been implemented in the Python package WELLFARE and integrated in the web application WELLINVERTER. Funded by a grant from the Institut Français de Bioinformatique (IFB), we improved WellInverter by developing a parallel computational architecture with a load balancer to distribute the analysis queries over several back-end servers, a new graphical user interface, and a plug-in system for defining high-level routines for parsing data files produced by microplate readers from different manufacturers. This has resulted in a scalable and user-friendly web service providing a guaranteed quality of service, in terms of availability and response time. This year the web service has been redeployed on the new IFB cloud and on an Inria server, accompanied by extensive user documentation, online help, and a tutorial. Moreover, we submitted a journal paper on WELLINVERTER illustrating the use of the tool by analyzing data of the expression of a fluorescent reporter gene controlled by a phage promoter in growing *Escherichia coli* populations. We notably show that the expression pattern in different growth media, supporting different growth rates, corresponds to the pattern expected for a constitutive gene.

Compared to most reporter gene assays based on fluorescence proteins, luciferase reporters have a superior signal-to-noise ratio, since they do not suffer from the high autofluorescence background of the bacterial cell. At the same time, however, luciferase reporters have the drawback of constant light emission, which leads to undesired cross-talk between neighbouring wells on a microplate. To overcome this limitation, Marco Mauri in collaboration with colleagues from the Philipps-Universität Marburg developed a computational method to correct for luminescence bleed-through and to estimate the "true" luminescence activity for each well of a microplate. As the sole input our algorithm uses the signals measured from a calibration plate, in which the light emitted from a single luminescent well serves as an estimate for the "light-spread function". We show that this light-spread function can be used to deconvolve any other measurement obtained under the same technical conditions. Our analysis demonstrates that the correction preserves low-level signals close to the background and shows that it is universally applicable to different kinds of microplate readers and plate types. A journal article on this work was submitted this year.

6.2. Microdomain formation of bacterial membrane proteins

Fluorescent reporters can be used not only to quantify gene expression, but also to localize proteins in different compartments of the cell, In particular, in bacteria proteins within the cytoplasmic membrane display distinct localization patterns and arrangements. While multiple models exist describing the dynamics of membrane proteins, to date there have been few systematic studies, particularly in bacteria, to evaluate how protein size, number of transmembrane domains, and temperature affect their diffusion, and if conserved localization patterns exist.

Marco Mauri in collaboration with colleagues from the Philipps-Universität in Marburg has used fluorescence microscopy, single-molecule tracking (SMT), and computer-aided visualization methods to obtain a better understanding of the three-dimensional organization of bacterial membrane proteins, using the model bacterium *Bacillus subtilis*. First, we carried out a systematic study of the localization of over 200 *B. subtilis* membrane proteins, tagged with monomeric mVenus-YFP at their original gene locus. Their subcellular localization could be discriminated in polar, septal, patchy, and punctate patterns. Almost 20% of membrane proteins specifically localized to the cell poles, and a vast majority of all proteins localized in distinct structures, which we term microdomains. Dynamics were analyzed for selected membrane proteins, using SMT. Diffusion coefficients of the analyzed transmembrane proteins did not correlate with protein molecular weight, but correlated inversely with the number of transmembrane helices, *i.e.*, transmembrane radius. We observed that temperature can strongly influence diffusion on the membrane, in that upon growth temperature upshift, diffusion coefficients of membrane proteins increased and still correlated inversely to the number of transmembrane domains, following the Saffman–Delbrück relation.

The vast majority of membrane proteins were observed to localize to distinct multimeric assemblies. Diffusion of membrane proteins can be suitably described by discriminating diffusion coefficients into two protein populations, one mobile and one immobile, the latter likely constituting microdomains. Moreover, in this study published in *BMC Biology* [18], we provided a method to correct the diffusion coefficient for the membrane curvature. Our results show there is high heterogeneity and yet structural order in the cell membrane, and provide a roadmap for our understanding of membrane organization in prokaryotes. Given the exceptionally richness of the data obtained, both further analysis on membrane lateral movements and a more detailed theory on the effect of membrane curvature is possible.

6.3. Stochastic modeling and identification of gene regulatory networks in bacteria

At the single-cell level, the processes that govern single-cell dynamics in general and gene expression in particular are better described by stochastic models rather than the deterministic models underlying the linear inversion methods discussed in Section 6.6. Modern techniques for the real-time monitoring of gene expression in single cells enable one to apply stochastic modelling to study the origins and consequences of random noise in response to various environmental stresses, and the emergence of phenotypic variability. The potential impact of single-cell stochastic analysis and modelling ranges from a better comprehension of the biochemical regulatory mechanisms underlying cellular phenotypes to the development of new strategies for the (computer assisted or genetically engineered) control of cell populations and even of single cells.

Work in IBIS on gene expression and interaction dynamics at the level of individual cells is addressed in terms of identification of parametric intrinsic noise models, on the one hand, and the nonparametric inference of gene expression statistics, on the other hand, from population snapshot data. Along with modelling and inference, identifiability analysis is dedicated special attention. Other problems related with single-cell modelling, extracellular variability and inheritance of traits at cell division are considered also through external collaborations, as discussed below (Section 6.4).

Concerning identification of intrinsic noise dynamics in single cells, previous results on the contribution of stochasticity to parameter identifiability and on reconstruction of unknown gene regulatory networks have been taken further. For the case of population snapshot meaurements, where the dynamics of the population statistics are observed by simple time-lapse experiments, our earlier results showing that variance measurements may provide tremendous improvement in network reconstruction relative to sole mean measurements have been developed into a full-blown method for first-order gene network reconstruction. Additionally, parameter identifiability methods and results initially developed for gene expression models have been generalized to the whole class of first-order stochastic reaction networks. These developments have been presented and demonstrated by simulation in a paper published in the journal *Processes* [15].

Reconstruction of promoter activity statistics from reporter gene population snapshot data has been further investigated, leading to a full-blown spectral analysis and reconstruction method for reporter gene systems. In the context of the ANR project MEMIP (Section 7.2), we have characterized reporter systems as noisy linear systems operating on a stochastic input (promoter activity), and developed an inversion method for estimation of promoter activation statistics from reporter population snapshots that is nonparametric, in the sense that it does not assume any parametric model for the unknown promoter dynamics. This analysis rests on a more general, original generalization of moment equations that we have developed for stochastic reaction networks with state-affine rates subject to random input processes. The theoretical results, together with a demonstration of the reporter gene inversion method on simulated data, have been accepted for publication in *Automatica* this year [16]. In addition to utilization of the method on real gene-expression data, the results lend themselves to several additional applications, among which the study of extrinsic noise and the optimal design of reporter systems.

6.4. Modelling and analysis of cellular trait dynamics over lineage trees

The investigation of cellular populations at a single-cell level has already led to the discovery of important phenomena, such as the occurrence of different phenotypes in an isogenic population. Nowadays, several experimental techniques, such as microscopy combined with the use of microfluidic devices, enable one to take investigation further by providing time-profiles of the dynamics of individual cells over entire lineage trees. The difficulty, and at the same time the opportunity, in exploiting these data and inferring mathematical models from them is the fact that the behavior of different cells is correlated because of inheritance.

From the modelling point of view, lineage trees are well described by structured branching population models where the life cycle of each cell depends on individual characteristics, such as size and internal protein dynamics, which play a key role in the mechanisms of cell division. One important aspect in the analysis of these population models consists in the investigation of biases arising from the sampling of a finite set of observed individuals. In order to characterize bias, we studied the dynamics of a structured branching population where the trait of each individual evolves in accordance with a Markov process. We assumed that the rate of division of each individual is a function of its trait and when a branching event occurs, the trait of the descendants at birth depends on their number and on the trait of the mother. We explicitly described the Markov process, named auxiliary process, corresponding to the dynamics of the trait of a "typical" individual by deriving its associated infinitesimal generator. In particular, we proved that this process characterizes exactly the process of the trait of a uniformly sampled individual in a large population approximation. This work, carried out by Aline Marguet, has been accepted for publication in the journal *Bernoulli* [19].

We also investigated the long-time behavior of the population and proved that a typical individual in the population asymptotically behaves like the auxiliary process previously introduced. These results have been submitted for publication [22]. Structured branching processes and their analysis also provide the basis for identification tools for lineage-tree data. In particular, in the context of a bifurcating Markov chain, where each individual is characterized by a trait evolving in accordance with a scalar diffusion, we proved that the maximum-likelihood estimator of the division rate is asymptotically efficient and demonstrate the method on simulated data. This work, in collaboration with M. Hoffmann at Univ Paris-Dauphine, has also been submitted for publication [21].

Along the same lines, modelling and identification of gene expression models with mother-daughter inheritance are being investigated in the context of the ANR project MEMIP. Starting from an earlier work of the group [7], with reference to an application on osmotic shock response by yeast, the key question is to what extent leveraging an inheritance model improves inference of individual cell dynamics as well as of inheritance dynamics themselves, relative to state-of-art approaches where inheritance is not accounted for at a modelling stage.

6.5. Models of carbon metabolism in bacteria

Adaptation of bacterial growth to changes in environmental conditions, such as the availability of specific carbon sources, is triggered at the molecular level by the reorganization of metabolism and gene expression: the concentration of metabolites is adjusted, as well as the concentration and activities of enzymes, the rate of metabolic reactions, the transcription and translation rates, and the stability of proteins and RNAs. This reprogramming of the bacterial cell is carried out by i) specific interactions involving regulatory proteins or RNAs that specifically respond to the change of environmental conditions and ii) global regulation involving changes in the concentration of RNA polymerase, ribosomes, and metabolite pools that globally affect the rates of transcription, translation, and degradation of all RNAs and proteins.

A quantitative description and understanding of this complex network, cutting across metabolism, gene expression, and signalling, can be accessed through mathematical modelling only. In collaboration with Andreas Kremling, professor at TU München and former visiting scientist in the IBIS project-team, Hans Geiselmann, Delphine Ropers and Hidde de Jong developed an ensemble of variants of a simple core model of carbon catabolite repression. The model variants, with two substrate assimilation pathways and four intracellular metabolites only, differ from one another in only a single aspect, each breaking the symmetry between the two pathways in a different manner. Interestingly, all model variants are able to reproduce the data from a reference diauxic growth experiment. For each of the model variants, we predicted the behaviour in two new experimental conditions. When qualitatively comparing these predictions with experimental data, a number of models could be excluded while other model variants are still not discriminable. The best-performing model variants are based on inducer inclusion and activation of enzymatic genes by a global transcription factor, but the other proposed factors may complement these well-known regulatory mechanisms. The model ensemble, which was described in a study pubmished in *BMC Systems Biology* this year, offers a better understanding of the variety of mechanisms that have been proposed to play a role in carbon catabolite repression, but is also useful as an educational resource for systems biology.

The same focus on the dynamics of physiological processes has shaped a project on the post-transcriptional control of carbon central metabolism in *E. coli*. In the framework of the PhD thesis of Manon Morin, supported by a Contrat Jeune Scientifique INRA-Inria, the collaboration of Delphine Ropers with Muriel Cocaign-Bousquet and Brice Enjalbert at INRA/INSA Toulouse has demonstrated the key role played by the post-transcriptional regulatory system CSR in growth transitions in a series of publications in the past few years (*e.g.*, [9]). The collaboration with INRA/INSA de Toulouse is continued in the context of the PhD thesis of Thibault Etienne, funded by an INRA-Inria PhD grant, with the objective of developing models able to explain how cells coordinate their physiology and the functioning of the transcription, translation, and degradation machineries following changes in the availability of carbon sources in the environment. This work is further supported by the ANR project ECORIB accepted this year and an IXXI grant in collaboration with the ERABLE project-team (Section 7.2).

6.6. Modelling bacterial growth

Various mathematical approaches have been used in the literature to describe the networks of biochemical reactions involved in microbial growth. With various levels of detail, the resulting models provide an integrated view of these reaction networks, including the transport of nutrients from the environment and the metabolism and gene expression allowing the conversion of these nutrients into biomass. The models hence bridge the scale between individual reactions to the growth of cell populations. Analysing the dynamics of some of these models mentioned above becomes quickly intractable, when mathematical functions are for instance given by complex algebraic expressions resulting from the mass balance of biochemical reactions. In a paper published in the *Bulletin of Mathematical Biology* [13], Edith Grac, former post-doc in IBIS, Delphine Ropers, and Stefano Casagranda and Jean-Luc Gouzé from the BIOCORE project-team, have studied how monotone system theory and time-scale arguments can be used to reduce high-dimension models based on the mass-action law. Applying the approach to an important positive feedback loop regulating the expression of RNA polymerase in *E. coli*, made it possible to study the stability of the system steady states and relate the dynamical behaviour of the system to observations on the physiology of the bacterium *E. coli*.

In another paper published in *BMC Systems Biology* [14], Delphine Ropers and BIOCORE members Stefano Casagranda, Jean-Luc Gouzé, and Suzanne Touzeau, have developed a new approach to deal with model complexity. The approach, named Principle Process Analysis, allows to identify processes playing a key role in the model dynamics and to reduce the complex dynamics to these core processes, omitting processes that are inactive. In particular, it has allowed the reduction of a well-known model of circadian rhythms in mammals into a succession of simpler submodels. Their analysis has resulted in the identification of the source of circadian oscillations, the main oscillator being the negative feedback loop involving proteins PER, CRY, CLOCK-BMAL1, in agreement with previous modelling and experimental studies.

Recent work has shown that coarse-grained models of resource allocation can account for a number of empirical regularities relating the the macromolecular composition of the cell to the growth rate. Some of these models hypothesize control strategies enabling microorganisms to optimize growth. While these studies focus on steady-state growth, such conditions are rarely found in natural habitats, where microorganisms are continually challenged by environmental fluctuations. In recent years, in the framework of the PhD thesis of Nils Giordano, we extended the study of microbial growth strategies to dynamical environments, using a self-replicator model. In collaboration with the BIOCORE project-team, we formulated dynamical growth maximization as an optimal control problem that can be solved using Pontryagin's Maximum Principle and we compared the theoretical results thus obtained with different possible implementations of growth control in bacterial cells [5]. The extension and experimental validation of some of these results are currently being carried out by Antrea Pavlou in the framework of her PhD project, funded by the ANR project Maximic (Section 7.2).

6.7. Growth control in bacteria and biotechnological applications

The ability to experimentally control the growth rate is crucial for studying bacterial physiology. It is also of central importance for applications in biotechnology, where often the goal is to limit or even arrest growth. Growth-arrested cells with a functional metabolism open the possibility to channel resources into the production of a desired metabolite, instead of wasting nutrients on biomass production. In recent years we obtained a foundation result for growth control in bacteria [6], in that we engineered an *E. coli* strain where the transcription of a key component of the gene expression machinery, RNA polymerase, is under the control of an inducible promoter. By changing the inducer concentration in the medium, we can adjust the RNA polymerase concentration and thereby switch bacterial growth between zero and the maximal growth rate supported by the medium. The publication also presented a biotechnological application of the synthetic growth switch in which both the wild-type *E. coli* strain and our modified strain were endowed with the capacity to produce glycerol when growing on glucose. Cells in which growth has been switched off continue to be metabolically active and harness the energy gain to produce glycerol at a twofold higher yield than in cells with natural control of RNA polymerase expression.

The experimental work underlying the growth switch has been continued in several directions in the context of the Maximic project by Célia Boyat. Moreover, in collaboration with colleagues from the BIOCORE project-team, we have formulated the maximization of metabolite production by means of the growth switch as a resource reallocation problem that can be analyzed by means of the self-replicator models of bacterial growth mentioned in Section 6.6 in combination with methods from optimal control theory. In a publication accepted for the *Journal of Mathematical Biology* this year [20], we study various optimal control problems by means of a combination of analytical and computational techniques. We show that the optimal solutions for biomass maximization and product maximization are very similar in the case of unlimited nutrient supply, but diverge when nutrients are limited. Moreover, external growth control overrides natural feedback growth control and leads to an optimal scheme consisting of a first phase of growth maximization followed by a second phase of product maximization. This two-phase scheme agrees with strategies that have been proposed in metabolic engineering. More generally, this work shows the potential of optimal control theory for better understanding and improving biotechnological production processes. Extensions concerning the effect on growth and bioproduction of the (biological or technological) costs associated with discontinuous control strategies, and of the time allotted to optimal substrate utilization, are described in a contribution to a control theory conference submitted this year.

7. Partnerships and Cooperations

7.1. Regional Initiatives

Project name	MuSE: MUlti-Omics and Metabolic models integration to study growth transition in Escherichia coli
Coordinator IBIS participants Type Web page	D. Ropers D. Ropers, T. Etienne IXXI/BioSyl project (2018-2020) http://www.biosyl.org/news/muse-2013-multi-omics-and-metabolic-models-integration-to-study-growth-transition-in-escherichia-coli

Project name	RNAfluo: Quantification d'ARN régulateurs *in vivo*
Coordinator IBIS participants Type	S. Lacour S. Lacour AGIR project Univ Grenoble Alpes (2016-2019)

7.2. National Initiatives

Project name	MEMIP – Modèles à effets mixtes de processus intracellulaires : méthodes, outils et applications
Coordinator IBIS participants Type	G. Batt E. Cinquemani, A. Marguet, D. Ropers ANR project (2016-2020)

Project name	ENZINVIVO – Détermination in vivo des paramètres enzymatiques dans une voie métabolique synthétique
Coordinator IBIS participants Type	G. Truan J. Geiselmann, H. de Jong ANR project (2016-2020)

Project name	MAXIMIC: Optimal control of microbial cells by natural and synthetic strategies
Coordinator IBIS participants Type Web page	H. de Jong C. Boyat, E. Cinquemani, J. Geiselmann, H. de Jong, A. Pavlou, C. Pinel, D. Ropers ANR project (2017-2021) https://project.inria.fr/maximic

Project name	RIBECO (RIBonucleotide ECOnomy): Engineering RNA life cycle to optimize economy of microbial energy
Coordinator IBIS participants Type Web page	M. Cocaign-Bousquet E. Cinquemani, T. Etienne, D. Ropers ANR project (2018-2022) https://project.inria.fr/ribeco/

Project name	COSY: real-time COntrol of SYnthetic microbial communities
Coordinator IBIS participants Type Web page	E. Cinquemani E. Cinquemani, H. de Jong, J. Geiselmann, M. Mauri, T. Muszbek, C. Pinel, D. Ropers Inria Project Lab (2017-2021) https://project.inria.fr/iplcosy/

Project name	CoSoft: Control software for a system of mini-bioreactors
Coordinator IBIS participants Type	E. Cinquemani E. Cinquemani, H. de Jong, J. Geiselmann, T. Muszbek Inria Hub (2017-2018)

Project name	AlgeaInSilico: Prédire et optimiser la productivité des microalgues en fonction de leur milieu de croissance
Coordinator IBIS participants Type Web page	O. Bernard H. de Jong, N. Giordano Inria Project Lab (2015-2019) https://project.inria.fr/iplalgaeinsilico/

Project name	Analyse intégrative de la coordination entre stabilité des ARNm et physiologie cellulaire chez Escherichia coli
Coordinators IBIS participants Type	D. Ropers, M. Cocaign-Bousquet (Inra, LISBP) T. Etienne, D. Ropers Contrat Jeune Scientifique Inra-Inria (2016-2019)

7.3. European Initiatives

7.3.1. Collaborations with Major European Organizations

Systems biotechnology department at Technische Universität Münich (Germany), Andreas Kremling

Modeling of carbon metabolism in bacteria

Automatic control laboratory at ETH Zürich (Switzerland), John Lygeros, and Cell cycle laboratory, School of Medicine, University of Patras (Greece), Zoe Lygeros

Control theory and systems identification with applications to single-cell systems biology

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific events organisation

8.1.1.1. Member of organizing committees

IBIS members	Conference, workshop, school	Date
Hidde de Jong	CompSysBio: Advanced Lecture Course on Computational Systems Biology, Aussois	Apr 2019

8.1.2. Scientific events selection

8.1.2.1. Chair of conference program committees

IBIS member	Conference, workshop, school	Role
Eugenio Cinquemani	European Control Conference (ECC 2019)	Associate editor

8.1.2.2. Member of conference program committees

IBIS member	Conference, workshop, program
Eugenio Cinquemani	ECC 2019, CMSB 2018 and 2019, HSB 2019, SASB 2018
Hidde de Jong	CMSB 2108 and 2019, FOSBE 2019, HSB 2019, MLCSB 2018, ISMB Network Biology 2018, WML 2018

8.1.3. Journal

8.1.3.1. Member of editorial boards

IBIS member	Journal
Johannes Geiselmann	Frontiers in Microbiology (review editor)
Hidde de Jong	Journal of Mathematical Biology
Hidde de Jong	Biosystems (reviews editor)
Hidde de Jong	ACM/IEEE Transactions on Computational Biology and Bioinformatics

8.1.4. Scientific evaluation and expertise

IBIS member	Organism	Role
Johannes Geiselmann	INRA	Member scientific advisory committee Microbiologie, Adaptation, Pathogénie
Johannes Geiselmann	UMR5240 CNRS-UCBL-INSA-BayerCropScience	Member scientific council
Hidde de Jong	Microbiology and Food Chain Department, Inra	Member scientific council
Hidde de Jong	Univ Grenoble Alpes	Member scientific council of Pôle MSTIC
Delphine Ropers	INRA-Inria	Member selection committee PhD grants

8.1.5. Invited talks and other presentations

Eugenio Cinquemani

Title	Event and location	Date
Estimation du bruit extrinsèque dans l'expression génique dans des cellules individuelles	Invited talk Journées Apprentissage de modèles statistiques et stochastiques à partir de données biologiques (ASSTABIO), Rennes	Mar 2018
Real-time control of synthetic microbial communities	Presentation Journée InBio, Lyon	Apr 2018
Towards automated control of synthetic E. coli communities	Presentation Journées GDR BIOSS-IA	Dec 2018

Hidde de Jong

Title	Event and location	Date
Carbon catabolite repression and diauxic growth in E. coli	Invited talk Principles of Microbial Adaptation, Lorentz Center, Leiden, the Netherlands	Mar 2018
Optimal control of bacterial growth for the maximization of metabolite production	Presentation SMAI-MODE 2018, Autrans	Mar 2018
Analysis and control of bacterial regulatory networks	Presentation Journée InBio, Lyon	Apr 2018
Global physiological effects and the analysis of gene regulatory networks	Invited talk 1st International Plant Systems Biology Meeting, Conférence Jacques Monod, Roscoff	Sep 2018
Natural and synthetic control of resource allocation in bacteria	Invited talk Molecular Logic and Computational Synthetic Biology workshop, Santiago de Chile	Dec 2018

Stephan Lacour

Title	Event and location	Date
Influence des ARNs non-codants réprimant CsgD sur la production de curli par Escherichia coli	Poster Ecole Thématique de Microbiologie Moléculaire, Carry-Le-Rouet	Oct 2018

Aline Marguet

Title	Event and location	Date
Estimation statistique dans une population structurée branchante	Seminar Institut Fourier, Grenoble	Jan 2018
Loi des grands nombres pour des processus de branchement structurés	Seminar Institut de Mathématiques de Toulouse	Mar 2018
Loi des grands nombres pour des processus de branchement structurés	Seminar Institut de Recherche Mathématique Avancée, Strasbourg	Mar 2018
Loi des grands nombres pour des processus de branchement structurés	Seminar Institut Montpelliérain Alexander Grothendieck, Montpellier	Mar 2018
Estimation statistique dans une population branchante structurée par une diffusion	Presentation journées MAS du groupe Modélisation Aléatoire et Statistique, Dijon	Aug 2018
Estimation statistique dans une population branchante structurée par une diffusion	Seminar Institut de Mathématiques de Dijon, Dijon	Aug 2018

Marco Mauri

Title	Event and location	Date
Analysis of dynamics of membrane-protein microdomains in bacteria	Poster annual Deutsche Physikalische Gesellschaft (DPG) meeting, Berlin	Mar 2018

Michel Page

Title	Event and location	Date
WellInverter: A web application for the analysis of fluorescent reporter gene data	Poster journéé Institut Français de Bioinformatique (IFB), Paris	Dec 2018

Corinne Pinel

Title	Event and location	Date
Présentation des techniques de génie génétique utilisées au sein du LIPhy	Presentation at the kick-off meeting of the network Ingénierie des Génomes des Micro-organismes (IGM), Paris	Sep 2018

Delphine Ropers

Title	Event and location	Date
Quantitative modelling of metabolism, gene expression and growth	Presentation Journée InBio, Lyon	Apr 2018

Etienne Thibault

Title	Event and location	Date
Coordination de la stabilité des ARNms et la physiologie cellulaire durant des transitions de taux de croissance	Presentation working group BIOSS of GDR IM/BIM, Marseille	Jul 2018
Coordination de la stabilité des ARNms et la physiologie cellulaire durant des transitions de taux de croissance	Poster at ICSB 2018, Lyon	Oct 2018
Coordination de la stabilité des ARNms et la physiologie cellulaire durant des transitions de taux de croissance	Seminar LISBP, INRA-INSA, Toulouse	Dec 2018

8.1.6. Research administration

IBIS member	Committee	Role
Eugenio Cinquemani	Inria Grenoble - Rhône-Alpes	Member Comité des Emplois Scientifiques (CES)
Eugenio Cinquemani	Inria Grenoble - Rhône-Alpes	Member Comité des Utilisateurs des Moyens Informatiques (CUMI)
Eugenio Cinquemani	Inria	Member Comité Administrative Paritaire (CAP)
Eugenio Cinquemani	Inria Grenoble - Rhône-Alpes	Member Comité Développement Technologique (CDT)
Hidde de Jong	Inria Grenoble - Rhône-Alpes	Member scientific council (COS)
Hidde de Jong	Inria	Member working group on International Relations of Conseil d'Orientation Scientifique et Technique (COST)
Delphine Ropers	Inria	Member of Commission d'évaluation d'Inria
Delphine Ropers	Inria	Member of Groupe de travail plan stratégique
Delphine Ropers	Inria Grenoble - Rhône-Alpes	Référente chercheurs
Delphine Ropers	Inria Grenoble - Rhône-Alpes	Member of Comité des études doctorales (CED)

8.1.7. Recruitment committees

IBIS member	Organism	Recruitment
Johannes Geiselmann Delphine Ropers	Inria	Chargés de recherche (jury d'admissibilité)
Delphine Ropers	Inria Grenoble - Rhône-Alpes	Chargés de recherche (jury d'admissibilité)
Delphine Ropers	INSA de Lyon	Assistant professor

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Four members of the IBIS team are either full professor or associate professor at Univ Grenoble Alpes. They therefore have a full teaching service (at least 192 hours per year) and administrative duties related to the organization and evaluation of the university course programs on all levels (from BSc to PhD). Besides the full-time academic staff in IBIS, the following people have contributed to courses last year.

Eugenio Cinquemani

Course: Stochastic modelling of gene regulatory networks, M2, BIM, INSA de Lyon (6 h)
Course: Statistics for systems biology, M1, Master Approches Interdisciplinaires du Vivant, CRI/Univ Paris Descartes (20 h, also in charge of 20 h of practicals)
Course: Modelling and identification of metabolic networks, M1, Phelma, INP Grenoble (4 h)

Hidde de Jong

Course and practicals: Modeling and simulation of gene regulatory networks, M2, BIM, INSA de Lyon (20 h)

Delphine Ropers

Course and practicals: Modelling in systems biology, M1, Phelma, INP Grenoble (16 h)
Course and practicals: Cell systems biology and modelling cell functions, M1, Master ingéniérie de la santé, Univ Grenoble Alpes (14 h)
Course: Modeling and simulation of genetic regulatory networks, M2, INSA de Toulouse (4 h)

Thibault Etienne

Practicals: Mathématiques et statistique appliquées aux sciences humaines et sociales (practicals), L1 and L3, Univ Grenoble Alpes (56 h)
Course and practicals: Construction et analyse de plasmides in silico, M1, Master ingéniérie de la santé, Univ Grenoble Alpes (10 h)

8.2.2. Supervision

PhD in progress: **Thibault Etienne**, Analyse intégrative de la coordination entre stabilité des ARNm et physiologie cellulaire chez *Escherichia coli*. Supervisors: Delphine Ropers and Muriel Cocaign-Bousquet (INRA Toulouse)
PhD in progress: **Joël Espel**, RNA engineering: Design of the dynamical folding of RNA and of RNA switches. Supervisors: Alexandre Dawid (Univ Grenoble Alpes) and Johannes Geiselmann
PhD in progress: **Antrea Pavlou**, Experimental and computational analysis of bacterial self-replicators. Supervisors: Hidde de Jong and Johannes Geiselmann

8.2.3. Juries

PhD thesis committees

IBIS member	Role	PhD student	University	Date
Johannes Geiselmann	Rapporteur	Clea Lachaux	Univ Paul Sabatier Toulouse	Sep 2018
Johannes Geiselmann	Président	Tomas Andersen	Univ Grenoble Alpes	Oct 2018
Hidde de Jong	Président	Alexandre Rocca	Univ Grenoble Alpes	May 2018
Hidde de Jong	Rapporteur	Jonathan Behaegel	Univ Côte d'Azur	Oct 2018

Habilitation (HDR) committees

IBIS member	Role	PhD student	University	Date
Johannes Geiselmann	Rapporteur	Fabien Letisse	Univ Paul Sabatier Toulouse	Oct 2018

PhD advisory committees

IBIS member	PhD student	University
Johannes Geiselmann	Alain Lombard	Univ Grenoble Alpes
Johannes Geiselmann	Shiny Martis	INSA de Lyon
Hidde de Jong	Céline Hernandez	ENS Paris
Hidde de Jong	Charlotte Coton	INRA Moulon
Stephan Lacour	Julien Trouillon	CEA Grenoble
Delphine Ropers	Manon Barthe	Univ de Toulouse
Delphine Ropers	Irene Ziska	Univ de Lyon

8.2.4. Teaching administration

Yves Markowicz is director of the BSc department at Univ Grenoble Alpes.

Michel Page is coordinator of the master Systèmes d'information et d'organisation at the Institut d'Adminstration des Entreprises (IAE), Univ Grenoble Alpes.

Eugenio Cinquemani organizes a module on statistics in systems biology at CRI/Univ Paris Descartes.

Delphine Ropers organizes a module on the mathematical modeling of biological systems at PHELMA, INP Grenoble.

Hidde de Jong organizes with Daniel Kahn a module on the modeling of genetic and metabolic networks at INSA de Lyon.

8.3. Popularization

8.3.1. Creation of media or tools for science outreach

Hidde de Jong presented the work of IBIS in the framework of the communication action "Mon équipe en 180 secondes". The video is on-line at the following address: https://videotheque.inria.fr/videotheque/media/47816;jsessionid=D615F60B85E6C88DEDD129AD56E4431E

9. Bibliography

Major publications by the team in recent years

[1] V. Baldazzi, D. Ropers, Y. Markowicz, D. Kahn, J. Geiselmann, H. de Jong. *The carbon assimilation network in Escherichia coli is densely connected and largely sign-determined by directions of metabolic fluxes*, in "PloS Computational Biology", 2010, vol. 6, n° 6, e1000812

[2] S. Berthoumieux, M. Brilli, H. de Jong, D. Kahn, E. Cinquemani. *Identification of metabolic network models from incomplete high-throughput datasets*, in "Bioinformatics", 2011, vol. 27, n° 13, p. i186-i195

[3] S. Berthoumieux, H. De Jong, G. Baptist, C. Pinel, C. Ranquet, D. Ropers, J. Geiselmann. *Shared control of gene expression in bacteria by transcription factors and global physiology of the cell*, in "Molecular Systems Biology", January 2013, vol. 9, n° 1, 11 [*DOI :* 10.1038/msb.2012.70], http://hal.inria.fr/hal-00793352

[4] E. CINQUEMANI, V. LAROUTE, M. BOUSQUET, H. DE JONG, D. ROPERS.*Estimation of time-varying growth, uptake and excretion rates from dynamic metabolomics data*, in "Bioinformatics", 2017, vol. 33, n^{o} 14, p. i301-i310 [*DOI :* 10.1093/BIOINFORMATICS/BTX250], https://hal.archives-ouvertes.fr/hal-01607919

[5] N. GIORDANO, F. MAIRET, J.-L. GOUZÉ, J. GEISELMANN, H. DE JONG.*Dynamical allocation of cellular resources as an optimal control problem: Novel insights into microbial growth strategies*, in "PLoS Computational Biology", March 2016, vol. 12, n^{o} 3, e1004802 [*DOI :* 10.1371/JOURNAL.PCBI.1004802], https://hal.inria.fr/hal-01332394

[6] J. IZARD, C. GOMEZ-BALDERAS, D. ROPERS, S. LACOUR, X. SONG, Y. YANG, A. B. LINDNER, J. GEISELMANN, H. DE JONG.*A synthetic growth switch based on controlled expression of RNA polymerase*, in "Molecular Systems Biology", November 2015, vol. 11, n^{o} 11, 16, https://hal.inria.fr/hal-01247993

[7] A. LLAMOSI, A. GONZALEZ, C. VERSARI, E. CINQUEMANI, G. FERRARI-TRECATE, P. HERSEN, G. BATT.*What population reveals about individual cell identity: Single-cell parameter estimation of models of gene expression in yeast*, in "PLoS Computational Biology", February 2016, vol. 12, n^{o} 2, e1004706, https://hal.inria.fr/hal-01248298

[8] M. MORIN, D. ROPERS, F. LETISSE, S. LAGUERRE, J.-C. J.-C. PORTAIS, M. COCAIGN-BOUSQUET, B. ENJALBERT.*The post-transcriptional regulatory system CSR controls the balance of metabolic pools in upper glycolysis of Escherichia coli*, in "Molecular Microbiology", January 2016, vol. 4, 15 [*DOI :* 10.1111/MMI.13343], https://hal.inria.fr/hal-01418224

[9] M. MORIN, D. ROPERS, F. LETISSE, S. LAGUERRE, J.-C. J.-C. PORTAIS, M. COCAIGN-BOUSQUET, B. ENJALBERT.*The Csr system regulates Escherichia coli fitness by controlling glycogen accumulation and energy levels*, in "mBio", 2017, vol. 8, n^{o} 5, p. 1-14, https://hal.inria.fr/hal-01672038

[10] C. PEANO, J. WOLF, J. DEMOL, E. ROSSI, L. PETITI, G. DE BELLIS, J. GEISELMANN, T. EGLI, S. LACOUR, P. LANDINI.*Characterization of the Escherichia coli σ(S) core regulon by Chromatin Immunoprecipitation-sequencing (ChIP-seq) analysis*, in "Scientific Reports", 2015, vol. 5, 15 [*DOI :* 10.1038/SREP10469], https://hal.inria.fr/hal-01217828

[11] D. STEFAN, C. PINEL, S. PINHAL, E. CINQUEMANI, J. GEISELMANN, H. DE JONG.*Inference of quantitative models of bacterial promoters from time-series reporter gene data*, in "PLoS Computational Biology", 2015, vol. 11, n^{o} 1, e1004028, https://hal.inria.fr/hal-01097632

[12] V. ZULKOWER, M. PAGE, D. ROPERS, J. GEISELMANN, H. DE JONG.*Robust reconstruction of gene expression profiles from reporter gene data using linear inversion*, in "Bioinformatics", 2015, vol. 31, n^{o} 12, p. i71-i79, https://hal.inria.fr/hal-01217800

Publications of the year

Articles in International Peer-Reviewed Journal

[13] I. BELGACEM, S. CASAGRANDA, E. GRAC, D. ROPERS, J.-L. GOUZÉ.*Reduction and stability analysis of a transcription-translation model of RNA polymerase*, in "Bulletin of Mathematical Biology", 2018, vol. 80, n^{o} 2, p. 294-318 [*DOI :* 10.1007/S11538-017-0372-4], https://hal.inria.fr/hal-01655367

[14] S. CASAGRANDA, S. TOUZEAU, D. ROPERS, J.-L. GOUZÉ.*Principal process analysis of biological models*, in "BMC Systems Biology", 2018, vol. 12, 68 [*DOI :* 10.1186/S12918-018-0586-6], https://hal.inria.fr/hal-01818033

[15] E. CINQUEMANI.*Identifiability and Reconstruction of Biochemical Reaction Networks from Population Snapshot Data*, in "Processes", August 2018, vol. 6, n^{o} 9, p. 1-24 [*DOI :* 10.3390/PR6090136], https://hal.inria.fr/hal-01909765

[16] E. CINQUEMANI.*Stochastic reaction networks with input processes: Analysis and application to gene expression inference*, in "Automatica", 2018, p. 1-9, https://hal.inria.fr/hal-01925923

[17] A. KREMLING, J. GEISELMANN, D. ROPERS, H. DE JONG.*An ensemble of mathematical models showing diauxic growth behaviour*, in "BMC Systems Biology", December 2018, vol. 12, n^{o} 1, p. 1-16 [*DOI :* 10.1186/S12918-018-0604-8], https://hal.inria.fr/hal-01893978

[18] D. LUCENA, M. MAURI, F. SCHMIDT, B. ECKHARDT, P. L. GRAUMANN.*Microdomain formation is a general property of bacterial membrane proteins and induces heterogeneity of diffusion patterns*, in "BMC Biology", December 2018, vol. 16, n^{o} 1 [*DOI :* 10.1186/S12915-018-0561-0], https://hal.inria.fr/hal-01966627

[19] A. MARGUET.*Uniform sampling in a structured branching population*, in "Bernoulli", 2019, https://arxiv.org/abs/1609.05678 , https://hal.archives-ouvertes.fr/hal-01362366

[20] I. YEGOROV, F. MAIRET, H. DE JONG, J.-L. GOUZÉ.*Optimal control of bacterial growth for the maximization of metabolite production*, in "Journal of Mathematical Biology", October 2018, p. 1-48 [*DOI :* 10.1007/S00285-018-1299-6], https://hal.inria.fr/hal-01929475

Other Publications

[21] M. HOFFMANN, A. MARGUET.*Statistical estimation in a randomly structured branching population*, November 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01662203

[22] A. MARGUET.*A law of large numbers for branching Markov processes by the ergodicity of ancestral lineages*, March 2018, https://arxiv.org/abs/1707.07993 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01567317

Project-Team IMAGINE

Intuitive Modeling and Animation for Interactive Graphics & Narrative Environments

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

IN PARTNERSHIP WITH:
CNRS

Institut polytechnique de Grenoble

Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Interaction and visualization

Table of contents

Project-Team IMAGINE

Creation of the Team: 2012 January 01, updated into Project-Team: 2013 January 01

Keywords:

Computer Science and Digital Science:

A5. - Interaction, multimedia and robotics
A5.5. - Computer graphics
A5.5.1. - Geometrical modeling
A5.5.3. - Computational photography
A5.5.4. - Animation
A5.6. - Virtual reality, augmented reality
A5.7. - Audio modeling and processing
A9.3. - Signal analysis

Other Research Topics and Application Domains:

B2. - Health
B2.2. - Physiology and diseases
B3. - Environment and planet
B3.3. - Geosciences
B5. - Industry of the future
B5.2. - Design and manufacturing
B5.7. - 3D printing
B9.1. - Education
B9.2.2. - Cinema, Television
B9.2.3. - Video games
B9.2.4. - Theater
B9.6.6. - Archeology, History

1. Team, Visitors, External Collaborators

Research Scientists

Rémi Ronfard [Team leader, Inria, Senior Researcher, HDR]
Frédéric Devernay [Inria, Researcher, until Sep 2018]
Mélina Skouras [Inria, Researcher]

Faculty Members

Stefanie Hahmann [Institut polytechnique de Grenoble, Professor, HDR]
Jean-Claude Léon [Institut polytechnique de Grenoble, Professor, HDR]
Olivier Palombi [Univ Grenoble Alpes, Associate Professor, HDR]

Technical Staff

Maguelonne Beaud de Brive [Inria, until Nov 2018]
Rémi Colin de Verdière [Inria, from Oct 2018]
Julien Daval [Inria]
Harold Vilmart [Univ Grenoble Alpes, until Sept 2018]

PhD Students

Thomas Buffet [Inria]
Pierre Casati [Inria]
Guillaume Cordonnier [Univ. Grenoble Alpes, until Sep 2018]
Pablo Coves [Univ Grenoble Alpes, until Mar 2018]
Amelie Fondevilla [Univ Grenoble Alpes]
Qianqian Fu [Univ Grenoble Alpes, from Nov 2018]
Maxime Garcia [Univ Grenoble Alpes]
Geoffrey Guingo [Institut polytechnique de Grenoble]
Youna Le Vaou [Cifre PSA]
Vaishnavi Ameya Murukutla [Univ Grenoble Alpes]
Sandra Nabil Mahrous Yacoub [Inria, until Nov 2018]

Post-Doctoral Fellow
Musaab Khalid Osman Mohammed [Institut polytechnique de Grenoble, from Oct 2018]

Visiting Scientist
David Jourdan [Inria, from Jun 2018 until Jul 2018]

Administrative Assistant
Marion Ponsot [Inria]

2. Overall Objectives

2.1. Context

With the fast increase of computational power and of memory space, increasingly complex and detailed 3D content is expected for virtual environments. Unfortunately, 3D modeling methodologies did not evolve as fast: most users still use standard CAD or 3D modeling software (such as Maya, 3DS or Blender) to design each 3D shape, to animate them and to manually control cameras for movie production. This is highly time consuming when large amounts of detailed content need to be produced. Moreover the quality of results is fully left in the user's hand, which restricts applicability to skilled professional artists. More intuitive software such as Z-Brush are restricted to shape design and still require a few months for being mastered by sculpture practitioners. Reducing user load can be done by capturing and re-using real objects or motions, at the price of restricting the range of possible content. Lastly, procedural generation methods can be used in specific cases to automatically get some detailed, plausible content. Although they save user's time, these procedural methods typically come at the price of control: indirect parameters need to be tuned during a series of trial and errors until the desired result is reached. Stressing that even skilled digital artists tend to prefer pen and paper than 3D computerized tools during the design stages of shapes, motion, and stories, Rob Cook, vice president of technology at Pixar animation studios notoriously stated (key-note talk, Siggraph Asia 2009): *new grand challenge in Computer Graphics is to make tools as transparent to the artists as special effects were made transparent to the general public*. This remains true ten years later.

Could digital modeling be turned into a tool, even more expressive and simpler to use than a pen, to quickly convey and refine shapes, motions and stories? This is the long term vision towards which we would like to advance.

2.2. Scientific goals

The goal of the IMAGINE project is to develop **a new generation of models, algorithms and interactive environments for the interactive creation of animated 3D content and its communication through virtual cinematography**.

Our insight is to revisit models for shapes, motion, and narration from a user-centred perspective, i.e. to give models an intuitive, predictable behaviour from the user's view-point. This will ease both semi-automatic generation of animated 3D content and fine tuning of the results. The three main fields will be addressed:

1. **Shape design**: We aim to develop intuitive tools for designing and editing 3D shapes and their assemblies, from arbitrary ones to shapes that obey application-dependent constraints - such as, for instance, developable surfaces representing cloth or paper, or shape assemblies used for CAD of mechanical prototypes.
2. **Motion synthesis**: Our goal is to ease the interactive generation and control of 3D motion and deformations, in particular by enabling intuitive, coarse to fine design of animations. The applications range from the simulation of passive objects to the control of virtual creatures.
3. **Narrative design**: The aim is to help users to express, refine and convey temporal narrations, from stories to educational or industrial scenarios. We develop both virtual direction tools such as interactive storyboarding frameworks, and high-level models for virtual cinematography, such as rule-based cameras able to automatically follow the ongoing action and automatic film editing techniques.

In addition to addressing specific needs of digital artists, this research contributes to the development of new expressive media for 3D content. The long term goal would be to enable any professional or scientist to model and interact with their object of study, to provide educators with ways to quickly express and convey their ideas, and to give the general public the ability to directly create animated 3D content.

3. Research Program

3.1. Methodology

As already stressed, thinking of future digital modeling technologies as an Expressive Virtual Pen enabling to seamlessly design, refine and convey animated 3D content, leads to revisit models for shapes, motions and stories from a user-centered perspective. More specifically, inspiring from the user-centered interfaces developed in the Human Computer Interaction domain, we introduced the new concept of user-centered graphical models. Ideally, such models should be designed to behave, under any user action, the way a human user would have predicted. In our case, user's actions may include creation gestures such as sketching to draft a shape or direct a motion, deformation gestures such as stretching a shape in space or a motion in time, or copy-paste gestures to transfer some of the features from existing models to other ones. User-centered graphical models need to incorporate knowledge in order to seamlessly generate the appropriate content from such actions. We are using the following methodology to advance towards these goals:

- Develop high-level models for shapes, motion and stories that embed the necessary knowledge to respond as expected to user actions. These models should provide the appropriate handles for conveying the user's intent while embedding procedural methods that seamlessly take care of the appropriate details and constraints.
- Combine these models with expressive design and control tools such as gesture-based control through sketching, sculpting, or acting, towards interactive environments where users can create a new virtual scene, play with it, edit or refine it, and semi-automatically convey it through a video.

3.2. Validation

Validation is a major challenge when developing digital creation tools: there is no ideal result to compare with, in contrast with more standard problems such as reconstructing existing shapes or motions. Therefore, we had to think ahead about our validation strategy: new models for geometry or animation can be validated, as usually done in Computer Graphics, by showing that they solve a problem never tackled before or that they provide a more general or more efficient solution than previous methods. The interaction methods we are developing for content creation and editing rely as much as possible on existing interaction design principles already validated within the HCI community. We also occasionally develop new interaction tools, most often in collaboration with this community, and validate them through user studies. Lastly, we work with expert users from various application domains through our collaborations with professional artists, scientists from other domains, and industrial partners: these expert users validate the use of our new tools compared to their usual pipeline.

4. Application Domains

4.1. Visual arts

Our research can be applied to any situation where users need to create new, imaginary, 3D content. Our work should be instrumental, in the long term, for the visual arts, from the creation of 3D films and games to the development of new digital planning tools for theater or cinema directors. Our models can also be used in interactive prototyping environments for engineering. They can help promoting interactive digital design to scientists, as a tool to quickly express, test and refine models, as well as an efficient way for conveying them to other people. Lastly, we expect our new methodology to put digital modeling within the reach of the general public, enabling educators, media and other practitioners to author their own 3D content.

The diversity of users these domains bring, from digital experts to other professionals and novices, gives us excellent opportunities to validate our general methodology with different categories of users. Our ongoing projects in these various application domains are listed in Section 6.

- Sculpture.
- Modeling and animation for 3D films and games.
- Virtual cinematography and tools for theater directors.

4.2. Engineering

- Industrial design.
- Mechanical & civil engineering.

4.3. Natural sciences

- Geology.
- Virtual functional anatomy.

4.4. Education and creative tools

- Sketch-based teaching.
- Creative environments for novice users.
- Museography

5. Highlights of the Year

5.1. Highlights of the Year

This is a transition year where the team has been actively involved in starting new projects with new PhD students along new research directions, which will be further emphasized in the future team ANIMA, due to start after the termination of the IMAGINE team in July 2019.

Figure 1. Filming rehearsals of jean-Francois Peyret's La fabrique des monstres, Théâtre de Vidy, Lausanne, January 2018.

We are now actively involved in the Performance Lab, a joint cross-disciplinary research program of IDEX Univ. Grenoble Alpes. In this new project started in January 2018 for three years, we will investigate "digital dramaturgies" mixing real-time computer graphics, augmented and virtual reality with live performances. We will also continue to develop our Kino Ai video capture, analysis and editing system.

As a follow-up to ADT ULTRAHD, we recorded three weeks of rehearsals from the play "La fabrique des monstres", a theatre adaptation of Mary Shelley's Frankenstein by Jean-Francois Peyret (Fig. 1). Our Kino Ai system was used to automatically generate six hours of cinematographic rushes from those recordings. Those rushes were edited by professional film editors into three short documentaries and published online (see episode 1, episode 2 and episode 3 to watch the full movies).

5.1.1. Awards

In December 2018, PhD laureate Guillaume Cordonnier was awarded the prestigious ETH Zurich Postdoctoral Fellowship and will join the Computer Graphics Lab's simulation group in 2019.

5.1.2. Patents

As part of Youna Le Vaou's CIFRE PhD thesis with PSA, we filed a joint patent application:

Y La Vaou, S Masfrand, M Mika, S Hahmann, J-C Léon: Procédé de modification de la forme d'un objet virtuel tridimensionnel représenté dans un espace immersif et système immersif mettant en œuvre ledit procédé, December 2018.

This new result will also be submitted for publication at an international conference in 2019.

6. New Software and Platforms

6.1. MyCF

My Corporis Fabrica

KEYWORDS: 3D modeling - Simulation - Health - Ontologies - Anatomy - Patientspecific - Medical imaging
FUNCTIONAL DESCRIPTION: Knowledge-based 3D anatomical modeling using MyCF The MyCF software eases the creation of 3D anatomical models for visualization and mechanical simulation. As input, the user provides a list of anatomical entities or functions to simulate, using keywords or navigating in reference 3D model. As output, she gets a 3D model ready to visualize, or to simulate.

- Participants: Ali Hamadi Dicko, Federico Ulliana, François Faure and Olivier Palombi
- Partner: Université Joseph-Fourier
- Contact: Olivier Palombi

6.2. Kino AI

Artificial intelligence for cinematography
KEYWORDS: Video analysis - Post-production
FUNCTIONAL DESCRIPTION: Kino AI is an implementation of the method described in our patent "automatic generation of cinematographic rushes using video processing". Starting from a single ultra high definitiion (UltraHD) recording of a live performance, we track and recognize all actors present on stage and generate one or more rushes suitable for cinematographic editing of a movie.

- Partner: IIIT Hyderabad
- Contact: Rémi Ronfard
- Publications: Multi-Clip Video Editing from a Single Viewpoint - Zooming On All Actors: Automatic Focus+Context Split Screen Video Generation

6.3. Platforms

6.3.1. AANALYSIS

AANALYSIS is a platform developed from 2015 to 2018 as a module of the SALOME platform and as a plugin of SHAPER (a software under development by EDF, CEA and OpenCascade). It processes CAD assemblies to derive symmetry properties, similarities of solids and groups of solids, interface features between solids and is operational on industrial models. The plugin version implements a shape similarity selection feature. The module is connected with a knowledge database MyProductFabrica developed in our group in partnership with GraphiK Inria team. It enables the automatic generation/annotation of functional knowledge from 3D geometry using 350 concepts, 90 relations and more than 100 inferences.

6.3.2. RUMBA

RUMBA is a next-generation 3D animation software targeted to professional animation studios, developed jointly by Mercenaries Engineering, TEAMTO and IMAGINE. Development was funded in part by FUI projects COLLODI 1 and COLLODI 2. RUMBA has been used in production by TEAMTO since 2017. RUMBA will be made commercially available to other animation studio by Mercenaries Engineering in 2019. We are using RUMBA as a platform for developing new algorithms in sketch-based animation, based on our previous work during Martin Guay's PhD thesis. This includes many improvements to allow those methods to work in a professional workflow.

6.3.3. SKY ENGINE

Sky Engine is a realtime game engine developped by Maxime Garcia as part of his PhD thesis, which incorporates several shape modleing and animation tools developped within the team. It is hoped that it will supersede Expressive as a platform for future integration of research results of the team involving real-time, story-driven shape modeling, animation and cinematography.

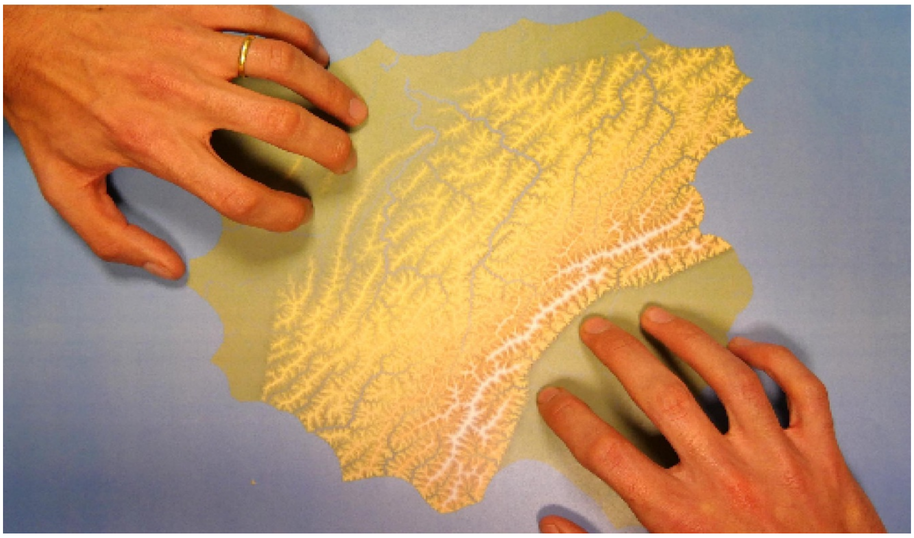


Figure 2. Sculpting Mountains: Interactive Terrain Modeling Based on Subsurface Geology

7. New Results

7.1. Sculpting Mountains: Interactive Terrain Modeling Based on Subsurface Geology

Most mountain ranges are formed by the compression and folding of colliding tectonic plates. Subduction of one plate causes large-scale asymmetry while their layered composition (or stratigraphy) explains the multi-scale folded strata observed on real terrains. As part of Guillaume Cordonnier's PhD thesis, we introduced a novel interactive modeling technique to generate visually plausible, large scale terrains that capture these phenomena (illustrated in Fig. 2). Our method draws on both geological knowledge for consistency and on sculpting systems for user interaction. The user is provided hands-on control on the shape and motion of tectonic plates, represented using a new geologically-inspired model for the Earth crust. The model captures their volume preserving and complex folding behaviors under collision, causing mountains to grow. It generates a volumetric uplift map representing the growth rate of subsurface layers. Erosion and uplift movement are jointly simulated to generate the terrain. The stratigraphy allows us to render folded strata on eroded cliffs. We validated the usability of our sculpting interface through a user study, and compare the visual consistency of the earth crust model with geological simulation results and real terrains.

7.2. Exploratory design of mechanical devices with motion constraints

Mechanical devices are ubiquitous in our daily lives, and the motion they are able to transmit is often a critical part of their function. While digital fabrication devices facilitate their realization, motion-driven mechanism design remains a challenging task. We take drawing machines as a case study in exploratory design. Devices such as the Spirograph can generate intricate patterns from an assembly of simple mechanical elements. Trying to control and customize these patterns, however, is particularly hard, especially when the number of parts increases. We propose a novel constrained exploration method that enables a user to easily explore feasible drawings by directly indicating pattern preferences at different levels of control. This is (illustrated in Fig. 3). The user starts by selecting a target pattern with the help of construction lines and rough sketching, and

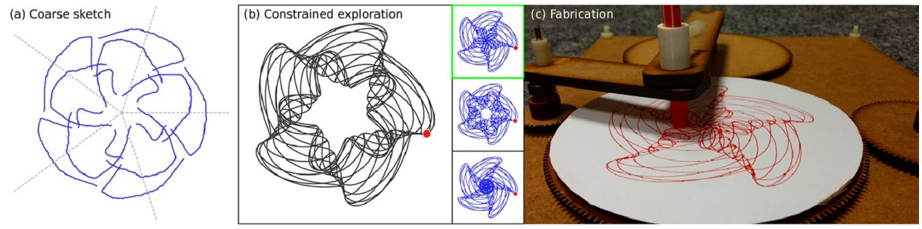


Figure 3. Exploratory design of mechanical devices with motion constraints. (a) The user first selects a mechanically feasible drawing by providing a rough sketch. (b) The user is then able to interactively explore local alternatives (b) by defining visual constraints directly on the pattern (here, the cusp position). (c) The resulting machine is automatically exported to laser cutter profiles for fabrication.

then fine-tunes it by prescribing geometric features of interest directly on the drawing. The designed pattern can then be directly realized with an easy-to-fabricate drawing machine. The key technical challenge is to facilitate the exploration of the high dimensional configuration space of such fabricable machines. To this end, we propose a novel method that dynamically reparameterizes the local configuration space and allows the user to move continuously between pattern variations, while preserving user-specified feature constraints. We tested our framework on several examples, conducted a user study, and fabricated a sample of the designed examples.

7.3. Automatic Generation of Geological Stories from a Single Sketch

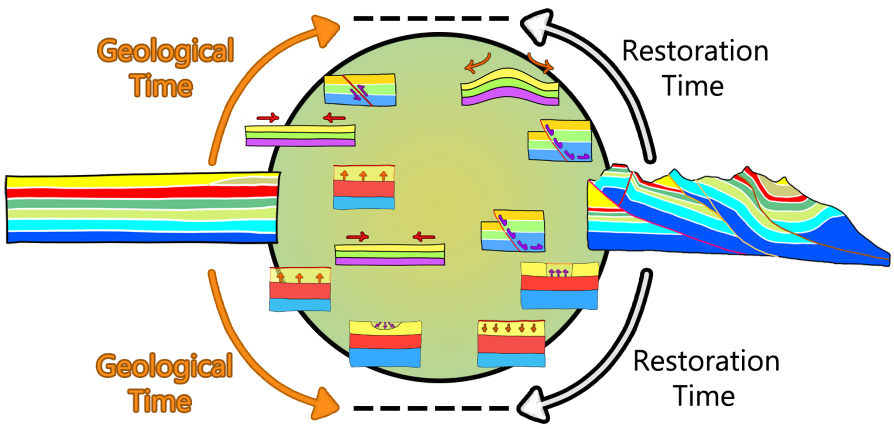


Figure 4. Automatic Generation of Geological Stories from a Single Sketch. From left to right, the original terrain from several million years ago undergoes events that will transform it to its current state. From right to left, the current terrain is restored and undergoes undo events that will transform it back to its original state.

Describing the history of a terrain from a vertical geological cross-section is an important problem in geology, called geological restoration. Designing the sequential evolution of the geometry is usually done manually, involving many trials and errors. In this work, we recast this problem as a storyboarding problem, where the diifferent stages in the restoration are automatically generated as storyboard panels and displayed as geological stories. Our system allows geologists to interactively explore multiple scenarios by selecting plausible geological event sequences and backward simulating them at interactive rate, causing the terrain layers to be progressively un-deposited, un-eroded, un-compacted, unfolded and un-faulted. Storyboard sketches are generated along the way. When a restoration is complete, the storyboard panels can be used for automatically generating a forward animation of the terrain history, enabling quick visualization and validation of hypotheses. As a proof-of-concept, we describe how our system was used by geologists to restore and animate cross-sections in real examples at various spatial and temporal scales and with different levels of complexity, including the Chartreuse region in the French Alps.

7.4. 3D Shape Decomposition and Sub-parts Classification

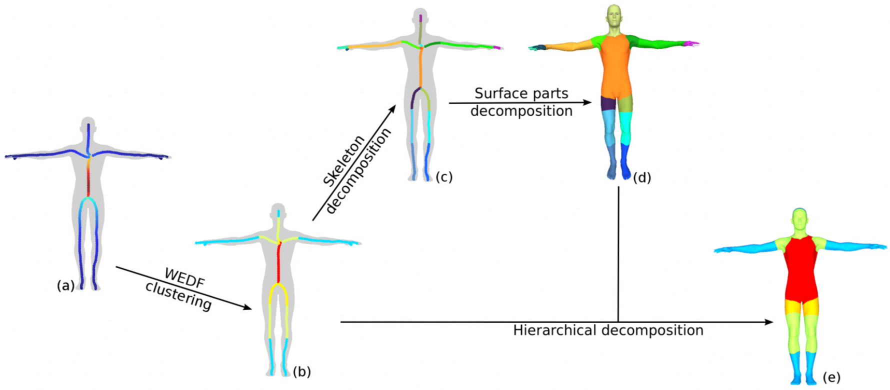


Figure 5. 3D Shape Decomposition and Sub-parts Classification. Starting from a 3D shape and its curve skeleton, we compute a new measure called WEDF on the curve skeleton (a) and, by clustering WEDF values, we decompose the skeleton into hierarchical parts (b). To each connected part on the skeleton –shown with a different color (c)– a connected region of the surface mesh is assigned (d). Then, a salience value according to the hierarchy is assigned to each corresponding surface part (e) –parts of same importance get a similar color.

This paper (illustrated in Fig. 5) introduces a measure of significance on a curve skeleton of a 3D piecewise linear shape mesh, allowing the computation of both the shape's parts and their saliency. We begin by reformulating three existing pruning measures into a non-linear PCA along the skeleton. From this PCA, we then derive a volume-based salience measure, the 3D WEDF, that determines the relative importance to the global shape of the shape part associated to a point of the skeleton. First, we provide robust algorithms for computing the 3D WEDF on a curve skeleton, independent on the number of skeleton branches. Then, we cluster the WEDF values to partition the curve skeleton, and coherently map the decomposition to the associated surface mesh. Thus, we develop an unsupervised hierarchical decomposition of the mesh faces into visually meaningful shape regions that are ordered according to their degree of perceptual salience. The shape analysis tools introduced in this paper are important for many applications including shape comparison, editing, and compression.

7.5. Interactive Generation of Time-evolving, Snow-Covered Landscapes with Avalanches

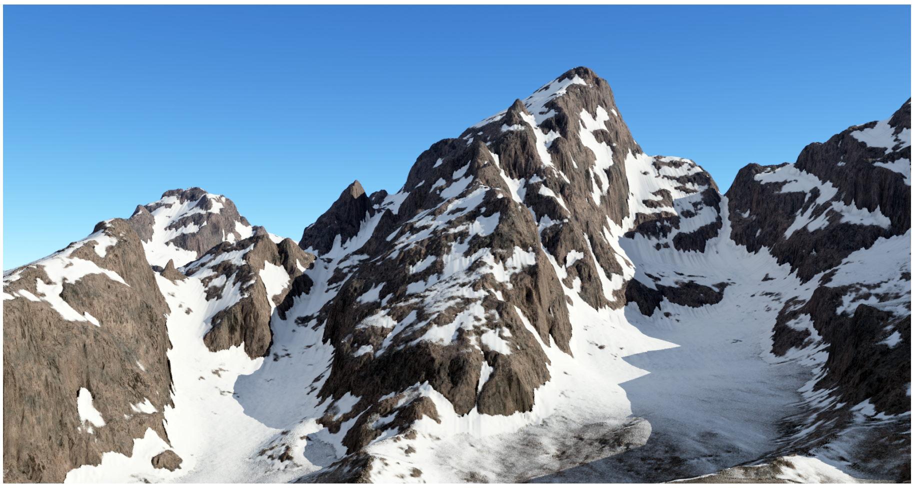


Figure 6. Interactive Generation of Time-evolving, Snow-Covered Landscapes with Avalanches

As part of Guillaume Cordonnier's PhD thesis, we also introduced a novel method for interactive generation of visually consistent, snow-covered landscapes, which provides control of their dynamic evolution over time. Our main contribution (illustrated in Fig. 6) was the real-time phenomenological simulation of avalanches and other user-guided events, such as tracks left by Nordic skiing, which can be applied to interactively sculpt the landscape. The terrain is modeled as a height field with additional layers for stable, compacted, unstable, and powdery snow, which behave in combination as a semi-viscous fluid. We incorporate the impact of several phenomena, including sunlight, temperature, prevailing wind direction, and skiing activities. The snow evolution includes snow-melt and snowdrift, which affect stability of the snow mass and the probability of avalanches. A user can shape landscapes and their evolution either with a variety of interactive brushes, or by prescribing events along a winter season time-line. Our optimized GPU-implementation allows interactive updates of snow type and depth across a large (10×10 km) terrain, including real-time avalanches, making this suitable for visual assets in computer games. We evaluated our method through perceptual comparison against exiting methods and real snow-depth data.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

We have on going contract with EDF R & D Saclay (Raphael Marc) on "Shape analysis of mechanical assemblies and their components". This has funded the work of Harold Vilmar until August 2018. The project has been renewed for 2019 and is funding the development of the AANALYSIS software (Jean-Claude Léon).

8.2. Bilateral Grants with Industry

We have an ongoing CIFRE PhD contract with PSA on the topic of aesthetic shape modeling in immersive virtual reality environments, which is funding the PhD of Youna Le Vaou.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Performance Lab (January 2018 - June 2021)

Participants: Rémi Ronfard, Qianqian Fu, Mélina Skouras, Maxime Garcia, Pierre Casati, Vaishnavi Ameya Murukutla, Rémi Colin de Verdière.

Performance Lab is a cross-disciplinary project (CDP) by IDEX Univ. Grenoble Alpes, started in January 2018, which is funding the Phd thesis of Qianqian Fu.

Conceived as an international platform, the Performance Lab brings together a community of researchers who are exploring contemporary issues that link embodiment, society and technology. The ambition of the project is to renew the ways in which research is conceived and practiced at Univ. Grenoble Alpes by developing new methods inspired by Anglo-Saxon notions of Performance as Research (PAR), research creation, practice-led and based research.

As part of the Performance Lab, tIMAGINE is actively involved in the research group on "digital dramaturgies" co-led by Remi Ronfard and Julie Valero.

9.2. National Initiatives

9.2.1. InriaHub ADT Kino Ai (October 2018-September 2020)

Participants: Rémi Ronfard, Rémi Colin de Verdière, Qianqian Fu.

This two-year contract is a follow up to the one-year InriaHub ULTRAHD project which was successfully completed in December 2017. Kino Ai is a joint research project of the IMAGINE team at Inria Grenoble Alpes, and the Performance Lab at Univ. Grenoble Alpes. Following our previous work in “multiclip video editing” and “Split Screen Video Generation”, we are working to provide a user-friendly environment for editing and watching ultra-high definition movies online, with an emphasis on recordings of live performances.

The code from Vineet Gandhi’s PhD thesis was r entirely re-designed for supporting ultra high definition video. T The software was extensively tested in 2017 on a large dataset of 4K video recordings of theatre rehearsals, in collaboration with the Litt&Arts team at Univ. Grenoble Alpes, theatre director Jean-Francois Peyret in Paris, Theatre de l’Hexagone in Meylan and Theatre de Vidy in Lausanne. The goal of the Kino AI ADT is to allow the Kino Ai python code to run in a web server, and to provide a redesigned user interface (in javascript) running on a web client. The user interface will be designed, tested and evaluated with the Litt&Arts team at Univ. Grenoble Alpes, as part of CDP project Performance Lab.

9.2.2. FUI LIVE360 TV(December 2015 - December 2018)

Participants: Frédéric Devernay, Sandra Nabil Mahrous Yacoub.

L’objectif de ce projet collaboratif est de développer une solution bout-en-bout pour la création « live », la diffusion et la restitution d’audio/vidéo 360° multi-écrans, et ce avec une qualité répondant aux exigences du marché « broadcast ».

Ce projet est né sous l’impulsion d’un consortium formé de PME (Arkamys, ATEME, Aviwest et Kolor) et de laboratoires (Inria et Télécom ParisTech). Il bénéficie du programme FUI19, le Fonds Unique Interministériel. This 3-year contract with industrial partners Arkamys, ATEME, Aviwest et Kolor (now GoPRO) was dedicated to creating an end-to-end solution for recording and broadcasting immersive multi-screen 360 degree audio/video movies with a professional quality.

The project has funded the PhD thesis of Sandra Nabil. It was completed in November 2018, with the Phd defense of Sandra Nabil and the closing FUI project meeting.

9.2.3. FUI Collodi 2 (December 2016 - April 2019)

Participants: Rémi Ronfard, Maguelonne Beaud de Brive, Julien Daval.

This 2-year contract with two industrial partners: TeamTo and Mercenaries Engineering (software for production rendering), is a follow-up and a generalization of Dynam'it and Collodi 1. The goal is to propose an integrated software for the animation and final rendering of high-quality movies, as an alternative to the ever-ageing Maya. The project is funding 2 engineers for 2 years.

The project was extended for four additional months from January to April 2019 to allow extended expert evaluation of our sketch-based animation toolkit.

9.2.4. FUI 3D-Oncochip (October 2018 - September 2021)

Participants: Jean-Claude Léon, Musaab Khalid Osman Mohammed.

3D-Oncochip project is a collaboration with Microlight 3D, with the objective of fabricating nanoscale 3D microtumors, which are human biological models of real tumors. This 3-year contract is funding the postdoc position of Musaab Khalid Osman Mohammed.

9.2.5. ANR E-ROMA (November 2017 - October 2020)

Participants: Rémi Ronfard, Stefanie Hahmann, Pierre Casati.

The eRoma project aims at revisiting the digitization and virtual restoration of archaeological and fine arts artefacts by taking advantage of the sites from which they were retrieved and the eras they belong to. To do so, e-Roma will develop a new virtual representation both versatile and unified enough to be used for both restoration and animation of digitized artworks. Traditional cardboard models with a fixed and rigid representation will therefore be replaced by interactive dynamic virtual prototypes, to help restore statues and illustrate changes over time.

This 3-year contract is a joint project with GeoMod team at LIRIS and the musée gallo-romain in Lyon. The contract started in November 2017 and is funding the PhD thesis of Pierre Casati.

9.2.6. ANR FOLD-DYN (November 2017 - October 2020)

Participant: Thomas Buffet.

The FOLDDyn project (Field-Oriented Layered Dynamics animating 3D characters) proposes the study of new theoretical approaches for the effective generation of virtual characters deformations, when they are animated. These deformations are twofolds: character skin deformations (skinning) and garment simulations. We propose to explore the possibilities offered by a novel theoretical way of addressing character deformations: the implicit skinning. This method jointly uses meshes (the standard representation for 3D animations) and volumetric scalar functions (an unusual representation in this community).

This 3-year contract is a joint project with the University of Toulouse.The contract started in November 2017 and is funding the PhD thesis of Thomas Buffet.

9.2.7. ANR ANATOMY2020 (November 2017 - October 2020)

Participants: Olivier Palombi, Rémi Ronfard, Vaishnavi Ameya Murukutla.

Anatomy2020 aims at developing an innovative educational platform to facilitate learning of functional anatomy. This platform will integrate recent advances in computer graphics, human-computer interaction together with recent insights in educational and cognitive sciences to design and test optimal scenarios for anatomy learning. The approach is based on evidences that body movements could improve learning of different knowledge by "augmenting" or "enriching" traces in long-term memory. This "embodied" perspective is particularly relevant for learning of functional anatomy as the knowledge to acquire could be specifically related to the learner's body in motion.

This 3-year contract is a joint project with TIMC (Computer-Assisted Medical Intervention team), Anatoscope, Gipsa-Lab (speech and cognition dept.), LIBM and LIG (Engineering Human-Computer Interaction team). The contract started in November 2017 and is funding the PhD thesis of Ameya Murukutla.

10. Dissemination

10.1. Promoting Scientific Activities

Remi Ronfard is a member of the selection committee for Inria-MCC (Ministry of Culture and Communication) activitites.

Stefanie Hahmann is a member of the Comité d'Etudes Doctorales (CED) at Inria Grenoble.

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Stefanie Hahmann was Program Chair of Conference Shape Modeling International (SMI18) held in LIsbon 2018.

10.1.1.2. Member of the Organizing Committees

Rémi Ronfard is a member of the steering committee for the Eurographics workshop on intelligence cinematography and editing (WICED).

Rémi Ronfard was a co-organizer of the international workshop Attention machine ! Pratiques artistiques et recyclages médiatiques at Univ. Grenoble Alpes in February 2018.

Stefanie Hahmann served in the International Program Committee for the Symposium on Solid and Physical Modeling (SPM18), Bilbao 2018

Jean-Claude Léon was a member of the program committee for the SPM 2018 conference.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

Mélina Skouras was the Poster Chair for ACM/Eurographics Symposium on Computer Animation 2018.

10.1.2.2. Member of the Conference Program Committees

Rémi Ronfard was a member of the Program Committees for Expressive 2018, International Conference on Interactive Storytelling (ICIDS) 2018, and Eurographics workshop on intelligence cinematography and editing (WICED) 2018.

Mélina Skouras was a member of the Program Committees for Eurographics 2018, Advances in Architectural Geometry 2018, Shape Modeling International 2018, and ACM/Eurographics Symposium on Computer Animation 2018.

10.1.2.3. Reviewer

Remi Ronfard was an external reviewer for Siggraph 2018 and Siggraph Asia 2018.

Jean-Claude Léon was a reviewer for the CAD 2018 conference.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Stefanie Hahmann is an Associate Editor of CAG (Computers and Graphics, Elsevier) and CAD (Computer Aided Design, Elsevier). She also was a guest Editor of the journal Computers and Graphics Vol. 74 (Elsevier): Special Issue on Shape Modeling International 2018,

Jean-Claude Léon is a member of the editorial board of the CAD journal.

10.1.3.2. Reviewer - Reviewing Activities

Rémi Ronfard was a reviewer for Siggraph 2018, Siggraph Asia 2018 and ACM transactions on Graphics (TOG).

Mélina Skouras was a reviewer for Siggraph Asia 2018 and ACM Transactions on Graphics.

Stefanie Hahmann was a reviewer for the journals CAD, Computers and Graphics,

10.1.4. Invited Talks

Mélina Skouras gave invited guest talks at Bellairs Workshop on Computer Animation, Fields Workshop on Robust Geometric Algorithms for Computational Fabrication, and Dagstuhl Seminar on Computational Aspects of Fabrication.

Remi Ronfard Remi Ronfard gave an invited talk on "Authoring and directing animated story worlds" at the national meeting of the GTAS (Groupe de Travail sur l'Animation et la Simulation) organized by GdR IG-RV (Informatique Géométrique et Graphique, Réalité Virtuelle et Visualisation) in Paris, July 2018.

10.1.5. Leadership within the Scientific Community

Stefanie Hahmann serves as a member of the Advisory Board (2014-2018) for the European Marie-Curie Training Network ARCADES.

Remi Ronfard and Julie Valéro (Litt & Arts, Univ. Grenoble Alpes) started a research group on Digital Dramaturgies as part of the Performance Lab, IDEX Univ. Grenoble Alpes (2018-2021). This research group is actively investigating (i) how contemporary dramaturgies represent digital worlds on stage; (ii) how contemporary dramaturgies use digital tools for writing and staging performances combining real actors with virtual actors and scenographies; and (iii) how contemporary dramaturgies can be digitally captured, indexed and analyzed for a better comprehension of the creative processes at work during pre-production and rehearsals. The research group is composed of researchers of Univ. Grenoble Alpes from multiple disciplines, i.e. litterature, theatre, choreography, film studies, social sciences, geography, computer science and applied mathematics.

10.1.6. Scientific Expertise

Remi Ronfard was a member of the scientific committee of IMAGINOVE in 2018.

Remi Ronfard was a member of the scientific committee for EXPERIMENTA 2018.

10.1.7. Research Administration

Stefanie Hahmann is the head of the French working group "GTMG" (Groupe de travail en Modélisation Géométrique) part of the CNRS GDR IM and GDR IGRV.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence : Stefanie Hahmann, Numerical Methods, Ensimag - Grenoble INP, 240 students, 3rd year Bachelor level, 42h,

Master : Stefanie Hahmann, Geometric Modeling, Ensimag - Grenoble INP, 60 students, Master 1st year, 47h.

Master : Stefanie Hahmann, Surface Modeling, Ensimag - Grenoble INP, 30 students, Master 2nd year, 51h.

Master : Rémi Ronfard, Advanced Computer Graphics, 18 HETD, M2, Univ. Grenoble Alpes, France

Master : Rémi Ronfard, Game Engine Programming, 18 HETD, M2, Univ. Montpellier, France

Licence : Mélina Skouras, Surface modeling, 14.5 HETD, Ensimag, Grenoble, France

Doctorate : Rémi Ronfard, Spatial Media, 90 HETD, ENSAD, Paris Sciences et Lettres (PSL), France

Stefanie Hahmann is co-responsible of the department MMIS (Images and Applied Maths) at Grenoble INP with 120 students. Stefanie Hahmann was also president of the jury for over 30 Masters (PFE) thesis defences in 2018.

Jean-Claude Léon is in charge of the module Mechanical Systems at Grenoble-INP ENSE3 (300 students, 64h, coordination of three courses).

Olivier Palombi is responsible for the French Campus numérique of anatomy. He is responsible and national leader of the project SIDES. All the French medical schools (43) have planed to use the same e-learning framework (SIDES) to manage evaluations and to create a large shared database of questions.

10.2.2. Supervision

PhD : Romain Brégier, Détection et estimation de pose d'instances d'objet rigide pour la manipulation robotisée, Univ. Grenoble Alpes, June 11, 2018, supervised by James Crowley and Frédéric Devernay

PhD : Even Entem, Interprétation et modélisation 3D automatique à partir de dessins au trait de formes organiques, Univ. Grenoble Alpes, October 26, 2018, supervised by Marie-Paule Cani and Loic Barthes

PhD : Sandra Nabil, Evaluation de la qualité de vidéos panoramiques synthétisées, Univ. Grenoble Alpes, November 27, 2018, supervised by James Crowley and Frédéric Devernay

PhD : Guillaume Cordonnier, Modèles à couches pour simuler l'évolution de paysages à grande échelle, Univ. Grenoble Alpes, December 6, 2018, supervised by Marie-Paule Cani and Eric Galin

PhD in progress : Maxime Garcia, Animation transfer: character animation by playing and acting, since October 2016, supervised by Rémi Ronfard

PhD in progress : Youna Le Vaou, Virtual Sculpture: shape creation and modification through immersive CAVE-like systems, since March 2017, supervised by Jean-Claude Léon and Stefanie Hahmann. Funded by CIFRE contract with PSA

PhD in progress : Amélie Fondevilla, Modélisation et animation de surfaces développables, since September 2016, supervised by Stefanie Hahmann and Damien Rohmer

PhD in progress : Ameya Murukutla, Storyboarding augmented reality anatomy lessons, since Octobre 2017, supervised by Rémi Ronfard and Olivier Palombi

PhD in progress : Pierre Casati, Modeling and animation of antique statues, since October 2017, supervised by Rémi Ronfard and Stéfanie Hahmann

PhD in progress : Qianqian Fu, Computational video editing of live performances, since November 2018, supervised by Rémi Ronfard and Benjamin Lecouteux (GETALP, LIG).

PhD in progress : Thomas Buffet, Efficient multi-layered cloth animation using implicit surfaces, since December 2017, supervised by Marie-Paule Cani and Damien Rohmer.

PhD in progress : Robin Roussel, Function-aware design for objects to be fabricated, since October 2015, supervised by Niloy Mitra, Marie-Paule Cani and Jean-Claude Léon.

PhD in progress : Geoffrey Guingo, Synthesis of animated textures, since October 2015, supervised by Marie-Paule Cani, Jean-Michel Dischler and Basile Sauvage.

PhD in progress : Nachwa Aboubakr, Observation and modeling of human activities, since October 2016, supervised by James Crowley and Rémi Ronfard.

PhD in progress: David Jourdan, Support optimization for tensile membrane structures, since October 2018, supervised by Adrien Bousseau and Mélina Skouras.

PhD in progress: Mickaël Ly, Inverse elastic shell design with contact and friction with applications to garment design, since October 2017, supervised by Florence Descoubes and Mélina Skouras.

Master's thesis : David Jourdan, Optimization of support structures for tensile architecture, June 2018, supervised by Adrien Bousseau and Mélina Skouras.

Master's thesis : Rémi Colin de Verdière, Mutli-target inverse kinematics for real-time character animation, June 2018, supervised by Remi Ronfard.

Master's thesis : Sami Chibane, What makes motion beautiful?, June 2018, supervised by Rémi Ronfard.

10.2.3. Juries

Rémi Ronfard was the president of the jury for Sandra Nabil's Phd Thesis.

Rémi Ronfard was part of the selection committee for recruiting an assistant professor at Univ. Montpellier in May 2018.

Stefanie Hahmann was member of the HDR committee of Julie Digne at Univ. Lyon 1 and a reviewer of the HDR Thesis of Alexandra Bac at Univ. Aix-Marseille. She also was a reviewer of 2 PhD thesis, Univ. Claude Bernard Lyon 1 (Mathieu Giroux) and Univ. Nice (Jean-Dominique Favreau) in 2018.

Stefanie Hahmann also participated in 2018 in 2 committees for recruiting assistent professors, one at UGA, Grenoble, the other at Univ. Aix-Marseille.

10.3. Popularization

Rémi Ronfard and Julie Valéro organized a panel at Experimenta on "Virtual reality storytelling between cinema and theatre", with invited guests Christelle Derré, (collectif Or Norme), Philippe Fuchs (MINES ParisTech), Pauline Bouchet (Univ. Grenoble Alpes), Pierre-Emmanuel Le Goff (France TV) and Annick Jakobowicz (France TV).

Rémi Ronfard presented new results of the COLLODI2 project to Computer Animation professionals at MIFA in June 2018, Next-generation computer animation with RUMBA, International Animation Film Festival Annecy.

Marie-Paule Cani and Rémi Ronfard were featured in a special report on Intelligence artificielle et création artistique, Science et Avenir, aout 2018.

Rémi Ronfard and Julie Valéro published online three new video documentaries shot using our KINO AI virtual cinematography system during rehearsals of the stage adaptation of Mary Shelley's Frankenstein, written and directed by Jean-Francois Peyret.

Rémi Ronfard and Julie Valéro wrote a blog entry, L'informatique rentre en scène, Blog Binaire, Le Monde, 4 décembre 2018.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] R. BRÉGIER.*Detection and pose estimation of instances of a rigid object for robotic bin-picking*, Univ. Grenoble Alpes, June 2018, https://hal.inria.fr/tel-01977050

[2] G. CORDONNIER.*Layered Models for Large Scale Time-Evolving Landscapes*, Université Grenoble Alpes (France), December 2018, https://tel.archives-ouvertes.fr/tel-01971947

[3] E. ENTEM.*Automatic interpretation and 3D modeling from line drawings of organic shapes*, UNIVERSITÉ GRENOBLE ALPES, October 2018, https://tel.archives-ouvertes.fr/tel-01977694

[4] S. NABIL.*Evaluation de la qualité des vidéos panoramiques assemblées*, Université Grenoble Alpes, November 2018, https://hal.archives-ouvertes.fr/tel-01979682

Articles in International Peer-Reviewed Journal

[5] E. BARTZOS, V. BORRELLI, R. DENIS, F. LAZARUS, D. ROHMER, B. THIBERT.*An Explicit Isometric Reduction of the Unit Sphere into an Arbitrarily Small Ball*, in "Foundations of Computational Mathematics", August 2018, vol. 18, n^{o} 4, p. 1015–1042 [*DOI :* 10.1007/s10208-017-9360-1], https://hal.inria.fr/hal-01647062

[6] T. BLANC-BEYNE, G. MORIN, K. LEONARD, S. HAHMANN, A. CARLIER.*A Salience Measure for 3D Shape Decomposition and Sub-parts Classification*, in "Graphical Models", September 2018, vol. 99, p. 22-30 [*DOI :* 10.1016/J.GMOD.2018.07.003], https://hal.inria.fr/hal-01840016

[7] R. BRÉGIER, F. DEVERNAY, L. LEYRIT, J. CROWLEY.*Defining the Pose of any 3D Rigid Object and an Associated Distance*, in "International Journal of Computer Vision", June 2018, vol. 126, n^o 6, p. 571–596, https://arxiv.org/abs/1612.04631 [*DOI :* 10.1007/s11263-017-1052-4], https://hal.inria.fr/hal-01415027

[8] G. CORDONNIER, M.-P. CANI, B. BENES, J. BRAUN, E. GALIN.*Sculpting Mountains: Interactive Terrain Modeling Based on Subsurface Geology*, in "IEEE Transactions on Visualization and Computer Graphics", May 2018, vol. 24, n^o 5, p. 1756-1769 [*DOI :* 10.1109/TVCG.2017.2689022], https://hal.archives-ouvertes.fr/hal-01517343

[9] G. CORDONNIER, P. ECORMIER, E. GALIN, J. GAIN, B. BENES, M.-P. CANI.*Interactive Generation of Time-evolving, Snow-Covered Landscapes with Avalanches*, in "Computer Graphics Forum", May 2018, vol. 37, n^o 2, p. 497-509 [*DOI :* 10.1111/CGF.13379], https://hal.inria.fr/hal-01736971

[10] M. GRARD, R. BRÉGIER, F. SELLA, E. DELLANDRÉA, L. CHEN.*Object segmentation in depth maps with one user click and a synthetically trained fully convolutional network*, in "Springer Proceedings in Advanced Robotics", 2019, vol. 7, p. 207-221, https://arxiv.org/abs/1801.01281 - This is a pre-print of an article published in Human Friendly Robotics, 10th International Workshop, Springer Proceedings in Advanced Robotics, vol 7. The final authenticated version is available online [*DOI :* 10.1007/978-3-319-89327-3_16], https://hal.archives-ouvertes.fr/hal-01674511

[11] M. LY, R. CASATI, F. BERTAILS-DESCOUBES, M. SKOURAS, L. BOISSIEUX.*Inverse Elastic Shell Design with Contact and Friction*, in "ACM Transactions on Graphics", November 2018, vol. 37, n^o 6, p. 1-16 [*DOI :* 10.1145/3272127.3275036], https://hal.inria.fr/hal-01883655

[12] G. NIETO, F. DEVERNAY, J. L. CROWLEY.*Rendu basé image avec contraintes sur les gradients*, in "Traitement du Signal", 2018, p. 1-26 [*DOI :* 10.3166/HSP.X.1-26], https://hal.archives-ouvertes.fr/hal-01900200

[13] R. ROUSSEL, M.-P. CANI, J.-C. LÉON, N. MITRA.*Exploratory design of mechanical devices with motion constraints*, in "Computers and Graphics", August 2018, vol. 74, p. 244-256 [*DOI :* 10.1016/J.CAG.2018.05.023], https://hal.archives-ouvertes.fr/hal-01828500

Invited Conferences

[14] S. NABIL, R. BALZARINI, F. DEVERNAY, J. L. CROWLEY.*Designing objective quality metrics for panoramic videos based on human perception*, in "IMVIP 2018 - Irish Machine Vision and Image Processing Conference", Belfast, United Kingdom, August 2018, p. 1-4, https://hal.inria.fr/hal-01849043

International Conferences with Proceedings

[15] N. ABOUBAKR, R. RONFARD, J. L. CROWLEY.*Recognition and Localization of Food in Cooking Videos*, in "Joint Workshop on Multimedia for Cooking and Eating Activities and Multimedia Assisted Dietary Management (CEA-MADiMa 2018)", Stockholm, Sweden, CEA/MADiMa '18 Proceedings of the Joint Workshop on Multimedia for Cooking and Eating Activities and Multimedia Assisted Dietary Management, ACM, July 2018, p. 21-24, In conjunction with the 27th International Joint Conference on Artificial Intelligence [*DOI :* 10.1145/3230519.3230590], https://hal.archives-ouvertes.fr/hal-01815512

[16] E. Entem, A. D. Parakkat, M.-P. Cani, L. Barthe.*Structuring and layering contour drawings of organic shapes*, in "Expressive 2018 - Joint Symposium on Computational Aesthetics and Sketch-Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering", Victoria, Canada, ACM Press, August 2018, vol. 18, p. 1-14 [*DOI :* 10.1145/3229147.3229155], https://hal.inria.fr/hal-01853410

[17] M. Garcia, M.-P. Cani, R. Ronfard, C. Gout, C. Perrenoud.*Automatic Generation of Geological Stories from a Single Sketch*, in "Expressive '18 - The Joint Symposium on Computational Aesthetics and Sketch Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering", Victoria, Canada, ACM, August 2018, p. 17-19 [*DOI :* 10.1145/3229147.3229161], https://hal.inria.fr/hal-01817923

[18] F. Morin, H. Courtecuisse, I. Reinertsen, F. Le Lann, O. Palombi, Y. Payan, M. Chabanas.*Resection-induced brain-shift compensation using vessel-based methods*, in "SPIE Medical Imaging 2018", Houston, TX, United States, Medical Imaging 2018: Physics of Medical Imaging, International Society for Optics and Photonics, February 2018, vol. 10576, p. 105760Q-1 - 105760Q-6 [*DOI :* 10.1117/12.2293640], https://hal.archives-ouvertes.fr/hal-01930274

[19] T. Stanko, N. Saguin-Sprynski, L. Jouanet, S. Hahmann, G.-P. Bonneau.*Morphorider: a new way for Structural Monitoring via the shape acquisition with a mobile device equipped with an inertial node of sensors*, in "EWSHM 2018 - 9th European Workshop on Structural Health Monitoring", Manchester, United Kingdom, July 2018, p. 1-11, https://hal.inria.fr/hal-01962789

Books or Proceedings Editing

[20] A. Gomes, S. Hahmann, E. Kalogerakis (editors). *Proceedings of Shape Modeling International (SMI) 2018 - Special Issue of Computers & Graphics Vol. 74. Jun 2018, Lisbon, Portugal, Elsevier, pp.33-243, 2018*, Elsevier, France, May 2018 [*DOI :* 10.1016/j.cag.2018.05.022], https://hal.archives-ouvertes.fr/hal-01840061

Research Reports

[21] C. Barillot, I. Fijalkow, I. Queinnec, F. Theoleyre, H. Touzet, M. Beaudouin-Lafon, M.-P. Cani, F. Chaumette, G. Conreur, V. Cortier, C. Gavoille, G. –. Gogniat, J. Gossa, A. Herzig, J. Krivine, P. Lamarre, F. Lamnabhi-Lagarrigue, A.-C. Letournel, L. Seinturier, I. Tellier, S. Torres, C. Truchet, A. Tsoukias, M. Verleysen.*Prospective Report of the scientific council of the Information Sciences and their interactions Institute of CNRS*, CNRS, December 2018, p. 1-87, https://hal.archives-ouvertes.fr/hal-01956087

Other Publications

[22] S. Nabil, F. Devernay, J. L. Crowley.*An objective quality metric for panoramic videos*, July 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01849261

Project-Team MAVERICK

Models and Algorithms for Visualization and Rendering

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

IN PARTNERSHIP WITH:
CNRS

Institut polytechnique de Grenoble

Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Interaction and visualization

Table of contents

Project-Team MAVERICK

Creation of the Team: 2012 January 01, updated into Project-Team: 2014 January 01

Keywords:

Computer Science and Digital Science:

A5.2. - Data visualization
A5.5. - Computer graphics
A5.5.1. - Geometrical modeling
A5.5.2. - Rendering
A5.5.3. - Computational photography
A5.5.4. - Animation

Other Research Topics and Application Domains:

B5.5. - Materials
B5.7. - 3D printing
B9.2.2. - Cinema, Television
B9.2.3. - Video games
B9.2.4. - Theater
B9.6.6. - Archeology, History

1. Team, Visitors, External Collaborators

Research Scientists

Nicolas Holzschuch [Team leader, Inria, Senior Researcher, HDR]
Fabrice Neyret [CNRS, Senior Researcher]
Cyril Soler [Inria, Researcher, HDR]

Faculty Members

Georges-Pierre Bonneau [Univ Grenoble Alpes, Professor]
Joëlle Thollot [Institut polytechnique de Grenoble, Associate Professor, HDR]
Romain Vergne [Univ Grenoble Alpes, Associate Professor]

PhD Students

Alexandre Bleron [Univ Grenoble Alpes, until Sep 2018]
Alban Fichet [Inria]
Morgane Gerardin [Univ Grenoble Alpes, from Oct 2018]
Guillaume Loubet [Inria, until May 2018]
Ronak Molazem [Inria, from Sep 2018]
Vincent Tavernier [Univ Grenoble Alpes]
Jérémy Wambecke [Univ Grenoble Alpes, until Sep 2018]
Sunrise Wang [Inria, from Sep 2018]

Visiting Scientist

Beibei Wang [Nanjing University of Science and Technology, from Apr 2018 until May 2018]

Administrative Assistant

Diane Courtiol [Inria]

2. Overall Objectives

2.1. Overall Objectives

Computer-generated pictures and videos are now ubiquitous: both for leisure activities, such as special effects in motion pictures, feature movies and video games, or for more serious activities, such as visualization and simulation.

Maverick was created as a research team in January 2012 and upgraded as a research project in January 2014. We deal with image synthesis methods. We place ourselves at the end of the image production pipeline, when the pictures are generated and displayed (see figure 1). We take many possible inputs: datasets, video flows, pictures and photographs, (animated) geometry from a virtual world... We produce as output pictures and videos.

These pictures will be viewed by humans, and we consider this fact as an important point of our research strategy, as it provides the benchmarks for evaluating our results: the pictures and animations produced must be able to convey the message to the viewer. The actual message depends on the specific application: data visualization, exploring virtual worlds, designing paintings and drawings... Our vision is that all these applications share common research problems: ensuring that the important features are perceived, avoiding cluttering or aliasing, efficient internal data representation, etc.

Computer Graphics, and especially Maverick is at the crossroad between fundamental research and industrial applications. We are both looking at the constraints and needs of applicative users and targeting long term research issues such as sampling and filtering.

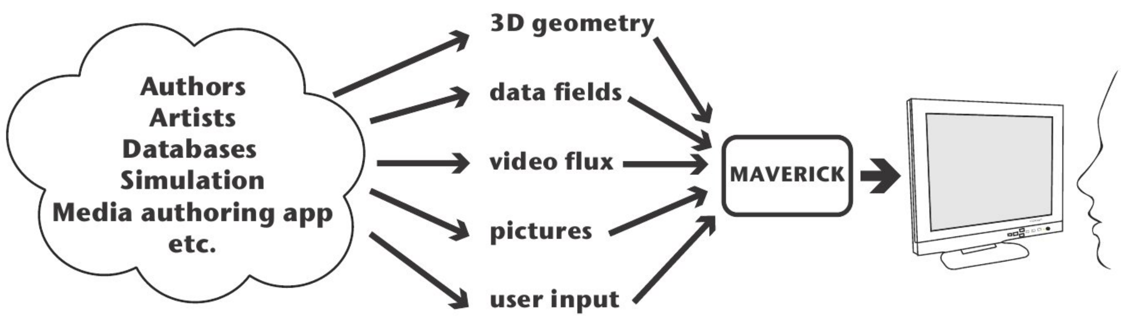


Figure 1. Position of the Maverick research team inside the graphics pipeline.

The Maverick project-team aims at producing representations and algorithms for efficient, high-quality computer generation of pictures and animations through the study of four *Research problems*:

- *Computer Visualization*, where we take as input a large localized dataset and represent it in a way that will let an observer understand its key properties,
- *Expressive Rendering*, where we create an artistic representation of a virtual world,
- *Illumination Simulation*, where our focus is modelling the interaction of light with the objects in the scene.
- *Complex Scenes*, where our focus is rendering and modelling highly complex scenes.

The heart of Maverick is *understanding* what makes a picture useful, powerful and interesting for the user, and designing algorithms to create these pictures.

We will address these research problems through three interconnected approaches:

- working on the *impact* of pictures, by conducting perceptual studies, measuring and removing artefacts and discontinuities, evaluating the user response to pictures and algorithms,
- developing *representations* for data, through abstraction, stylization and simplification,
- developing new methods for *predicting* the properties of a picture (*e.g.* frequency content, variations) and adapting our image-generation algorithm to these properties.

A fundamental element of the Maverick project-team is that the research problems and the scientific approaches are all cross-connected. Research on the *impact* of pictures is of interest in three different research problems: *Computer Visualization*, *Expressive rendering* and *Illumination Simulation*. Similarly, our research on *Illumination simulation* will gather contributions from all three scientific approaches: impact, representations and prediction.

3. Research Program

3.1. Introduction

The Maverick project-team aims at producing representations and algorithms for efficient, high-quality computer generation of pictures and animations through the study of four **research problems**:

- *Computer Visualization* where we take as input a large localized dataset and represent it in a way that will let an observer understand its key properties. Visualization can be used for data analysis, for the results of a simulation, for medical imaging data...
- *Expressive Rendering*, where we create an artistic representation of a virtual world. Expressive rendering corresponds to the generation of drawings or paintings of a virtual scene, but also to some areas of computational photography, where the picture is simplified in specific areas to focus the attention.
- *Illumination Simulation*, where we model the interaction of light with the objects in the scene, resulting in a photorealistic picture of the scene. Research include improving the quality and photorealism of pictures, including more complex effects such as depth-of-field or motion-blur. We are also working on accelerating the computations, both for real-time photorealistic rendering and offline, high-quality rendering.
- *Complex Scenes*, where we generate, manage, animate and render highly complex scenes, such as natural scenes with forests, rivers and oceans, but also large datasets for visualization. We are especially interested in interactive visualization of complex scenes, with all the associated challenges in terms of processing and memory bandwidth.

The fundamental research interest of Maverick is first, *understanding* what makes a picture useful, powerful and interesting for the user, and second *designing* algorithms to create and improve these pictures.

3.2. Research approaches

We will address these research problems through three interconnected research approaches:

3.2.1. Picture Impact

Our first research axis deals with the *impact* pictures have on the viewer, and how we can improve this impact. Our research here will target:

- *evaluating user response:* we need to evaluate how the viewers respond to the pictures and animations generated by our algorithms, through user studies, either asking the viewer about what he perceives in a picture or measuring how his body reacts (eye tracking, position tracking).
- *removing artefacts and discontinuities:* temporal and spatial discontinuities perturb viewer attention, distracting the viewer from the main message. These discontinuities occur during the picture creation process; finding and removing them is a difficult process.

3.2.2. Data Representation

The data we receive as input for picture generation is often unsuitable for interactive high-quality rendering: too many details, no spatial organisation... Similarly the pictures we produce or get as input for other algorithms can contain superfluous details.

One of our goals is to develop new data representations, adapted to our requirements for rendering. This includes fast access to the relevant information, but also access to the specific hierarchical level of information needed: we want to organize the data in hierarchical levels, pre-filter it so that sampling at a given level also gives information about the underlying levels. Our research for this axis include filtering, data abstraction, simplification and stylization.

The input data can be of any kind: geometric data, such as the model of an object, scientific data before visualization, pictures and photographs. It can be time-dependent or not; time-dependent data bring an additional level of challenge on the algorithm for fast updates.

3.2.3. Prediction and simulation

Our algorithms for generating pictures require computations: sampling, integration, simulation... These computations can be optimized if we already know the characteristics of the final picture. Our recent research has shown that it is possible to predict the local characteristics of a picture by studying the phenomena involved: the local complexity, the spatial variations, their direction...

Our goal is to develop new techniques for predicting the properties of a picture, and to adapt our image-generation algorithms to these properties, for example by sampling less in areas of low variation.

Our research problems and approaches are all cross-connected. Research on the *impact* of pictures is of interest in three different research problems: *Computer Visualization*, *Expressive rendering* and *Illumination Simulation*. Similarly, our research on *Illumination simulation* will use all three research approaches: impact, representations and prediction.

3.3. Cross-cutting research issues

Beyond the connections between our problems and research approaches, we are interested in several issues, which are present throughout all our research:

sampling is an ubiquitous process occurring in all our application domains, whether photorealistic rendering (*e.g.* photon mapping), expressive rendering (*e.g.* brush strokes), texturing, fluid simulation (Lagrangian methods), etc. When sampling and reconstructing a signal for picture generation, we have to ensure both coherence and homogeneity. By *coherence*, we mean not introducing spatial or temporal discontinuities in the reconstructed signal. By *homogeneity*, we mean that samples should be placed regularly in space and time. For a time-dependent signal, these requirements are conflicting with each other, opening new areas of research.

filtering is another ubiquitous process, occuring in all our application domains, whether in realistic rendering (*e.g.* for integrating height fields, normals, material properties), expressive rendering (*e.g.* for simplifying strokes), textures (through non-linearity and discontinuities). It is especially relevant when we are replacing a signal or data with a lower resolution (for hierarchical representation); this involves filtering the data with a reconstruction kernel, representing the transition between levels.

performance and scalability are also a common requirement for all our applications. We want our algorithms to be usable, which implies that they can be used on large and complex scenes, placing a great importance on scalability. For some applications, we target interactive and real-time applications, with an update frequency between 10 Hz and 120 Hz.

coherence and continuity in space and time is also a common requirement of realistic as well as expressive models which must be ensured despite contradictory requirements. We want to avoid flickering and aliasing.

animation: our input data is likely to be time-varying (*e.g.* animated geometry, physical simulation, time-dependent dataset). A common requirement for all our algorithms and data representation is that they must be compatible with animated data (fast updates for data structures, low latency algorithms...).

3.4. Methodology

Our research is guided by several methodological principles:

Experimentation: to find solutions and phenomenological models, we use experimentation, performing statistical measurements of how a system behaves. We then extract a model from the experimental data.

Validation: for each algorithm we develop, we look for experimental validation: measuring the behavior of the algorithm, how it scales, how it improves over the state-of-the-art... We also compare our algorithms to the exact solution. Validation is harder for some of our research domains, but it remains a key principle for us.

Reducing the complexity of the problem: the equations describing certain behaviors in image synthesis can have a large degree of complexity, precluding computations, especially in real time. This is true for physical simulation of fluids, tree growth, illumination simulation... We are looking for *emerging phenomena* and *phenomenological models* to describe them (see framed box "Emerging phenomena"). Using these, we simplify the theoretical models in a controlled way, to improve user interaction and accelerate the computations.

Transferring ideas from other domains: Computer Graphics is, by nature, at the interface of many research domains: physics for the behavior of light, applied mathematics for numerical simulation, biology, algorithmics... We import tools from all these domains, and keep looking for new tools and ideas.

Develop new fondamental tools: In situations where specific tools are required for a problem, we will proceed from a theoretical framework to develop them. These tools may in return have applications in other domains, and we are ready to disseminate them.

Collaborate with industrial partners: we have a long experiment of collaboration with industrial partners. These collaborations bring us new problems to solve, with short-term or medium-term transfert opportunities. When we cooperate with these partners, we have to find *what they need*, which can be very different from *what they want*, their expressed need.

4. Application Domains

4.1. Application Domains

The natural application domain for our research is the production of digital images, for example for movies and special effects, virtual prototyping, video games... Our research have also been applied to tools for generating and editing images and textures, for example generating textures for maps. Our current application domains are:

- Offline and real-time rendering in movie special effects and video games;
- Virtual prototyping;
- Scientific visualization;
- Content modeling and generation (e.g. generating texture for video games, capturing reflectance properties, etc);
- Image creation and manipulation.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- the paper "MNPR: A Framework for Real-Time Expressive Non-Photorealistic Rendering of 3D Computer Graphics" [13], co-authored by Santiago Montesdeoca, Hock Soon Seah, Amir Semmo, Pierre Bénard, Romain Vergne, Joëlle Thollot and Davide Benvenuti, has received the "Best Paper Award" during the conference Expressive 2018.
- the paper "High-performance By-Example Noise using a Histogram-Preserving Blending Operator" [4], co-authored by Eric Heitz and Fabrice Neyret, has received the "Best Paper Award" during the conference High-performance Graphics 2018.
- the paper "A New Microflake Model with Microscopic Self-Shadowing for Accurate Volume Downsampling" [5], co-authored by Guillaume Loubet and Fabrice Neyret, has received the "Best Paper Award" during the conference Eurographics 2018.

6. New Software and Platforms

6.1. Diffusion curves

KEYWORDS: Vector-based drawing - Shading
FUNCTIONAL DESCRIPTION: Diffusion Curves is a vector-based design tool for creating complex shaded images. This prototype is composed of the Windows binary, along with the required shader programs (ie. in source code).

- Participants: Adrien Bousseau, Alexandrina Orzan, David Salesin, Holger Winnemoeller, Joëlle Thollot and Pascal Barla
- Partners: CNRS - LJK - INP Grenoble - Université Joseph-Fourier
- Contact: Joëlle Thollot
- URL: http://maverick.inria.fr/Publications/2008/OBWBTS08/index.php

6.2. GigaVoxels

FUNCTIONAL DESCRIPTION: Gigavoxel is a software platform which goal is the real-time quality rendering of very large and very detailed scenes which couldn't fit memory. Performances permit showing details over deep zooms and walk through very crowdy scenes (which are rigid, for the moment). The principle is to represent data on the GPU as a Sparse Voxel Octree which multiscale voxels bricks are produced on demand only when necessary and only at the required resolution, and kept in a LRU cache. User defined producer lays accross CPU and GPU and can load, transform, or procedurally create the data. Another user defined function is called to shade each voxel according to the user-defined voxel content, so that it is user choice to distribute the appearance-making at creation (for faster rendering) or on the fly (for storageless thin procedural details). The efficient rendering is done using a GPU differential cone-tracing using the scale corresponding to the 3D-MIPmapping LOD, allowing quality rendering with one single ray per pixel. Data is produced in case of cache miss, and thus only whenever visible (accounting for view frustum and occlusion). Soft-shadows and depth-of-field is easily obtained using larger cones, and are indeed cheaper than unblurred rendering. Beside the representation, data management and base rendering algorithm themself, we also worked on realtime light transport, and on quality prefiltering of complex data. Ongoing researches are addressing animation. GigaVoxels is currently used for the quality real-time exploration of the detailed galaxy in ANR RTIGE. Most of the work published by Cyril Crassin (and al.) during his PhD (see http://maverick.inria.fr/Members/Cyril.Crassin/) is related to GigaVoxels. GigaVoxels is available for Windows and Linux under the BSD-3 licence.

- Participants: Cyril Crassin, Eric Heitz, Fabrice Neyret, Jérémy Sinoir, Pascal Guehl and Prashant Goswami
- Contact: Fabrice Neyret
- URL: http://gigavoxels.inrialpes.fr

6.3. GRATIN

FUNCTIONAL DESCRIPTION: Gratin is a node-based compositing software for creating, manipulating and animating 2D and 3D data. It uses an internal direct acyclic multi-graph and provides an intuitive user interface that allows to quickly design complex prototypes. Gratin has several properties that make it useful for researchers and students. (1) it works in real-time: everything is executed on the GPU, using OpenGL, GLSL and/or Cuda. (2) it is easily programmable: users can directly write GLSL scripts inside the interface, or create new C++ plugins that will be loaded as new nodes in the software. (3) all the parameters can be animated using keyframe curves to generate videos and demos. (4) the system allows to easily exchange nodes, group of nodes or full pipelines between people.

- Participants: Pascal Barla and Romain Vergne
- Partner: UJF
- Contact: Romain Vergne
- URL: http://gratin.gforge.inria.fr/

6.4. HQR

High Quality Renderer
KEYWORDS: Lighting simulation - Materials - Plug-in
FUNCTIONAL DESCRIPTION: HQR is a global lighting simulation platform. HQR software is based on the photon mapping method which is capable of solving the light balance equation and of giving a high quality solution. Through a graphical user interface, it reads X3D scenes using the X3DToolKit package developed at Maverick, it allows the user to tune several parameters, computes photon maps, and reconstructs information to obtain a high quality solution. HQR also accepts plugins which considerably eases the developpement of new algorithms for global illumination, those benefiting from the existing algorithms for handling materials, geometry and light sources.

- Participant: Cyril Soler
- Contact: Cyril Soler
- URL: http://artis.imag.fr/~Cyril.Soler/HQR

6.5. libylm

LibYLM
KEYWORD: Spherical harmonics
FUNCTIONAL DESCRIPTION: This library implements spherical and zonal harmonics. It provides the means to perform decompositions, manipulate spherical harmonic distributions and provides its own viewer to visualize spherical harmonic distributions.

- Author: Cyril Soler
- Contact: Cyril Soler
- URL: https://launchpad.net/~csoler-users/+archive/ubuntu/ylm

6.6. MobiNet

KEYWORDS: Co-simulation - Education - Programmation

FUNCTIONAL DESCRIPTION: The MobiNet software allows for the creation of simple applications such as video games, virtual physics experiments or pedagogical math illustrations. It relies on an intuitive graphical interface and language which allows the user to program a set of mobile objects (possibly through a network). It is available in public domain for Linux,Windows and MacOS.

- Participants: Fabrice Neyret, Franck Hétroy-Wheeler, Joëlle Thollot, Samuel Hornus and Sylvain Lefebvre
- Partners: CNRS - LJK - INP Grenoble - Inria - IREM - Cies - GRAVIR
- Contact: Fabrice Neyret
- URL: http://mobinet.imag.fr/index.en.html

6.7. PLANTRAD

KEYWORDS: Bioinformatics - Biology
FUNCTIONAL DESCRIPTION: PlantRad is a software program for computing solutions to the equation of light equilibrium in a complex scene including vegetation. The technology used is hierarchical radiosity with clustering and instantiation. Thanks to the latter, PlantRad is capable of treating scenes with a very high geometric complexity (up to millions of polygons) such as plants or any kind of vegetation scene where a high degree of approximate self-similarity permits a significant gain in memory requirements.

- Participants: Cyril Soler, François Sillion and George Drettakis
- Contact: Cyril Soler

6.8. PROLAND

PROcedural LANDscape
KEYWORDS: Real time - 3D - Realistic rendering - Masses of data - Atmosphere - Ocean
FUNCTIONAL DESCRIPTION: The goal of this platform is the real-time quality rendering and editing of large landscapes. All features can work with planet-sized terrains, for all viewpoints from ground to space. Most of the work published by Eric Bruneton and Fabrice Neyret (see http://evasion.inrialpes.fr/Membres/Eric.Bruneton/) has been done within Proland and integrated in the main branch. Proland is available under the BSD-3 licence.

- Participants: Antoine Begault, Eric Bruneton, Fabrice Neyret and Guillaume Piolet
- Contact: Fabrice Neyret
- URL: https://proland.inrialpes.fr/

6.9. ShwarpIt

KEYWORD: Warping
FUNCTIONAL DESCRIPTION: ShwarpIt is a simple mobile app that allows you to manipulate the perception of shapes in images. Slide the ShwarpIt slider to the right to make shapes appear rounder. Slide it to the left to make shapes appear more flat. The Scale slider gives you control on the scale of the warping deformation.

- Contact: Georges-Pierre Bonneau
- URL: http://bonneau.meylan.free.fr/ShwarpIt/ShwarpIt.html

6.10. Vrender

FUNCTIONAL DESCRIPTION: The VRender library is a simple tool to render the content of an OpenGL window to a vectorial device such as Postscript, XFig, and soon SVG. The main usage of such a library is to make clean vectorial drawings for publications, books, etc.

In practice, VRender replaces the z-buffer based hidden surface removal of OpenGL by sorting the geometric primitives so that they can be rendered in a back-to-front order, possibly cutting them into pieces to solve cycles.

VRender is also responsible for the vectorial snapshot feature of the QGLViewer library.

- Participant: Cyril Soler
- Contact: Cyril Soler
- URL: http://artis.imag.fr/Software/VRender/

6.11. X3D TOOLKIT

X3D Development pateform
FUNCTIONAL DESCRIPTION: X3DToolkit is a library to parse and write X3D files, that supports plugins and extensions.

- Participants: Gilles Debunne and Yannick Le Goc
- Contact: Cyril Soler
- URL: http://artis.imag.fr/Software/X3D/

7. New Results

7.1. Expressive Rendering

7.1.1. A workflow for designing stylized shading effects

Participants: Alexandre Bléron, Romain Vergne, Thomas Hurtut, Joëlle Thollot.

In this report [18], we describe a workflow for designing stylized shading effects on a 3D object, targeted at technical artists. Shading design, the process of making the illumination of an object in a 3D scene match an artist vision, is usually a time-consuming task because of the complex interactions between materials, geometry, and lighting environment. Physically based methods tend to provide an intuitive and coherent workflow for artists, but they are of limited use in the context of non-photorealistic shading styles. On the other hand, existing stylized shading techniques are either too specialized or require considerable hand-tuning of unintuitive parameters to give a satisfactory result. Our contribution is to separate the design process of individual shading effects in three independent stages: control of its global behavior on the object, addition of procedural details, and colorization. Inspired by the formulation of existing shading models, we expose different shading behaviors to the artist through parametrizations, which have a meaningful visual interpretation. Multiple shading effects can then be composited to obtain complex dynamic appearances. The proposed workflow is fully interactive, with real-time feedback, and allows the intuitive exploration of stylized shading effects, while keeping coherence under varying viewpoints and light configurations (see Fig. 2). Furthermore, our method makes use of the deferred shading technique, making it easily integrable in existing rendering pipelines.

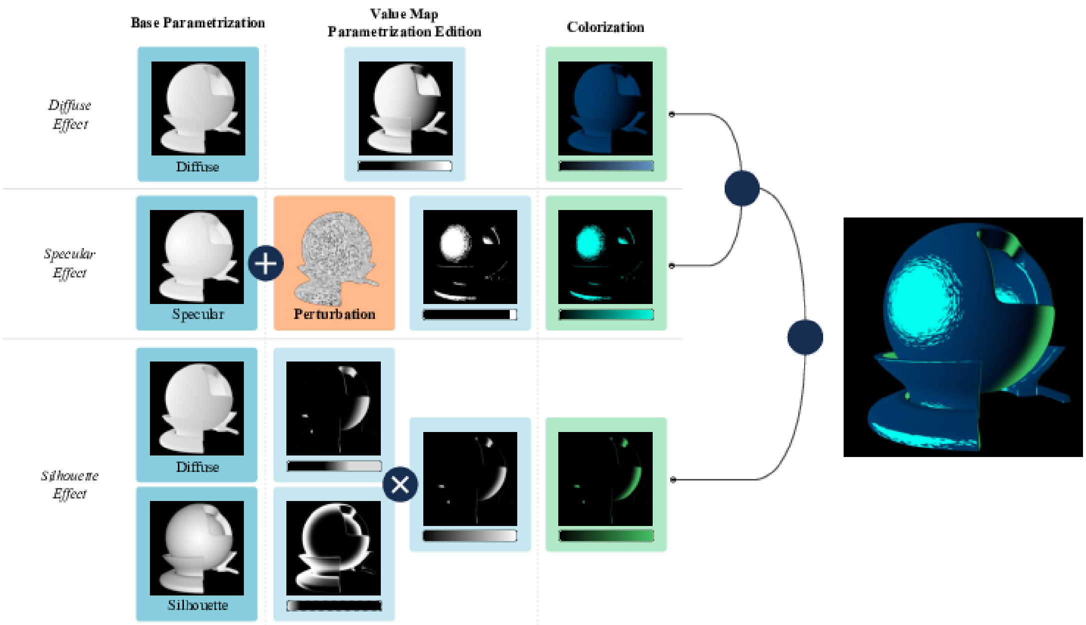


Figure 2. Illustration of our workflow showing an example with three appearance effects. A user can modify and combine base parametrizations to design the shading behavior (blue nodes) of an appearance effect, using value maps and combination operations. A color map (green nodes) is then applied on the designed behavior to colorize the effect. Output effects are then composited to obtain the final appearance. Perturbations (orange nodes) can be attached to every operation in order to add procedural details to an effect. The orientation of the perturbation can be controlled by the gradient of a shading behavior (as shown here), or by an external vector field, such as a tangent map.

7.1.2. MNPR: A framework for real-time expressive non-photorealistic rendering of 3D computer graphics

Participants: Santiago Montesdeoca, Hock Soon Seah, Amir Semmo, Pierre Bénard, Romain Vergne, Joëlle Thollot, Davide Benvenuti.

We propose a framework for expressive non-photorealistic rendering of 3D computer graphics: MNPR. Our work focuses on enabling stylization pipelines with a wide range of control, thereby covering the interaction spectrum with real-time feedback. In addition, we introduce control semantics that allow cross-stylistic art-direction, which is demonstrated through our implemented watercolor, oil and charcoal stylizations (see Fig. 3). Our generalized control semantics and their style-specific mappings are designed to be extrapolated to other styles, by adhering to the same control scheme. We then share our implementation details by breaking down our framework and elaborating on its inner workings. Finally, we evaluate the usefulness of each level of control through a user study involving 20 experienced artists and engineers in the industry, who have collectively spent over 245 hours using our system. MNPR is implemented in Autodesk Maya and open-sourced through this publication, to facilitate adoption by artists and further development by the expressive research and development community. This paper was presented at Expressive [13] and received the best paper award.

Figure 3. A scene rendered through MNPR in different styles. Baba Yaga's hut model, © Inuciian.

7.1.3. Motion-coherent stylization with screen-space image filters

Participants: Alexandre Bléron, Romain Vergne, Thomas Hurtut, Joëlle Thollot.

One of the qualities sought in expressive rendering is the 2D impression of the resulting style, called flatness. In the context of 3D scenes, screen-space stylization techniques are good candidates for flatness as they operate in the 2D image plane, after the scene has been rendered into G-buffers. Various stylization filters can be applied in screen-space while making use of the geometrical information contained in G-buffers to ensure motion coherence. However, this means that filtering can only be done inside the rasterized surface of the object. This can be detrimental to some styles that require irregular silhouettes to be convincing. In this paper, we describe a post-processing pipeline that allows stylization filters to extend outside the rasterized footprint of the object by locally *inflating* the data contained in G-buffers (see Fig. 4). This pipeline is fully implemented on the GPU and can be evaluated at interactive rates. We show how common image filtering techniques, when integrated in our pipeline and in combination with G-buffer data, can be used to reproduce a wide range of *digitally-painted* appearances, such as directed brush strokes with irregular silhouettes, while keeping enough motion coherence. This paper was presented at Expressive [11].

7.2. Illumination simulation and materials

7.2.1. Rendering homogeneous participating media

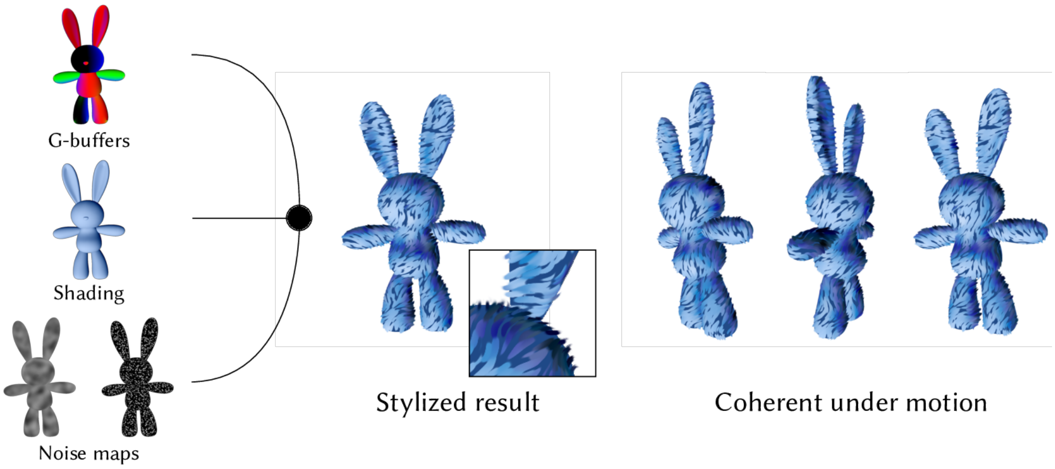


Figure 4. Using standard G-buffers and auxiliary buffers (noise, shading) as input, our pipeline can reproduce stylization effects that extend outside the original rasterized footprint of the object. Visual features produced by the filters stay coherent under motion or viewpoint changes.

Participants: Beibei Wang, Nicolas Holzschuch, Liangsheng Ge, Lu Wang.

Illumination effects in translucent materials are a combination of several physical phenomena: refraction at the surface, absorption and scattering inside the material. Because refraction can focus light deep inside the material, where it will be scattered, practical illumination simulation inside translucent materials is difficult. We have worked on a Point-Based Global Illumination method for light transport on homogeneous translucent materials with refractive boundaries. We start by placing light samples inside the translucent material and organizing them into a spatial hierarchy. At rendering, we gather light from these samples for each camera ray. We compute separately the sample contributions for single, double and multiple scattering, and add them. Multiple scattering effects are precomputed and stores in a table, accessed at runtime. An illustration of our approach is given in Fig 5. We present two implementations of our algorithm: an offline version for high-quality rendering and an interactive GPU implementation. The offline version provides significant speed-ups and reduced memory footprints compared to state-of-the-art algorithms, with no visible impact on quality. The GPU version yields interactive frame rates: 30 fps when moving the viewpoint, 25 fps when editing the light position or the material parameters. This work was published in IEEE Transactions on Visualization and Computer Graphics [9].

Storing the precomputed table for these multiple scattering effects is the largest memory cost for this algorithm. In a separate work, we used a neural network to encode these effects. We replaced the precomputed multiple scattering table with a trained neural network, with a cost of 6490 bytes (1623 floats). At runtime, the neural network is used to generate multiple scattering. The approach can be combined with many rendering algorithms, as illustrated in Fig. 6. This work was published as a Siggraph Talk [12].

7.2.2. Fast global illumination with discrete stochastic microfacets using a filterable model

Participants: Beibei Wang, Lu Wang, Nicolas Holzschuch.

Many real-life materials have a sparkling appearance, whether by design or by nature. Examples include metallic paints, sparkling varnish but also snow. These sparkles correspond to small, isolated, shiny particles reflecting light in a specific direction, on the surface or embedded inside the material. The particles responsible for these sparkles are usually small and discontinuous. These characteristics make it difficult to integrate them

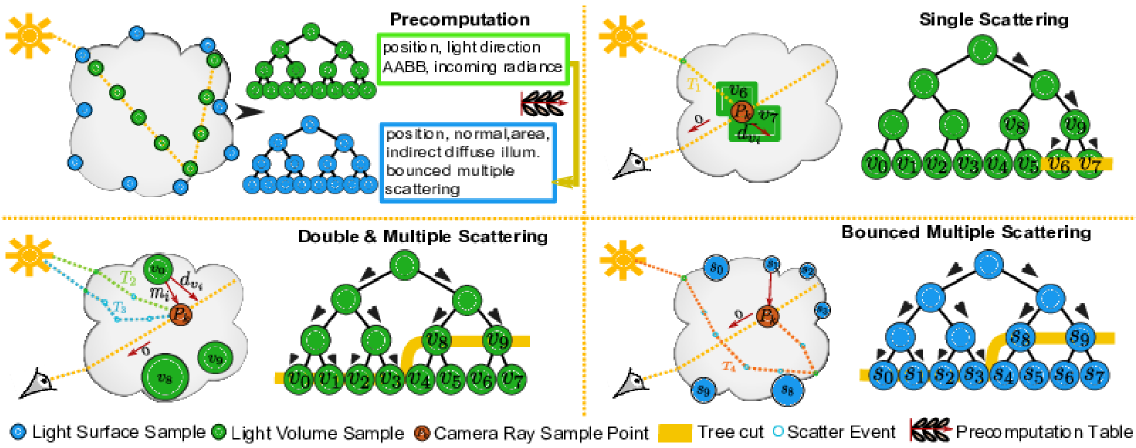


Figure 5. Our algorithm: we begin by computing incoming light at volume and surface samples. We then compute Single-, Double- and Multiple scattering effects for each camera ray using these volume and surface samples.

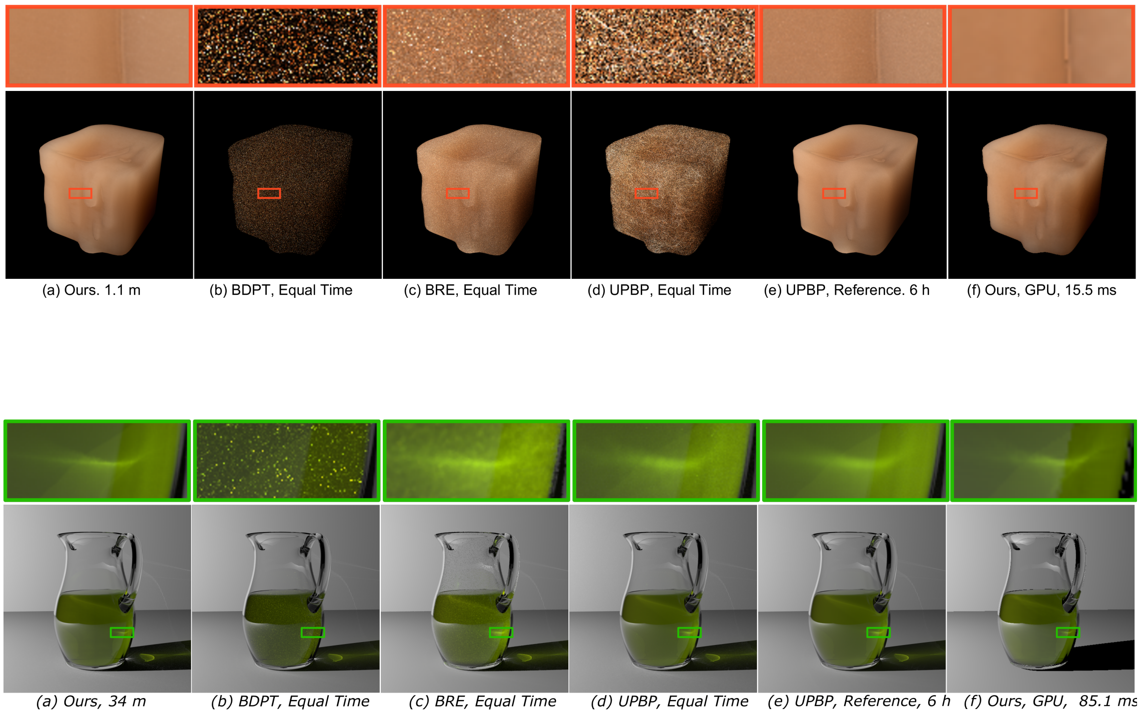


Figure 6. comparison between our algorithm, other algorithms with equal time or equal quality and reference images. Top row: wax. For this material, with a large albedo and a small mean free path, multiple scattering effects dominate. Bottom row: olive oil. For this material with low albedo and large mean-free-path, low-order scattering effects dominate.

efficiently in a standard rendering pipeline, especially for indirect illumination. Existing approaches use a 4-dimensional hierarchy, searching for light-reflecting particles simultaneously in space and direction. The approach is accurate, but still expensive. We have shown that this 4-dimensional search can be approximated using separate 2-dimensional steps. This approximation allows fast integration of glint contributions for large footprints, reducing the extra cost associated with glints be an order of magnitude, as illustrated in Fig. 7. This work was published in Computer Graphics Forum and presented at the Pacific Graphics conference [10].

(a) *Our model (with filtering)* Total: 6.09 h, Cost: +1.5%
(b) *Our model (no filter),* Total: 6.35 h, Cost: +6%
(c) *Original [JHY*14],* Total: 8.96 h, Cost: +49%
(d) *Without Glints,* Total: 6.0 h

Figure 7. Our algorithm, compared to the original Discrete Stochastic Microfacets model (c). Converting the 4D search to a product of 2D searches (b) produces almost identical results. This is the basis for our filterable model (a), which allows fast global illumination with negligible cost..

7.2.3. Handling fluorescence in a uni-directional spectral path tracer

Participants: Michal Mojzík, Alban Fichet, Alexander Wilkie

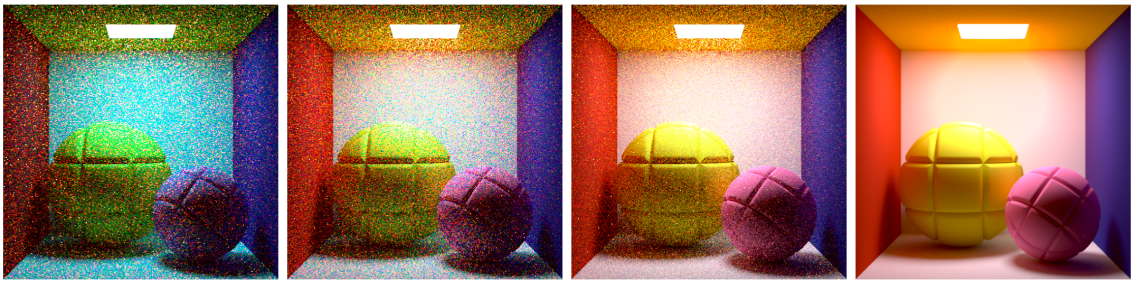
We present two separate improvements to the handling of fluorescence effects in modern uni-directional spectral rendering systems.

The first is the formulation of a new distance tracking scheme for fluorescent volume materials which exhibit a pronounced wavelength asymmetry. Such volumetric materials are an important and not uncommon corner case of wavelength-shifting media behaviour, and have not been addressed so far in rendering literature. This new tracking scheme (figure 8(b)) converges faster than a simple modification that can be added to the traditional exponential tracking (figure 8(a)).

The second one is that we introduce an extension of Hero wavelength sampling which can handle fluorescence events, both on surfaces, and in volumes. Both improvements are useful by themselves, and can be used separately: when used together, they enable the robust inclusion of arbitrary fluorescence effects in modern uni-directional spectral MIS path tracers (figure 8(c)). Our extension of Hero wavelength sampling is generally useful, while our proposed technique for distance tracking in strongly asymmetric media is admittedly not very efficient. However, it makes the most of a rather difficult situation, and at least allows the inclusion of such media in uni-directional path tracers, albeit at comparatively high cost. Which is still an improvement since up to now, their inclusion was not really possible at all, due to the inability of conventional tracking schemes to generate sampling points in such volume materials. This work was published in the journal Computer Graphics Forum [6].

7.2.4. A versatile parameterization for measured material manifolds

Participants: Cyril Soler, Kartic Subr, Derek Nowrouzezahrai.



(a) max(σ_t, σ_s, 16 spp, mono (b) Fluorescent-aware, 16 spp, mono (c) Fluo-aware, 16spp, HWSS (d) Fluo-aware, 65000 spp, HWSS

Figure 8. Comparison of proposed techniques to improve rendering of fluorescence.

A popular approach for computing photorealistic images of virtual objects requires applying reflectance profiles measured from real surfaces, introducing several challenges: the memory needed to faithfully capture realistic material reflectance is large, the choice of materials is limited to the set of measurements, and image synthesis using the measured data is costly. Typically, this data is either compressed by projecting it onto a subset of its linear principal components or by applying non-linear methods. The former requires many components to faithfully represent the input reflectance, whereas the latter necessitates costly extrapolation algorithms. We learn an underlying, low-dimensional non-linear reflectance manifold amenable to rapid exploration and rendering of real-world materials. We can express interpolated materials as linear combinations of the measured data, despite them lying on an inherently non-linear manifold. This allows us to efficiently interpolate and extrapolate measured BRDFs, and to render directly from the manifold representation. We exploit properties of Gaussian process latent variable models and use our representation for high-performance and offline rendering with interpolated real-world materials. This work has been published in the journal Computer Graphics Forum [7], and presented at Eurographics 2018.

7.3. Complex scenes

7.3.1. A new microflake model with microscopic self-shadowing for accurate volume downsampling

Participants: Guillaume Loubet, Fabrice Neyret.

In this work, we addressed the problem of representing the effect of internal self-shadowing in elements about to be filtered out at a given LOD, in the scope of volume of voxels containing density and phase-function (represented by a microflakes).

Naïve linear methods for downsampling high resolution microflake volumes often produce inaccurate results, especially when input voxels are very opaque. Preserving correct appearance at all resolutions requires taking into account inter- and intravoxel self-shadowing effects (see Figure 10). We introduce a new microflake model whose parameters characterize self-shadowing effects at the microscopic scale. We provide an anisotropic self-shadowing function and a microflake distribution for which scattering coefficients and phase functions of our model have closed-form expressions. We use this model in a new downsampling approach in which scattering parameters are computed from local estimations of self-shadowing in the input volume. Unlike previous work, our method handles datasets with spatially varying scattering parameters, semi-transparent volumes and datasets with intricate silhouettes. We show that our method generates LoDs with correct transparency and consistent appearance through scales for a wide range of challenging datasets, allowing for huge memory savings and efficient distant rendering without loss of quality. This work received the Best Paper Award at Eurographics 2018 and was published in the journal Computer Graphics Forum [5].

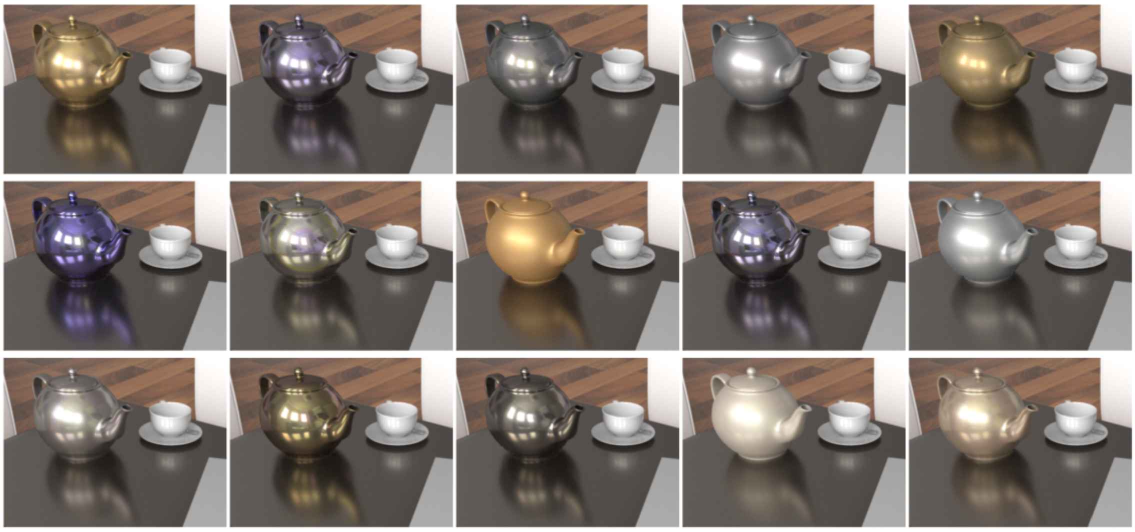


Figure 9. Four of the images above (Number 2, 4, 6 and 12 in reading order) are rendered with measured BRDFs from the MERL dataset, the remaining 11 being rendered with BRDFs randomly picked from our parameterization of the non-linear manifold containing MERL materials. We explore this manifold interactively to produce high-quality BRDFs which retain the physical properties and perceptual aspect of real materials.

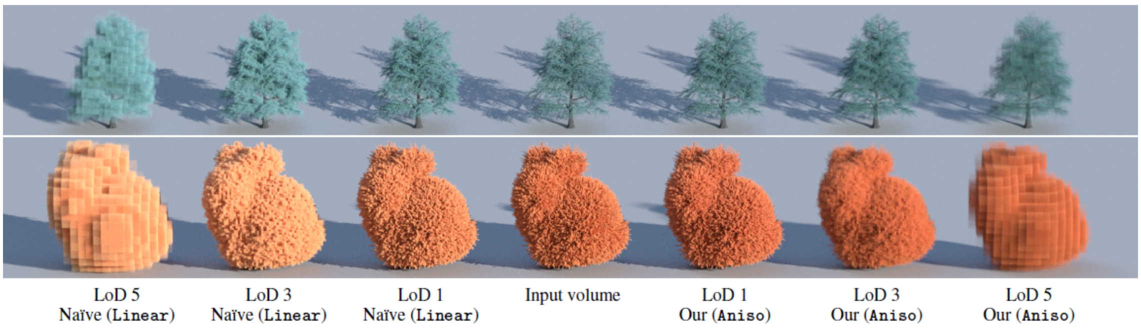


Figure 10. Comparison between naïve downsampling of microflake volumes and our method ("Aniso"). Naïve dowsampling of volumes with dense voxels often lead to inaccurate results due to the loss of inter- and intra-voxel self-shadowing effects. Our method is based on a new participating medium model and on local estimations of self-shadowing. It generates LoDs with correct transparency and consistent appearance through scales. Rendered with volume path tracing in Mitsuba (http://www.mitsuba-renderer.org/): the trunk of the cedar is a mesh.

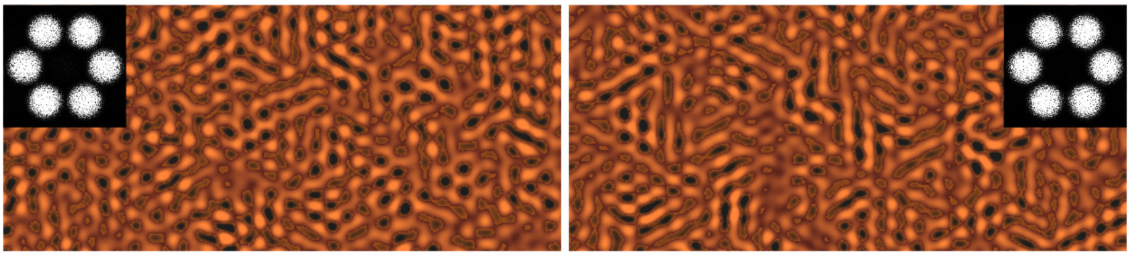
7.4. Texture synthesis

7.4.1. Gabor noise revisited

Participants: Vincent Tavernier, Fabrice Neyret, Romain Vergne, Joëlle Thollot.

Gabor Noise is a powerful procedural texture synthesis technique, but has two major drawbacks: It is costly due to the high required splat density and not always predictable because properties of instances can differ from those of the process. We bench performance and quality using alternatives for each Gabor Noise ingredient: point distribution, kernel weighting and kernel shape. For this, we introduce 3 objective criteria to measure process convergence, process stationarity, and instance stationarity. We show that minor implementation changes allow for $17-24\times$ speed-up with same or better quality (see Fig. 11).

This paper was presented at AFIG [17] and received the best paper award. An article has been submitted to Eurographics-short 2019.



(a) Seminal Gabor, $N = 45$ (b) Bernoulli+strat.+sin, $N = 3$

Figure 11. Real case with complex power spectrum (3 kernels, cf. inset) and non-linear post-treatment. Our optimized set of ingredients achieves the same visual quality in $1/17^{th}$ of the time required by the seminal method.

7.4.2. High-performance by-example noise using a histogram-preserving blending operator

Participants: Eric Heitz, Fabrice Neyret.

We propose a new by-example noise algorithm that takes as input a small example of a stochastic texture and synthesizes an infinite output with the same appearance. It works on any kind of random-phase inputs as well as on many non-random-phase inputs that are stochastic and non-periodic, typically natural textures such as moss, granite, sand, bark, etc. Our algorithm achieves high-quality results comparable to state-of-the-art procedural-noise techniques but is more than 20 times faster. Our approach is conceptually simple: we partition the output texture space on a triangle grid and associate each vertex with a random patch from the input such that the evaluation inside a triangle is done by blending 3 patches. The key to this approach is the blending operation that usually produces visual artifacts such as ghosting, softened discontinuities and reduced contrast, or introduces new colors not present in the input. We analyze these problems by showing how linear blending impacts the histogram and show that a blending operator that preserves the histogram prevents these problems. The main requirement for a rendering application is to implement such an operator in a fragment shader without further post-processing, i.e. we need a histogram-preserving blending operator that operates only at the pixel level. Our insight for the design of this operator is that, with Gaussian inputs, histogram-preserving blending boils down to mean and variance preservation, which is simple to obtain analytically. We extend this idea to non-Gaussian inputs by "Gaussianizing" them with a histogram transformation and "de-Gaussianizing" them with the inverse transformation after the blending operation. We show how to precompute and store these histogram transformations such that our algorithm can be implemented in a fragment shader, as illustrated in Fig. 12. This work received the Best Paper Award at High Performance Graphics 2018 [4].

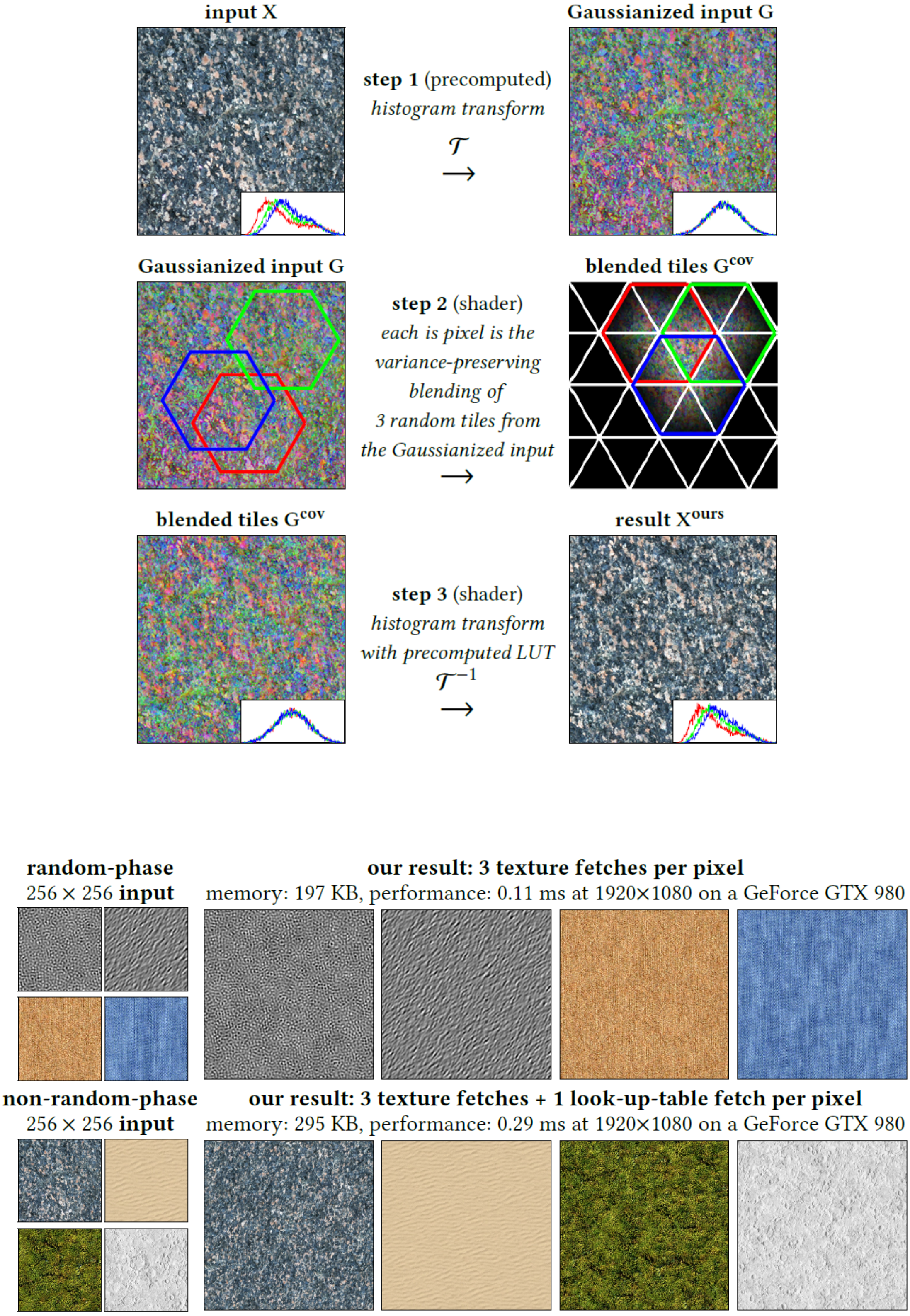


Figure 12. Top: method overview. Bottom: results and performances.

7.5. Visualization

7.5.1. A "What if" approach for eco-feedback

Participants: Jérémy Wambecke, Georges-Pierre Bonneau, Romain Vergne, Renaud Blanch.

Many households share the objective of reducing electricity consumption for either economic or ecological motivations. Eco-feedback technologies support this objective by providing users with a visualization of their consumption. However as pointed out by several studies, users encounter difficulties in finding concrete actions to reduce their consumption. To overcome this limitation, we introduce and evaluate Activelec, a system based on the visualization and interaction with user's behavior rather than raw consumption data. The user's behavior is modeled as the set of actions modifying the state of appliances over time. A key novelty of our solution is its focus on the What if approach applied to eco-feedback. Users can analyze and experiment scenarios by selecting and modifying their usage of electrical appliances over time and visualize the impact on the consumption, as illustrated in Fig. 13. In [16] we conducted two laboratory user studies that evaluate the usability of Activelec and the relevance of the What if approach for electricity consumption. Our results show that users understand the interaction paradigm and can easily find relevant modifications in their usage of appliances. Moreover participants judge these changes of behavior would require little effort to be adopted. In [15] we conducted an in-situ evaluation of Activelec, confirming these results in a real setting.

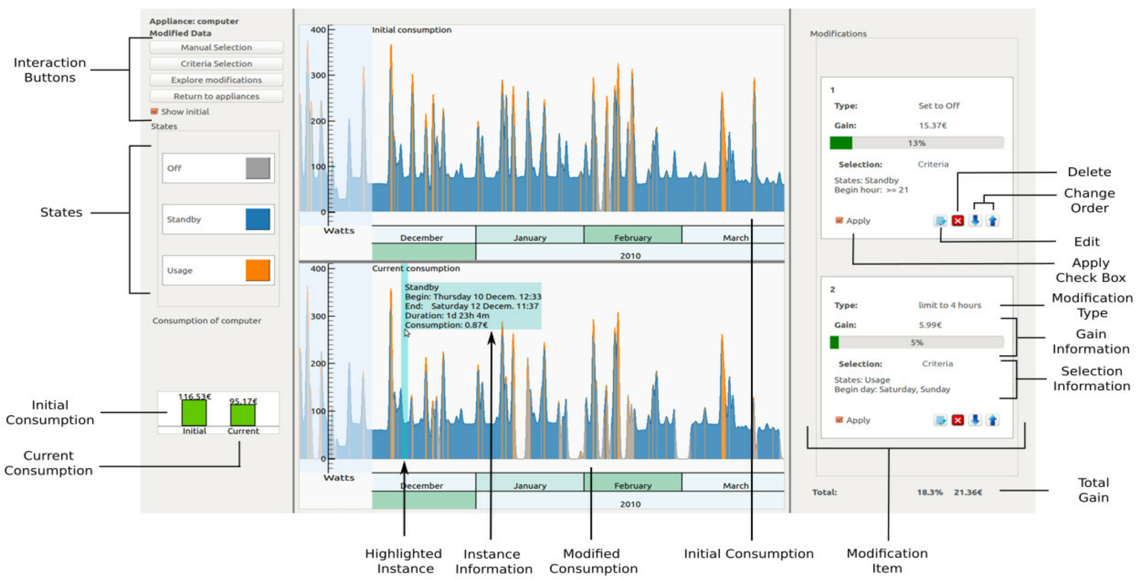


Figure 13. Interface of our system. A computer has been chosen by the user, whose states are Usage (orange) and Standby (blue). At the right, we can see that the user has applied two modifications, the first one to remove instances of Standby after 9 P.M, and the second one to limit the instances of On to 4 hours during the weekend. When the user is selecting instances, this panel displays information about the selection.

7.5.2. Morphorider: a new way for Structural Monitoring via Shape Acquisition

Participants: Tibor Stanko, Laurent Jouanet, Nathalie Saguin-Sprynski, Georges-Pierre Bonneau, Stefanie Hahmann.

In collaboration with CEA-Leti we introduce a new kind of monitoring device, illustrated in Fig. 14, allowing the shape acquisition of a structure via a single mobile node of inertial sensors and an odometer. Previous approaches used devices placed along a network with fixed connectivity between the sensor nodes (lines, grid). When placed onto a shape, this sensor network provides local surface orientations along a curve network on the shape, but its absolute position in the world space is unknown. The new mobile device provides a novel way of structures monitoring: the shape can be scanned regularly, and following the shape or some specific parameters along time may afford the detection of early signs of failure. Here, we present a complete framework for 3D shape reconstruction. To compute the shape, our main insight is to formulate the reconstruction as a set of optimization problems. Using discrete representations, these optimization problems are resolved efficiently and at interactive time rates. We present two main contributions. First, we introduce a novel method for creating well-connected networks with cell-complex topology using only orientation and distance measurements and a set of user-defined constraints. Second, we address the problem of surfacing a closed 3D curve network with given surface normals. The normal input increases shape fidelity and allows to achieve globally smooth and visually pleasing shapes. The proposed framework was tested on experimental data sets acquired using our device. A quantitative evaluation was performed by computing the error of reconstruction for our own designed surfaces, thus with known ground truth. Even for complex shapes, the mean error remains around 1%. This work was published at the 9th European Workshop on Structural Health Monitoring [14].

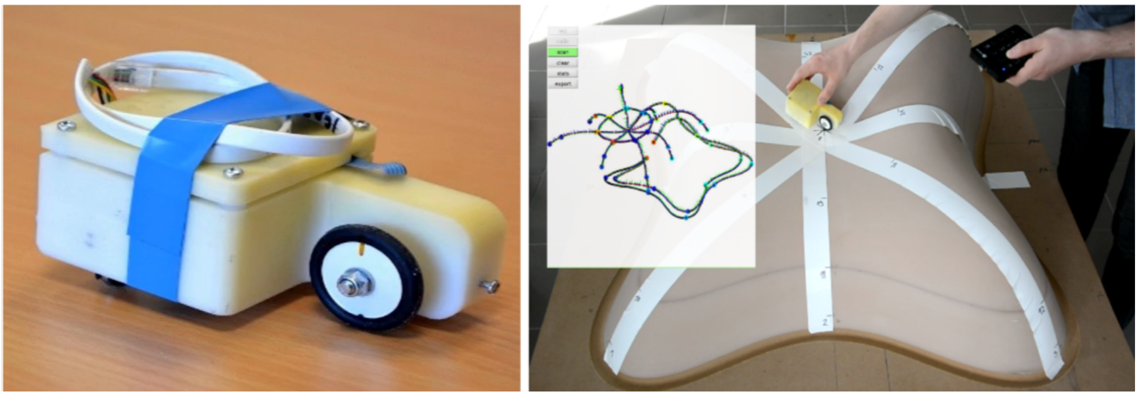


Figure 14. Morphorider: Structural Monitoring via Shape Acquisition (right) with a mobile device (left) equipped with an inertial node of sensors and an odometer.

8. Partnerships and Cooperations

8.1. Regional Initiatives

We have frequent exchanges and on-going collaborations with Cyril Crassin from nVIDIA-Research, and Eric Heitz, Laurent Belcour, Jonathan Dupuy and Kenneth Vanhoey from Unity-Research. Maverick is part of the GPU Research Center labeled by nVIDIA at Inria Grenoble. Team contact: Fabrice Neyret.

8.2. National Initiatives

8.2.1. ANR: Materials

Participants: Nicolas Holzschuch [contact], Romain Vergne.

We are funded by the ANR for a joint research project on acquisition and restitution of micro-facet based materials. This project is in cooperation with Océ Print Logic technologies, the Museum of Ethnography at the University of Bordeaux and the Manao team at Inria Bordeaux. The grant started in October 2015, for 48 months.

8.2.2. CDP: Patrimalp 2.0

Participants: Nicolas Holzschuch [contact], Romain Vergne.

The main objective and challenge of Patrimalp 2.0 is to develop a cross-disciplinary approach in order to get a better knowledge of the material cultural heritage in order to ensure its sustainability, valorization and diffusion in society. Carried out by members of UGA laboratories, combining skills in human sciences, geosciences, digital engineering, material sciences, in close connection with stakeholders of heritage and cultural life, curators and restorers, Patrimalp 2.0 intends to develop of a new interdisciplinary science: Cultural Heritage Science. The grant starts in January 2018, for a period of 48 months.

8.2.3. ANR: CaLiTrOp

Participant: Cyril Soler [contact].

Computing photorealistic images relies on the simulation of light transfer in a 3D scene, typically modeled using geometric primitives and a collection of reflectance properties that represent the way objects interact with light. Estimating the color of a pixel traditionally consists in integrating contributions from light paths connecting the light sources to the camera sensor at that pixel.

In this ANR we explore a transversal view of examining light transport operators from the point of view of infinite dimensional function spaces of light fields (imagine, e.g., reflectance as an operator that transforms a distribution of incident light into a distribution of reflected light). Not only are these operators all linear in these spaces but they are also very sparse. As a side effect, the sub-spaces of light distributions that are actually relevant during the computation of a solution always boil down to a low dimensional manifold embedded in the full space of light distributions.

Studying the structure of high dimensional objects from a low dimensional set of observables is a problem that becomes ubiquitous nowadays: Compressive sensing, Gaussian processes, harmonic analysis and differential analysis, are typical examples of mathematical tools which will be of great relevance to study the light transport operators.

Expected results of the fundamental-research project CALiTrOp, are a theoretical understanding of the dimensionality and structure of light transport operators, bringing new efficient lighting simulation methods, and efficient approximations of light transport with applications to real time global illumination for video games.

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Declared Inria International Partners

Title: “MAIS”: Mathematical Analysis of Image SynthesisInternational Partner (Institution - Laboratory - Researcher):

University of Montreal (Canada) - Département d’Informatique et Recherche Opérationnelle - Derek Nowrouzezahrai
Duration: 2015 - 2019

8.4. International Research Visitors

8.4.1. Visits to International Teams

8.4.1.1. Research Stays Abroad

Alban Fichet has returned in October 2018 from a 12 months research stay at Charles University in Prague, to work with Alexander Wilkie and Jaroslav Krivanek on material models.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

- Nicolas Holzschuch is a member of the steering committee for the Eurographics Working Group on Rendering.

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Georges-Pierre Bonneau: *Solid and Physical Modeling* 2018, *Shape Modeling International* 2018, *EnvirVis* 2018, Short paper track for *Eurographics Visualization Symposium* 2018,
- Joëlle Thollot: *Expressive* 2018.

9.1.3. Journal

9.1.3.1. Reviewer - Reviewing Activities

All members of the Maverick team work as reviewers for the most prestigious journals, including ACM TOG, IEEE TVCG, etc.

9.1.4. Research Administration

- Georges-Pierre Bonneau is member of the "conseil du Laboratoire Jean Kuntzmann".
- Romain Vergne is member of the "conseil du Laboratoire Jean Kuntzmann".
- Romain Vergne is co-responsible of the GT Rendu.
- Romain Vergne is co-responsible of the PhD students of the Laboratoire Jean Kuntzmann.
- Nicolas Holzschuch is an elected member of Inria Evaluation Committee (CE), an elected member of Inria Comité Technique (CTI) and a reserve member of Inria Scientific Council (CS).
- Nicolas Holzschuch is responsible for the department "Geometry and Images" of the Laboratoire Jean Kuntzmann.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Joëlle Thollot and Georges-Pierre Bonneau are both full Professor of Computer Science. Romain Vergne is an associate professor in Computer Science. They teach general computer science topics at basic and intermediate levels, and advanced courses in computer graphics and visualization at the master levels. Nicolas Holzschuch teaches advanced courses in computer graphics at the Master level.

Licence: Joëlle Thollot, Théorie des langages, 27h, L3, ENSIMAG, France

Licence: Joëlle Thollot, Séminaire d'innovation, 10h, L3, ENSE3, France

Master: Joëlle Thollot, Responsable de la flière MMIS (Modélisation mathématique, image, simulation) 24h, M1-M2, ENSIMAG, France

Master: Joëlle Thollot, Tutorat d'apprentis, 10h, M1-M2, ENSIMAG, France

Licence : Romain Vergne, Introduction to algorithms, 64h, L1, UGA, France.

Licence : Romain Vergne, WebGL, 29h, L3, IUT2 Grenoble, France.

Master : Romain Vergne, Geometric modeling, 18h, M1, UGA, France.

Master : Romain Vergne, Image synthesis, 27h, M1, UGA, France.

Master : Romain Vergne, Image synthesis, 15h, M1, Polytech, France.

Master : Romain Vergne, 3D graphics, 15h, M1, UGA, France.

Master : Nicolas Holzschuch, Computer Graphics II, 18h, M2 MoSIG, France.

Master : Nicolas Holzschuch, Synthèse d'Images et Animation, 32h, M2, ENSIMAG, France.

Licence: Georges-Pierre Bonneau, Algorithmique et Programmation Impérative, 23h, L3, Polytech-Grenoble, France.

Master: Georges-Pierre Bonneau, responsable de la 4ième année du département INFO, 32h, M1, Polytech-Grenoble, France

Master: Georges-Pierre Bonneau, Image Synthesis, 23h, M1, Polytech-Grenoble, France

Master: Georges-Pierre Bonneau, Data Visualization, 40h, M2, Polytech-Grenoble, France

Master: Georges-Pierre Bonneau, Digital Geometry, 23h, M1, UGA

Master: Georges-Pierre Bonneau, Information Visualization, 22h, Mastere, ENSIMAG, France.

Master: Georges-Pierre Bonneau, Scientific Visualization, M2, ENSIMAG, France.

Master: Georges-Pierre Bonneau, Computer Graphics II, 18h, M2 MoSiG, UGA, France.

9.2.2. Supervision

PhD: Guillaume Loubet, Efficient representations for sub-pixel appearence, 25/06/2018, Fabrice Neyret

PhD: Jérémy Wambecke, Visualisation de Données Temporelles Personnelles, 22/10/2018, Georges-Pierre Bonneau, Romain Vergne, Renaud Blanch

PhD: Alexandre Bléron, Stylization of animated 3D scenes in a painterly style, 8/11/2018, Joëlle Thollot, Romain Vergne

PhD in progress: Alban Fichet, Efficient representation for measured reflectance, 1/10/2015 Nicolas Holzschuch.

PhD in progress: Vincent Tavernier, Procedural stochastic textures, 1/10/2017, Fabrice Neyret, Joëlle Thollot, Romain Vergne.

PhD in progress: Sunrise Wang, Light transport operators simplification using neural networks, 1/9/2018, Nicolas Holzschuch

PhD in progress: Morgane Gérardin, Connecting physical and chemical properties with material appearance, 1/10/2018, Nicolas Holzschuch

PhD in progress: Ronak Molazem, Dimensional Analysis of Light Transport, 1/09/2018, Cyril Soler

9.2.3. Juries

Nicolas Holzschuch, member of the jury and referee, PhD of Laurent Gilles, Telecom Paristech, 12/11/2018

Nicolas Holzschuch, member of the jury and referee, PhD of Hélène Perrier, University of Lyon, 7/3/2018

Joëlle Thollot, member of the jury and referee, PhD of Jean-Dominique Favreau (Inria Sophia-Antipolis) 15/03/2018,

Joëlle Thollot, member of the jury, HdR of Aurélie BUGEAU, University of Bordeaux, 30/05/2018

Georges-Pierre Bonneau, president of the jury, PhD of Ali Jabbari, University of Grenoble, 4/07/2018

Georges-Pierre Bonneau, member of the jury, HdR of Basile Sauvage, University of Strasbourg, 16/07/2018

Georges-Pierre Bonneau, president of the jury, PhD of Even Enthem, University of Grenoble, 26/10/2018

9.3. Popularization

Every year, “MobiNet” (see section 6.7) classes are conducted with high school pupils of the large Grenoble area to practice initiation and intuition on Computer Science, Maths and Physics. Depending on the year, we have 2 to 4 groups in the scope of INP-Grenoble “Enginneering weeks”, and 0 to 2 groups in the scope of Math-C2+ operations.

Fabrice Neyret published a popularization article "Le mouvement sous toutes ses coutures" [19] in the booklet "Maths Mouvement Express 2018".

Fabrice Neyret maintains the blog shadertoyunofficial (https://shadertoyunofficial.wordpress.com/) and various shaders examples on Shadertoy site (https://www.shadertoy.com/) to popularize GPU technologies as well as disseminates academic models within computer graphics, computer science, applied math and physics fields.

Cyril Soler presented at FOSDEM2018 his work on privacy preserving data distribution over decentralized mesh networks ([https://www.youtube.com/watch?v=FR9SSdGN0K8]).

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. BLÉRON.*Real-time stylized rendering of 3D animated scenes*, Université Grenoble Alpes, November 2018, https://tel.archives-ouvertes.fr/tel-02061390

[2] G. LOUBET.*Efficient models for representing sub-pixel appearances*, Université Grenoble Alpes, June 2018, https://tel.archives-ouvertes.fr/tel-01849666

[3] J. WAMBECKE.*Visualization of personal time-dependent data*, Université Grenoble Alpes, October 2018, https://tel.archives-ouvertes.fr/tel-02007675

Articles in International Peer-Reviewed Journal

[4] E. HEITZ, F. NEYRET.*High-Performance By-Example Noise using a Histogram-Preserving Blending Operator*, in "Proceedings of the ACM on Computer Graphics and Interactive Techniques", August 2018, vol. 1, n° 2, Article No. 31:1-25 [*DOI :* 10.1145/3233304], https://hal.inria.fr/hal-01824773

[5] G. LOUBET, F. NEYRET.*A new microflake model with microscopic self-shadowing for accurate volume downsampling*, in "Computer Graphics Forum", May 2018, vol. 37, n° 2, p. 111-121 [*DOI :* 10.1111/CGF.13346], https://hal.archives-ouvertes.fr/hal-01702000

[6] M. MOJZÍK, A. FICHET, A. WILKIE.*Handling Fluorescence in a Uni-directional Spectral Path Tracer*, in "Computer Graphics Forum", July 2018, vol. 37, n° 4, p. 77 - 94 [*DOI :* 10.1111/CGF.13477], https://hal.inria.fr/hal-01818826

[7] C. SOLER, K. SUBR, D. NOWROUZEZAHRAI.*A Versatile Parameterization for Measured Material Manifolds*, in "Computer Graphics Forum", April 2018, vol. 37, n° 2, p. 135-144 [*DOI :* 10.1111/CGF.13348], https://hal.inria.fr/hal-01702722

[8] B. WANG, L. GE, N. HOLZSCHUCH.*Precomputed Multiple Scattering for Rapid Light Simulation in Participating Media*, in "Transactions on Visualization & Computer Graphics (TVCG)", December 2018, https://hal.inria.fr/hal-01970399

[9] B. WANG, N. HOLZSCHUCH.*Point-Based Rendering for Homogeneous Participating Media with Refractive Boundaries*, in "IEEE Transactions on Visualization and Computer Graphics", October 2018, vol. 24, n^o^ 10, p. 2743-2757 [*DOI :* 10.1109/TVCG.2017.2768525], https://hal.inria.fr/hal-01628188

[10] B. WANG, L. WANG, N. HOLZSCHUCH.*Fast Global Illumination with Discrete Stochastic Microfacets Using a Filterable Model*, in "Computer Graphics Forum", October 2018, vol. 37, n^o^ 7, p. 1-10, https://hal.inria.fr/hal-01856980

International Conferences with Proceedings

[11] A. BLÉRON, R. VERGNE, T. HURTUT, J. THOLLOT.*Motion-coherent stylization with screen-space image filters*, in "Expressive '18 - The Joint Symposium on Computational Aesthetics and Sketch Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering", Victoria, Canada, ACM, August 2018, p. 1-13 [*DOI :* 10.1145/3229147.3229163], https://hal.inria.fr/hal-01826558

[12] L. GE, B. WANG, L. WANG, N. HOLZSCHUCH.*A Compact Representation for Multiple Scattering in Participating Media using Neural Networks*, in "Siggraph 2018 Talks", Vancouver, Canada, August 2018, p. 1-2 [*DOI :* 10.1145/3214745.3214758], https://hal.inria.fr/hal-01809890

[13] S. E. MONTESDEOCA, H. S. SEAH, A. SEMMO, P. BÉNARD, R. VERGNE, J. THOLLOT, D. BENVENUTI.*MNPR: A Framework for Real-Time Expressive Non-Photorealistic Rendering of 3D Computer Graphics*, in "Expressive '18 - The Joint Symposium on Computational Aesthetics and Sketch Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering", Victoria, Canada, August 2018, 11 [*DOI :* 10.1145/3229147.3229162], https://hal.inria.fr/hal-01814332

[14] T. STANKO, N. SAGUIN-SPRYNSKI, L. JOUANET, S. HAHMANN, G.-P. BONNEAU.*Morphorider: a new way for Structural Monitoring via the shape acquisition with a mobile device equipped with an inertial node of sensors*, in "9th European Workshop on Structural Health Monitoring, 2018", Manchester, United Kingdom, July 2018, https://hal.inria.fr/hal-01962789

[15] J. WAMBECKE, G.-P. BONNEAU, R. BLANCH, R. VERGNE.*In-situ Evaluation of a "What if" Approach for Eco-feedback*, in "IHM 2018 - 30eme conférence francophone sur l'interaction homme-machine", Brest, France, AFIHM (editor), Articles Scientifiques, October 2018, p. 83-91, https://hal.archives-ouvertes.fr/hal-01899031

[16] J. WAMBECKE, G.-P. BONNEAU, R. BLANCH, R. VERGNE.*What if we use the "What if" Approach for Eco-Feedback? Designing an Electricity Consumption Analysis for Layman Users*, in "EnvirVis 2018 - Workshop on Visualisation in Environmental Sciences", Brno, Czech Republic, June 2018, p. 1-8 [*DOI :* 10.2312/ENVIRVIS.20181140], https://hal.inria.fr/hal-01813283

Conferences without Proceedings

[17] V. TAVERNIER, F. NEYRET, R. VERGNE, J. THOLLOT.*Gabor Noise Revisited*, in "j•FIG 2018 - Journées Françaises d'Informatique Graphique", Poitiers, France, November 2018, p. 1-4, https://hal.archives-ouvertes.fr/hal-01926451

Research Reports

[18] A. BLÉRON, R. VERGNE, T. HURTUT, J. THOLLOT.*A workflow for designing stylized shading effects*, Inria Grenoble Rhône-Alpes, November 2018, n^o RR-9225, p. 1-29, https://hal.inria.fr/hal-01919501

Scientific Popularization

[19] F. NEYRET.*Le mouvement sous toutes ses coutures*, May 2018, p. 1-6, Maths Mouvement Express: Brochure de vulgarisation mathématique éditée annuellement par le CIJM à l'occasion du Salon Culture & Jeux Mathématiqueshttps://www.cijm.org/accueil/productions-cijm/90-maths-express, https://hal.inria.fr/hal-01915380

Other Publications

[20] A. BIZARD, F. NEYRET.*Procedural modeling of multiscale galactic dust*, LJK / Grenoble University - Inria, June 2018, p. 1-46, https://hal.inria.fr/hal-01870165

Project-Team MISTIS

Modelling and Inference of Complex and Structured Stochastic Systems

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

IN PARTNERSHIP WITH:
Institut polytechnique de Grenoble

Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Optimization, machine learning and statistical methods

Table of contents

Project-Team MISTIS

Creation of the Project-Team: 2008 January 01

Keywords:

Computer Science and Digital Science:

A3.1.1. - Modeling, representation
A3.1.4. - Uncertain data
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.4. - Optimization and learning
A3.4.5. - Bayesian methods
A3.4.7. - Kernel methods
A5.3.3. - Pattern recognition
A5.9.2. - Estimation, modeling
A6.2. - Scientific computing, Numerical Analysis & Optimization
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.3. - Computation-data interaction
A6.3.1. - Inverse problems
A6.3.3. - Data processing
A6.3.5. - Uncertainty Quantification
A9.2. - Machine learning
A9.3. - Signal analysis

Other Research Topics and Application Domains:

B1.2.1. - Understanding and simulation of the brain and the nervous system
B2.6.1. - Brain imaging
B3.3. - Geosciences
B3.4.1. - Natural risks
B3.4.2. - Industrial risks and waste
B3.5. - Agronomy
B5.1. - Factory of the future
B9.5.6. - Data science
B9.11.1. - Environmental risks

1. Team, Visitors, External Collaborators

Research Scientists

Julyan Arbel [Inria, Researcher]
Florence Forbes [Team leader, Inria, Senior Researcher, HDR]
Stéphane Girard [Inria, Senior Researcher, HDR]

Faculty Member
Jean-Baptiste Durand [Institut polytechnique de Grenoble, Associate Professor]

External Collaborator
Eric Paul Marchand [Univ. Sherbrook, Canada, from Mar 2018 until Jun 2018]

Technical Staff
Jaime Eduardo Arias Almeida [Inria, until Jan 2018]
Fatima Fofana [Inria]
Steven Quinito Masnada [Inria, until Aug 2018]

PhD Students
Clément Albert [Inria]
Alexis Arnaud [Institut polytechnique de Grenoble]
Karina Ashurbekova [Univ Grenoble Alpes]
Meryem Bousebata [Univ Grenoble Alpes, from Oct 2018]
Fabien Boux [Univ Grenoble Alpes]
Alexandre Constantin [Univ Grenoble Alpes, from Oct 2018]
Benoit Kugler [Univ Grenoble Alpes, from Sep 2018]
Veronica Munoz Ramirez [Univ Grenoble Alpes]
Brice Olivier [Univ Grenoble Alpes]
Thibaud Rahier [Cifre Schneider Electric, until Nov 2018]
Pierre-Antoine Rodesch [CEA, until Oct 2018]
Mariia Vladimirova [Inria, from Oct 2018]

Post-Doctoral Fellows
Jean-Michel Bécu [Inria, until Mar 2018]
Marta Crispino [Inria, from Mar 2018]
Pascal Dkengne Sielenou [Inria, from Nov 2018]
Hongliang Lu [Inria]
Antoine Usseglio Carleve [Inria, from Oct 2018]
Fei Zheng [Inria, from Mar 2018]

Visiting Scientists
Aboubacrène Ag Ahmad [Univ. Gaston Berger, Senegal, from Sep 2018 until Nov 2018]
Hien Nguyen [Inria, from Sep 2018 until Oct 2018]

Administrative Assistant
Marion Ponsot [Inria]

2. Overall Objectives

2.1. Overall Objectives

The context of our work is the analysis of structured stochastic models with statistical tools. The idea underlying the concept of structure is that stochastic systems that exhibit great complexity can be accounted for by combining simple local assumptions in a coherent way. This provides a key to modelling, computation, inference and interpretation. This approach appears to be useful in a number of high impact applications including signal and image processing, neuroscience, genomics, sensors networks, etc. while the needs from these domains can in turn generate interesting theoretical developments. However, this powerful and flexible approach can still be restricted by necessary simplifying assumptions and several generic sources of complexity in data.

Often data exhibit complex dependence structures, having to do for example with repeated measurements on individual items, or natural grouping of individual observations due to the method of sampling, spatial or temporal association, family relationship, and so on. Other sources of complexity are related to the measurement process, such as having multiple measuring instruments or simulations generating high dimensional and heterogeneous data or such that data are dropped out or missing. Such complications in data-generating processes raise a number of challenges. Our goal is to contribute to statistical modelling by offering theoretical concepts and computational tools to handle properly some of these issues that are frequent in modern data. So doing, we aim at developing innovative techniques for high scientific, societal, economic impact applications and in particular via image processing and spatial data analysis in environment, biology and medicine.

The methods we focus on involve mixture models, Markov models, and more generally hidden structure models identified by stochastic algorithms on one hand, and semi and non-parametric methods on the other hand.

Hidden structure models are useful for taking into account heterogeneity in data. They concern many areas of statistics (finite mixture analysis, hidden Markov models, graphical models, random effect models, ...). Due to their missing data structure, they induce specific difficulties for both estimating the model parameters and assessing performance. The team focuses on research regarding both aspects. We design specific algorithms for estimating the parameters of missing structure models and we propose and study specific criteria for choosing the most relevant missing structure models in several contexts.

Semi and non-parametric methods are relevant and useful when no appropriate parametric model exists for the data under study either because of data complexity, or because information is missing. When observations are curves, they enable us to model the data without a discretization step. These techniques are also of great use for *dimension reduction* purposes. They enable dimension reduction of the functional or multivariate data with no assumptions on the observations distribution. Semi-parametric methods refer to methods that include both parametric and non-parametric aspects. Examples include the Sliced Inverse Regression (SIR) method which combines non-parametric regression techniques with parametric dimension reduction aspects. This is also the case in *extreme value analysis*, which is based on the modelling of distribution tails by both a functional part and a real parameter.

3. Research Program

3.1. Mixture models

Participants: Alexis Arnaud, Jean-Baptiste Durand, Florence Forbes, Stéphane Girard, Julyan Arbel, Jean-Michel Bécu, Hongliang Lu, Fabien Boux, Veronica Munoz Ramirez, Benoit Kugler, Alexandre Constantin, Fei Zheng.

Key-words: mixture of distributions, EM algorithm, missing data, conditional independence, statistical pattern recognition, clustering, unsupervised and partially supervised learning.

In a first approach, we consider statistical parametric models, θ being the parameter, possibly multi-dimensional, usually unknown and to be estimated. We consider cases where the data naturally divides into observed data $y = \{y_1, ..., y_n\}$ and unobserved or missing data $z = \{z_1, ..., z_n\}$. The missing data z_i represents for instance the memberships of one of a set of K alternative categories. The distribution of an observed y_i can be written as a finite mixture of distributions,

$$f(y_i; \theta) = \sum_{k=1}^{K} P(z_i = k; \theta) f(y_i \mid z_i; \theta) \ . \tag{1}$$

These models are interesting in that they may point out hidden variables responsible for most of the observed variability and so that the observed variables are *conditionally* independent. Their estimation is often difficult due to the missing data. The Expectation-Maximization (EM) algorithm is a general and now standard approach to maximization of the likelihood in missing data problems. It provides parameter estimation but also values for missing data.

Mixture models correspond to independent z_i's. They have been increasingly used in statistical pattern recognition. They enable a formal (model-based) approach to (unsupervised) clustering.

3.2. Markov models

Participants: Alexis Arnaud, Brice Olivier, Thibaud Rahier, Jean-Baptiste Durand, Florence Forbes, Karina Ashurbekova, Hongliang Lu, Pierre-Antoine Rodesch, Julyan Arbel, Mariia Vladimirova.

Key-words: graphical models, Markov properties, hidden Markov models, clustering, missing data, mixture of distributions, EM algorithm, image analysis, Bayesian inference.

Graphical modelling provides a diagrammatic representation of the dependency structure of a joint probability distribution, in the form of a network or graph depicting the local relations among variables. The graph can have directed or undirected links or edges between the nodes, which represent the individual variables. Associated with the graph are various Markov properties that specify how the graph encodes conditional independence assumptions.

It is the conditional independence assumptions that give graphical models their fundamental modular structure, enabling computation of globally interesting quantities from local specifications. In this way graphical models form an essential basis for our methodologies based on structures.

The graphs can be either directed, e.g. Bayesian Networks, or undirected, e.g. Markov Random Fields. The specificity of Markovian models is that the dependencies between the nodes are limited to the nearest neighbor nodes. The neighborhood definition can vary and be adapted to the problem of interest. When parts of the variables (nodes) are not observed or missing, we refer to these models as Hidden Markov Models (HMM). Hidden Markov chains or hidden Markov fields correspond to cases where the z_i's in (1) are distributed according to a Markov chain or a Markov field. They are a natural extension of mixture models. They are widely used in signal processing (speech recognition, genome sequence analysis) and in image processing (remote sensing, MRI, etc.). Such models are very flexible in practice and can naturally account for the phenomena to be studied.

Hidden Markov models are very useful in modelling spatial dependencies but these dependencies and the possible existence of hidden variables are also responsible for a typically large amount of computation. It follows that the statistical analysis may not be straightforward. Typical issues are related to the neighborhood structure to be chosen when not dictated by the context and the possible high dimensionality of the observations. This also requires a good understanding of the role of each parameter and methods to tune them depending on the goal in mind. Regarding estimation algorithms, they correspond to an energy minimization problem which is NP-hard and usually performed through approximation. We focus on a certain type of methods based on variational approximations and propose effective algorithms which show good performance in practice and for which we also study theoretical properties. We also propose some tools for model selection. Eventually we investigate ways to extend the standard Hidden Markov Field model to increase its modelling power.

3.3. Functional Inference, semi- and non-parametric methods

Participants: Clément Albert, Stéphane Girard, Florence Forbes, Jean-Michel Bécu, Antoine Usseglio Carleve, Pascal Dkengne Sielenou, Marta Crispino, Meryem Bousebata.

Key-words: dimension reduction, extreme value analysis, functional estimation.

We also consider methods which do not assume a parametric model. The approaches are non-parametric in the sense that they do not require the assumption of a prior model on the unknown quantities. This property is important since, for image applications for instance, it is very difficult to introduce sufficiently general parametric models because of the wide variety of image contents. Projection methods are then a way to decompose the unknown quantity on a set of functions (*e.g.* wavelets). Kernel methods which rely on smoothing the data using a set of kernels (usually probability distributions) are other examples. Relationships exist between these methods and learning techniques using Support Vector Machine (SVM) as this appears in the context of *level-sets estimation* (see section 3.3.2). Such non-parametric methods have become the cornerstone when dealing with functional data [77]. This is the case, for instance, when observations are curves. They enable us to model the data without a discretization step. More generally, these techniques are of great use for *dimension reduction* purposes (section 3.3.3). They enable reduction of the dimension of the functional or multivariate data without assumptions on the observations distribution. Semi-parametric methods refer to methods that include both parametric and non-parametric aspects. Examples include the Sliced Inverse Regression (SIR) method [79] which combines non-parametric regression techniques with parametric dimension reduction aspects. This is also the case in *extreme value analysis* [76], which is based on the modelling of distribution tails (see section 3.3.1). It differs from traditional statistics which focuses on the central part of distributions, *i.e.* on the most probable events. Extreme value theory shows that distribution tails can be modelled by both a functional part and a real parameter, the extreme value index.

3.3.1. Modelling extremal events

Extreme value theory is a branch of statistics dealing with the extreme deviations from the bulk of probability distributions. More specifically, it focuses on the limiting distributions for the minimum or the maximum of a large collection of random observations from the same arbitrary distribution. Let $X_{1,n} \leq ... \leq X_{n,n}$ denote n ordered observations from a random variable X representing some quantity of interest. A p_n-quantile of X is the value x_{p_n} such that the probability that X is greater than x_{p_n} is p_n, *i.e.*$P(X > x_{p_n}) = p_n$. When $p_n < 1/n$, such a quantile is said to be extreme since it is usually greater than the maximum observation $X_{n,n}$ (see Figure 1).

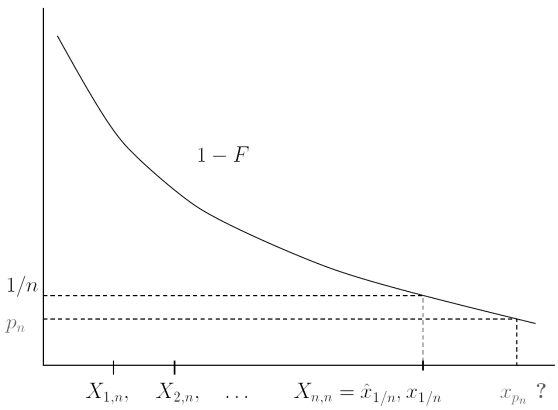


Figure 1. The curve represents the survival function $x \to P(X > x)$. The $1/n$-quantile is estimated by the maximum observation so that $\widehat{x}_{1/n} = X_{n,n}$. As illustrated in the figure, to estimate p_n-quantiles with $p_n < 1/n$, it is necessary to extrapolate beyond the maximum observation.

To estimate such quantiles therefore requires dedicated methods to extrapolate information beyond the observed values of X. Those methods are based on Extreme value theory. This kind of issue appeared in hydrology. One objective was to assess risk for highly unusual events, such as 100-year floods, starting from flows measured over 50 years. To this end, semi-parametric models of the tail are considered:

$$P(X > x) = x^{-1/\theta}\ell(x),\ x > x_0 > 0, \tag{2}$$

where both the extreme-value index $\theta > 0$ and the function $\ell(x)$ are unknown. The function ℓ is a slowly varying function *i.e.* such that

$$\frac{\ell(tx)}{\ell(x)} \to 1 \text{ as } x \to \infty \tag{3}$$

for all $t > 0$. The function $\ell(x)$ acts as a nuisance parameter which yields a bias in the classical extreme-value estimators developed so far. Such models are often referred to as heavy-tail models since the probability of extreme events decreases at a polynomial rate to zero. It may be necessary to refine the model (2,3) by specifying a precise rate of convergence in (3). To this end, a second order condition is introduced involving an additional parameter $\rho \leq 0$. The larger ρ is, the slower the convergence in (3) and the more difficult the estimation of extreme quantiles.
More generally, the problems that we address are part of the risk management theory. For instance, in reliability, the distributions of interest are included in a semi-parametric family whose tails are decreasing exponentially fast. These so-called Weibull-tail distributions [9] are defined by their survival distribution function:

$$P(X > x) = \exp\{-x^{\theta}\ell(x)\},\ x > x_0 > 0. \tag{4}$$

Gaussian, gamma, exponential and Weibull distributions, among others, are included in this family. An important part of our work consists in establishing links between models (2) and (4) in order to propose new estimation methods. We also consider the case where the observations were recorded with a covariate information. In this case, the extreme-value index and the p_n-quantile are functions of the covariate. We propose estimators of these functions by using moving window approaches, nearest neighbor methods, or kernel estimators.

3.3.2. Level sets estimation

Level sets estimation is a recurrent problem in statistics which is linked to outlier detection. In biology, one is interested in estimating reference curves, that is to say curves which bound 90% (for example) of the population. Points outside this bound are considered as outliers compared to the reference population. Level sets estimation can be looked at as a conditional quantile estimation problem which benefits from a non-parametric statistical framework. In particular, boundary estimation, arising in image segmentation as well as in supervised learning, is interpreted as an extreme level set estimation problem. Level sets estimation can also be formulated as a linear programming problem. In this context, estimates are sparse since they involve only a small fraction of the dataset, called the set of support vectors.

3.3.3. Dimension reduction

Our work on high dimensional data requires that we face the curse of dimensionality phenomenon. Indeed, the modelling of high dimensional data requires complex models and thus the estimation of high number of parameters compared to the sample size. In this framework, dimension reduction methods aim at replacing the original variables by a small number of linear combinations with as small as a possible loss of information. Principal Component Analysis (PCA) is the most widely used method to reduce dimension in data. However, standard linear PCA can be quite inefficient on image data where even simple image distortions can lead to highly non-linear data. Two directions are investigated. First, non-linear PCAs can be proposed, leading to semi-parametric dimension reduction methods [78]. Another field of investigation is to take into account the application goal in the dimension reduction step. One of our approaches is therefore to develop new Gaussian models of high dimensional data for parametric inference [75]. Such models can then be used in a Mixtures or Markov framework for classification purposes. Another approach consists in combining dimension reduction, regularization techniques, and regression techniques to improve the Sliced Inverse Regression method [79].

4. Application Domains

4.1. Image Analysis

Participants: Alexis Arnaud, Veronica Munoz Ramirez, Florence Forbes, Stéphane Girard, Jaime Eduardo Arias Almeida, Pierre-Antoine Rodesch, Hongliang Lu, Fabien Boux, Benoit Kugler, Alexandre Constantin.

As regards applications, several areas of image analysis can be covered using the tools developed in the team. More specifically, in collaboration with team PERCEPTION, we address various issues in computer vision involving Bayesian modelling and probabilistic clustering techniques. Other applications in medical imaging are natural. We work more specifically on MRI and functional MRI data, in collaboration with the Grenoble Institute of Neuroscience (GIN) and the NeuroSpin center of CEA Saclay. We also consider other statistical 2D fields coming from other domains such as remote sensing, in collaboration with the Institut de Planétologie et d'Astrophysique de Grenoble (IPAG) and the Centre National d'Etudes Spatiales (CNES). In this context, we worked on hyperspectral and/or multitemporal images. In the context of the "pole de competivité" project I-VP, we worked of images of PC Boards. We also address reconstruction problems in tomography with CEA Grenoble.

4.2. Multi sensor Data Analysis

Participants: Jean-Michel Bécu, Florence Forbes, Thibaud Rahier, Hongliang Lu, Fatima Fofana.

A number of our methods are at the the intersection of data fusion, statistics, machine learning and acoustic signal processing. The context can be the surveillance and monitoring of a zone acoustic state from data acquired at a continuous rate by a set of sensors that are potentially mobile and of different nature (eg WIFUZ project with the ACOEM company in the context of a DGA-rapid initiative). Typical objectives include the development of prototypes for surveillance and monitoring that are able to combine multi sensor data coming from acoustic sensors (microphones and antennas) and optical sensors (infrared cameras) and to distribute the processing to multiple algorithmic blocs. Our interest in acoustic data analysis mainly started from past European projects, POP and Humavips, in collaboration with the PERCEPTION team (PhD theses of Vassil Khalidov, Ramya Narasimha, Antoine Deleforge, Xavier Alameda, and Israel Gebru).

4.3. Biology, Environment and Medicine

Participants: Jaime Eduardo Arias Almeida, Alexis Arnaud, Florence Forbes, Stéphane Girard, Jean-Baptiste Durand, Clément Albert, Julyan Arbel, Jean-Michel Bécu, Thibaud Rahier, Brice Olivier, Karina Ashurbekova, Fabien Boux, Veronica Munoz Ramirez, Fei Zheng.

A third domain of applications concerns biology and medicine. We considered the use of mixture models to identify biomakers. We also investigated statistical tools for the analysis of fluorescence signals in molecular biology. Applications in neurosciences are also considered. In the environmental domain, we considered the modelling of high-impact weather events and the use of hyperspectral data as a new tool for quantitative ecology.

5. Highlights of the Year

5.1. Highlights of the Year

Scholarships:

- Alexandre Constantin supervised by S. Girard (MISTIS) and M. Fauvel (INRA Toulouse) was granted a PhD scholarship on "Analyse de séries temporelles massives d'images satellitaires: Applications à la cartographie des écosystèmes" from CNES and the IDEX Université Grenoble Alpes – Initiatives de Recherche Stratégiques (IRS).
- Meryem Bousebata supervised by S. Girard (MISTIS) and G. Enjolras (CERAG Grenoble) was granted a PhD scholarship on "Bayesian estimation of extreme risk measures: Implication for the insurance of natural disasters" from the Idex project named Risk@UGA.

Projects:

- In the context of another Idex project named Data@UGA, a 2-year multi-disciplinary project entitled "Tracking and analysis of large population of dynamic single molecules" was granted in November 2018 to MISTIS in collaboration with the GIN, coordinated by F. Forbes (MISTIS) and V. Stoppin-Mellet (GIN).

Editorial and publishing activities:

- A new book entitled *Handbook of mixture analysis*, edited at CRC Press by Gilles Celeux (Inria), Sylvia Früwirth-Schnatter (Wien University), and Christian P. Robert (Université Paris-Dauphine) is now available (December 2018). Florence Forbes and Julyan Arbel have written 2 of the chapters in the book [49], [51].
- Marianne Clausel and Jean-Baptiste Durand co-published a chapter [48] on generative models in data science in the book *Data Science. Cours et exercices*, edited by Eyrolles (Paris).
- Stéphane Girard and Julyan Arbel have co-edited a book of proceedings following the Summer School Stat4Astro they organized in Autrans in 2017 [64].

New appointments:

- Stéphane Girard has been hired as a research collaborator by the CMAP (Centre de Mathématiques Appliquées de l'école Polytechnique) in the context of the Chair Stress Test, RISK Management and Financial Steering, led by the French Ecole polytechnique and its Foundation and sponsored by BNP Paribas.

6. New Software and Platforms

6.1. BOLD model FIT

KEYWORDS: Functional imaging - FMRI - Health

SCIENTIFIC DESCRIPTION: Physiological and biophysical models have been proposed to link neuronal activity to the Blood Oxygen Level-Dependent (BOLD) signal in functional MRI (fMRI). Those models rely on a set of parameter values that are commonly estimated using gradient-based local search methods whose initial values are taken from the literature. In some applications, interesting insight into the brain physiology or physiopathology can be gained from an estimation of the model parameters from measured BOLD signals. In this work we focus on the extended Balloon model and propose the estimation of 15 parameters using seven different approaches: three versions of the Expectation Maximization Gauss-Newton (EM/GN) approach (the *de facto* standard in the neuroscientific community) and four metaheuristics (Particle Swarm Optimization (PSO), Differential Evolution (DE), Real-Coded Genetic Algorithms (GA), and a Memetic Algorithm (MA) combining EM/GN and DE). To combine both the ability to escape local optima and to incorporate prior knowledge, we derive the target function from Bayesian modeling. The general behavior of these algorithms is analyzed and compared, providing very promising results on challenging real and synthetic fMRI data sets involving rats with epileptic activity. These stochastic optimizers provided a better performance than EM/GN in terms of distance to the ground truth in 4 out of 6 synthetic data sets and a better signal fitting in 12 out of 12 real data sets. Non-parametric statistical tests showed the existence of statistically significant differences between the real data results obtained by DE and EM/GN. Finally, the estimates obtained from DE for these parameters seem both more realistic and more stable or at least as stable across sessions as the estimates from EM/GN. This is the largest comparison of optimizers for the estimation of biophysical parameters in BOLD fMRI

FUNCTIONAL DESCRIPTION: This Matlab toolbox performs the automatic estimation of biophysical parameters using the extended Balloon model and BOLD fMRI data. It takes as input a MAT file and provides as output the parameter estimates achieved by using stochastic optimization

NEWS OF THE YEAR: The main differences with our previous work: 1) we also use synthetic data, 2) we use stochastic GN and MCMC+DE, 3) We evaluate results not only in physiological terms but also comparing fitness function values. Also changes were made to allow running on the cluster via MPI

- Participants: Pablo Mesejo Santiago, Florence Forbes and Jan Warnking
- Partner: University of Granada, Spain
- Contact: Pablo Mesejo Santiago
- Publication: A differential evolution-based approach for fitting a nonlinear biophysical model to fMRI BOLD data
- URL: https://hal.archives-ouvertes.fr/hal-01221115v2/

6.2. PyHRF

KEYWORDS: Medical imaging - Health - Brain - IRM - Neurosciences - Statistic analysis - FMRI

SCIENTIFIC DESCRIPTION: Functional Magnetic Resonance Imaging (fMRI) is a neuroimaging technique that allows the non-invasive study of brain function. It is based on the hemodynamic variations induced by changes in cerebral synaptic activity following sensory or cognitive stimulation. The measured signal depends on the variation of blood oxygenation level (BOLD signal) which is related to brain activity: a decrease in deoxyhemoglobin concentration induces an increase in BOLD signal. The BOLD signal is delayed with respect to changes in synaptic activity, which can be modeled as a convolution with the Hemodynamic Response Function (HRF) whose exact form is unknown and fluctuates with various parameters such as age, brain region or physiological conditions. In this work we propose to analyze fMRI data using a Joint Detection-Estimation (JDE) approach. It jointly detects cortical activation and estimates the HRF. In contrast to existing tools, PyHRF estimates the HRF instead of considering it as a given constant in the entire brain.

FUNCTIONAL DESCRIPTION: As part of fMRI data analysis, PyHRF provides a set of tools for addressing the two main issues involved in intra-subject fMRI data analysis : (i) the localization of cerebral regions that elicit evoked activity and (ii) the estimation of the activation dynamics also referenced to as the recovery of the Hemodynamic Response Function (HRF). To tackle these two problems, PyHRF implements the Joint Detection-Estimation framework (JDE) which recovers parcel-level HRFs and embeds an adaptive spatio-temporal regularization scheme of activation maps.

NEWS OF THE YEAR: The framework to perform software tests has been further developed. Some unitary tests have been set.

- Participants: Aina Frau Pascual, Christine Bakhous, Florence Forbes, Jaime Eduardo Arias Almeida, Laurent Risser, Lotfi Chaari, Philippe Ciuciu, Solveig Badillo, Thomas Perret and Thomas Vincent
- Partners: CEA - NeuroSpin
- Contact: Florence Forbes
- Publications: Frontiers in Neuroinformatics Flexible multivariate hemodynamics fMRI data analyses and simulations with PyHRF - Fast joint detection-estimation of evoked brain activity in event-related fMRI using a variational approach - A Bayesian Non-Parametric Hidden Markov Random Model for Hemodynamic Brain Parcellation
- URL: http://pyhrf.org

6.3. xLLiM

High dimensional locally linear mapping
KEYWORDS: Clustering - Regression
SCIENTIFIC DESCRIPTION: Building a regression model for the purpose of prediction is widely used in all disciplines. A large number of applications consists of learning the association between responses and predictors and focusing on predicting responses for the newly observed samples. In this work, we go beyond simple linear models and focus on predicting low-dimensional responses using high-dimensional covariates when the associations between responses and covariates are non-linear.
FUNCTIONAL DESCRIPTION: This is an R package available on the CRAN at https://cran.r-project.org/web/packages/xLLiM/index.html

XLLiM provides a tool for non linear mapping (non linear regression) using a mixture of regression model and an inverse regression strategy. The methods include the GLLiM model (Deleforge et al (2015)) based on Gaussian mixtures and a robust version of GLLiM, named SLLiM (see Perthame et al (2016)) based on a mixture of Generalized Student distributions.
NEWS OF THE YEAR: A new Hierarchical version of GLLiM has been developed in collaboration with University of Michigan, USA.

- Participants: Antoine Deleforge, Emeline Perthame and Florence Forbes
- Partner: University of Michigan, Ann Arbor, USA
- Contact: Florence Forbes
- Publications: Inverse regression approach to robust nonlinear high-to-low dimensional mapping - High-Dimensional Regression with Gaussian Mixtures and Partially-Latent Response Variables
- URL: https://cran.r-project.org/web/packages/xLLiM/index.html

6.4. MMST

Mixtures of Multiple Scaled Student T distributions
KEYWORDS: Health - Statistics - Brain MRI - Medical imaging - Robust clustering
SCIENTIFIC DESCRIPTION: A new family of multivariate heavy-tailed distributions that allow variable marginal amounts of tailweight is proposed and implemented. The originality comes from introducing multidimensional instead of univariate scale variables for the mixture of scaled Gaussian family of distributions. In contrast to most existing approaches, the derived distributions can account for a variety of shapes and have a simple tractable form with a closed-form probability density function whatever the dimension. We provide maximum likelihood estimation of the parameters and illustrate their modelling flexibility.
FUNCTIONAL DESCRIPTION: The package implements mixtures of so-called multiple scaled Student distributions, which are generalisation of multivariate Student T distribution allowing different tails in each dimension. Typical applications include Robust clustering to analyse data with possible outliers. In this context, the model and package have been used on large data sets of brain MRI to segment and identify brain tumors. Recent additions include a Markov random field implementation to account for spatial dependencies between observations, and a Bayesian implementation that can be used to select the number of mixture components automatically.

RELEASE FUNCTIONAL DESCRIPTION: Recent additions include a Markov random field implementation to account for spatial dependencies between observations, and a Bayesian implementation that can be used to select the number of mixture components automatically.

NEWS OF THE YEAR: Recent additions include a Markov random field implementation to account for spatial dependencies between observations, and a Bayesian implementation that can be used to select the number of mixture components automatically.

- Participants: Alexis Arnaud, Darren Wraith, Florence Forbes, Steven Quinito Masnada and Stéphane Despréaux
- Partner: Institut des Neurosciences Grenoble
- Contact: Florence Forbes
- Publications: A new family of multivariate heavy-tailed distributions with variable marginal amounts of tailweights: Application to robust clustering - Fully Automatic Lesion Localization and Characterization: Application to Brain Tumors Using Multiparametric Quantitative MRI Data
- URL: https://team.inria.fr/mistis/software/

7. New Results

7.1. Mixture models

7.1.1. Hierarchical mixture of linear mappings in high dimension

Participant: Florence Forbes.

Joint work with: Benjamin Lemasson from Grenoble Institute of Neuroscience, Naisyin Wang and Chun-Chen Tu from University of Michigan, Ann Arbor, USA.

Regression is a widely used statistical tool. A large number of applications consists of learning the association between responses and predictors. From such an association, different tasks, including prediction, can then be conducted. To go beyond simple linear models while maintaining tractability, non-linear mappings can be handled through exploration of local linearity. The non-linear relationship can be captured by a mixture of locally linear regression models as proposed in the so-called Gaussian Locally Linear Mapping (GLLiM) model [6] that assumes Gaussian noise models. In the past year, we have been working on several extensions and applications of GLLiM as described below and the next two subsections.

We proposed a structured mixture model called Hierarchical Locally Linear Mapping (HGLLiM), to predict low-dimensional responses based on high dimensional covariates when the associations between the responses and the covariates are non-linear. For tractability, HGLLiM adopts inverse regression to handle the high dimension and locally-linear mappings to capture potentially non-linear relations. Data with similar associations are grouped together to form a cluster. A mixture is composed of several clusters following a hierarchical structure. This structure enables shared covariance matrices and latent factors across smaller clusters to limit the number of parameters to estimate. Moreover, HGLLiM adopts a robust estimation procedure for model stability. We used three real-world datasets to demonstrate different features of HGLLiM. With the face dataset, HGLLiM shows the ability of modeling non-linear relationship through mixtures. With the orange juice dataset, we show the prediction performance of HGLLiM is robust to the presence of outliers. Moreover, we demonstrated that HGLLiM is capable of handling large-scale complex data using the data acquired from a magnetic resonance vascular fingerprinting (MRvF) study. These examples illustrate the wide applicability of HGLLiM on handling different aspects of a complex data structure in prediction. A preliminary version of this work under revision for JRSS-C can be found in [72].

7.1.2. Dictionary-free MR fingerprinting parameter estimation via inverse regression

Participants: Florence Forbes, Fabien Boux, Julyan Arbel.

Joint work with: Emmanuel Barbier from Grenoble Institute of Neuroscience.

Magnetic resonance imaging (MRI) can map a wide range of tissue properties but is often limited to observe a single parameter at a time. In order to overcome this problem, Ma et al. introduced magnetic resonance fingerprinting (MRF), a procedure based on a dictionary of simulated couples of signals and parameters. Acquired signals called fingerprints are then matched to the closest signal in the dictionary in order to estimate parameters. This requires an exhaustive search in the dictionary, which even for moderately sized problems, becomes costly and possibly intractable . We propose an alternative approach to estimate more parameters at a time. Instead of an exhaustive search for every signal, we use the dictionary to learn the functional relationship between signals and parameters. This allows the direct estimation of parameters without the need of searching through the dictionary. We investigated the use of GLLiM [6] that bypasses the problems associated with high-to-low regression. The experimental validation of our method is performed in the context of vascular fingerprinting. The comparison between a standard grid search and the proposed approach suggest that MR Fingerprinting could benefit from a regression approach to limit dictionary size and fasten computation time. Preliminary tests and results have been presented at International Society for Magnetic Resonance in Medicine conference, ISMRM 2018 [35].

7.1.3. *Massive analysis of multi-angular hyperspectral images of the planet Mars by inverse regression of physical models*

Participants: Florence Forbes, Benoit Kugler.

Joint work with: Sylvain Douté from Institut de Planétologie et d'Astrophysique de Grenoble (IPAG).

In the starting PhD of Benoit Kugler, the objective is to develop a statistical learning technique capable of solving a complex inverse problem in planetary remote sensing. The challenges are 1) the large number of observations to to inverse, 2) their large dimension, 3) the need to provide predictions for correlated parameters and 4) the need to provide a quality index (eg. uncertainty). To achieve this goal, we have started to investigate a setting in which a physical model is available to provide simulations that can then be used for learning prior to inversion of real observed data. For the learning step to be as accurate as possible, an initial task is then to estimate the best fit of the theoretical model to the real data. We proposed an iterative procedure based on a combination of GLLiM [6] predictions and importance sampling steps.

7.1.4. *Quantitative MRI Characterization of Brain Abnormalities in de novo Parkinsonian patients*

Participants: Florence Forbes, Veronica Munoz Ramirez, Alexis Arnaud, Julyan Arbel.

Joint work with: Michel Dojat from Grenoble Institute of Neuroscience.

Currently there is an important delay between the onset of Parkinson's disease and its diagnosis. The detection of changes in physical properties of brain structures may help to detect the disease earlier. In this work, we proposed to take advantage of the informative features provided by quantitative MRI to construct statistical models representing healthy brain tissues. We used mixture models of non Gaussian distributions [8] to capture the non-standard shape of the data multivariate distribution. This allowed us to detect atypical values for these features in the brain of Parkinsonian patients following a procedure similar to that in [16]. Promising preliminary results demonstrate the potential of our approach in discriminating patients from controls and revealing the subcortical structures the most impacted by the disease. This work has been accepted at the IEEE International Symposium on Biological Imaging, ISBI 2019 [36].

7.1.5. *No structural differences are revealed by voxel-based morphometry in de novo Parkinsonian patients*

Participants: Florence Forbes, Veronica Munoz Ramirez.

Joint work with: Michel Dojat from Grenoble Institute of Neuroscience and Pierrick Coupé from Laboratoire Bordelais de Recherche en Informatique, UMR 5800, Univ. Bordeaux, Talence.

The identification of brain morphological alterations in newly diagnosed PD patients (i.e. de novo) could potentially serve as a biomarker and accelerate diagnosis. However, presently no consensus exists in the literature possibly due to several factors: small size cohorts, differences in segmentation techniques or bad control of false positive rates. In this study, we seek, using the Computational Anatomy Toolbox (CAT12) (University of Jena) pipeline, for morphological brain differences in gray and white matter of 66 controls and 144 de novo PD patients whose data were extracted from the PPMI (Parkinson Progressive Markers Initiative) database. Moreover, we searched for subcortical structure differences using the new online platform VolBrain (J. V. Manjón and P. Coupé, "volBrain: An Online MRI Brain Volumetry System," Front. Neuroinform., vol. 10, p. 30, Jul. 2016). We found no structural brain differences in this de novo Parkinsonian population, neither in tissues using a whole brain analysis nor in any of nine subcortical structures analyzed separately. We concluded that some results published in the literature appear as false positives and are not reproducible.

7.1.6. Characterization of daily glycemic variability in the patient with type 1 diabetes

Participants: Florence Forbes, Fei Zheng.

Joint work with: Stéphane Bonnet from CEA Leti and Pierre-Yves Benhamou, Manon Jalbert from CHU Grenoble Alpes.

Glycemic variability (GV) is an important component of glycemic control in patients with type 1 diabetes. Many metrics have been proposed to account for this variability but none is unanimous among physicians. One difficulty is that the variations in blood sugar levels are expressed very differently from one day to another in some subjects. Our goal was to develop and evaluate the performance of a daily GV index built by combining different known metrics (CV, MAGE, GVP etc). This in order to merge their descriptive power to obtain a more complete and more accurate index. This preliminary study will be presented at the Société Francophone du Diabète (SFD) in 2019 [46].

7.1.7. Glycemic variability improves after pancreatic islet transplantation in patients with type 1 diabetes

Participants: Florence Forbes, Fei Zheng.

Joint work with: Stéphane Bonnet from CEA Leti and Pierre-Yves Benhamou, Manon Jalbert from CHU Grenoble Alpes.

Glycemic variability (GV) must be taken into account in the efficacy of treatment of type 1 diabetes because it determines the quality of glycemic control, the risk of complication of the patient's disease. Our goal in this study was to describe GV scores in patients with pancreatic islet transplantation (PIT) type 1 diabetes in the TRIMECO trial, and change of thresholds, for each index. predictive of success of PIT.

7.1.8. Dirichlet process mixtures under affine transformations of the data

Participant: Julyan Arbel.

Joint work with: Riccardo Corradin from Milano Bicocca, Italy and Bernardo Nipoti from Trinity College Dublin, Ireland.

Location-scale Dirichlet process mixtures of Gaussians (DPM-G) have proved extremely useful in dealing with density estimation and clustering problems in a wide range of domains. Motivated by an astronomical application, in this work we address the robustness of DPM-G models to affine transformations of the data, a natural requirement for any sensible statistical method for density estimation. In [57], we first devise a coherent prior specification of the model which makes posterior inference invariant with respect to affine transformation of the data. Second, we formalize the notion of asymptotic robustness under data transformation and show that mild assumptions on the true data generating process are sufficient to ensure that DPM-G models feature such a property. As a by-product, we derive weaker assumptions than those provided in the literature for ensuring posterior consistency of Dirichlet process mixtures, which could reveal of independent interest. Our investigation is supported by an extensive simulation study and illustrated by the analysis of an astronomical dataset consisting of physical measurements of stars in the field of the globular cluster NGC 2419.

7.1.9. Applications of mixture models in Industry

Participant: Julyan Arbel.

Joint work with: Kerrie Mengersen, Earl Duncan, Clair Alston-Knox and Nicole White.

A very wide range of commonly encountered problems in industry are amenable to statistical mixture modelling and analysis. These include process monitoring or quality control, efficient resource allocation, risk assessment, prediction, and so on. Commonly articulated reasons for adopting a mixture approach include the ability to describe non-standard outcomes and processes, the potential to characterize each of a set of multiple outcomes or processes via the mixture components, the concomitant improvement in interpretability of the results, and the opportunity to make probabilistic inferences such as component membership and overall prediction.

In [51], We illustrate the wide diversity of applications of mixture models to problems in industry, and the potential advantages of these approaches, through a series of case studies.

7.1.10. Approximation results regarding the multiple-output mixture of the Gaussian-gated linear experts model

Participant: Florence Forbes.

Joint work with: Hien Nguyen, La Trobe University Melbourne Australia and Faicel Chamroukhi, Caen University, France.

Mixture of experts (MoE) models are a class of artificial neural networks that can be used for functional approximation and probabilistic modeling. An important class of MoE models is the class of mixture of linear experts (MoLE) models, where the expert functions map to real topological output spaces. Recently, Gaussian-gated MoLE models have become popular in applied research. There are a number of powerful approximation results regarding Gaussian-gated MoLE models, when the output space is univariate. These results guarantee the ability of Gaussian-gated MoLE mean functions to approximate arbitrary continuous functions, and Gaussian-gated MoLE models themselves to approximate arbitrary conditional probability density functions. We utilized and extended upon the univariate approximation results in order to prove a pair of useful results for situations where the output spaces are multivariate. We do this by proving a pair of lemmas regarding the combination of univariate MoLE models, which are interesting in their own rights.

7.1.11. Models for ranking data

Participant: Marta Crispino.

within the BigInsight project, Oslo.

We developed a new method and algorithms for working with ranking data. This kind of data is particularly relevant in applications involving personalized recommendations. In particular, we have invented a new Bayesian approach based on extensions of the Mallows model, which allows making personalized recommendations equipped with a level of uncertainty.

The Mallows model (MM) is a popular parametric family of models for ranking data, based on the assumption that a modal ranking, which can be interpreted as the consensus ranking of the population, exists. The probability of observing a given ranking is then assumed to decay exponentially fast as its distance from the consensus grows. The MM is therefore a two-parameter distance-based family of models. The scale or precision parameter, controlling the concentration of the distribution determines the rate of decay of the probability of individual ranks. Individual models with different properties can be obtained depending on the choice of distance on the space of permutations. A major drawback of the MM is that its computational complexity has limited its use to a particular form based on Kendall distance. We develop new computationally tractable methods for Bayesian inference in Mallows models that work with any right-invariant distance. Our method performs inference on the consensus ranking of the items, also when based on partial rankings, such as top-k items or pairwise comparisons. When assessors are many or heterogeneous, we propose a mixture model for clustering them in homogeneous subgroups, with cluster specific consensus rankings. We develop approximate stochastic algorithms that allow a fully probabilistic analysis, leading to coherent quantifications of uncertainties, make probabilistic predictions on the class membership of assessors based on their ranking of just some items, and predict missing individual preferences, as needed in recommendation systems. The methodology has been published in the Journal of Machine Learning Research, JMLR, in early 2018.

A generalization of the model above involves dealing with non-transitive and heterogeneous pairwise comparison data, coming from an experiment within the musicology domain. We thus develop a mixture model extension of the Bayesian Mallows model able to handle non-transitive data, with a latent layer of uncertainty which captures the generation of preference misreporting. This paper was recently accepted for publication in the Annals of Applied Statistics, AoAS.

Within this project, we also write a survey paper, whose main goal is to compare the performance of our method with other existing methodologies, including the Plackett-Luce, the Bradley-Terry, the collaborative filtering methods, and some of their variations. We illustrate and discuss the use of these models by means of an experiment in which assessors rank potatoes, and with a simulation. The purpose of this paper is not to recommend the use of one best method, but to present a palette of different possibilities for different questions and different types of data. This was recently accepted on the Annual Review of Statistics and Its Applications, ARSIA.

7.2. Semi and non-parametric methods

7.2.1. Estimation of extreme risk measures

Participant: Stéphane Girard.

Joint work with: A. Daouia (Univ. Toulouse), L. Gardes (Univ. Strasbourg), J. Elmethni (Univ. Paris 5) and G. Stupfler (Univ. Nottingham, UK).

One of the most popular risk measures is the Value-at-Risk (VaR) introduced in the 1990's. In statistical terms, the VaR at level $\alpha \in (0,1)$ corresponds to the upper α-quantile of the loss distribution. The Value-at-Risk however suffers from several weaknesses. First, it provides us only with a pointwise information: VaR(α) does not take into consideration what the loss will be beyond this quantile. Second, random loss variables with light-tailed distributions or heavy-tailed distributions may have the same Value-at-Risk. Finally, Value-at-Risk is not a coherent risk measure since it is not subadditive in general. A first coherent alternative risk measure is the Conditional Tail Expectation (CTE), also known as Tail-Value-at-Risk, Tail Conditional Expectation or Expected Shortfall in case of a continuous loss distribution. The CTE is defined as the expected loss given that the loss lies above the upper α-quantile of the loss distribution. This risk measure thus takes into account the whole information contained in the upper tail of the distribution. In [20], we investigate the extreme properties of a new risk measure (called the Conditional Tail Moment) which encompasses various risk measures, such as the CTE, as particular cases. We study the situation where some covariate information is available under some general conditions on the distribution tail. We thus has to deal with conditional extremes. However, the asymptotic normality of the empirical CTE estimator requires that the underlying distribution possess a finite variance; this can be a strong restriction in heavy-tailed models which constitute the favoured class of models in actuarial and financial applications. One possible solution in very heavy-tailed models where this assumption fails could be to use the more robust Median Shortfall, but this quantity is actually just a quantile, which therefore only gives information about the frequency of a tail event and not about its typical magnitude. In [65], we construct a synthetic class of tail L_p -medians, which encompasses the Median Shortfall (for $p = 1$) and Conditional Tail Expectation (for $p = 2$). We show that, for $1 < p < 2$, a tail L_p -median always takes into account both the frequency and magnitude of tail events, and its empirical estimator is, within the range of the data, asymptotically normal under a condition weaker than a finite variance. We extrapolate this estimator, along with another technique, to proper extreme levels using the heavy-tailed framework. The estimators are showcased on a simulation study and on a set of real fire insurance data showing evidence of a very heavy right tail.

A possible coherent alternative risk measure is based on expectiles [18], [63], [62]. Compared to quantiles, the family of expectiles is based on squared rather than absolute error loss minimization. The flexibility and virtues of these least squares analogues of quantiles are now well established in actuarial science, econometrics and statistical finance. Both quantiles and expectiles were embedded in the more general class of M-quantiles [19] as the minimizers of a generic asymmetric convex loss function. It has been proved very recently that the only M-quantiles that are coherent risk measures are the expectiles.

7.2.2. *Extrapolation limits associated with extreme-value methods*

Participants: Clément Albert, Stéphane Girard.

Joint work with: L. Gardes (Univ. Strasbourg) and A. Dutfoy (EDF R&D).

The PhD thesis of Clément Albert (co-funded by EDF) is dedicated to the study of the sensitivity of extreme-value methods to small changes in the data and to their extrapolation ability. Two directions are explored:

(i) In [54], we investigate the asymptotic behavior of the (relative) extrapolation error associated with some estimators of extreme quantiles based on extreme-value theory. It is shown that the extrapolation error can be interpreted as the remainder of a first order Taylor expansion. Necessary and sufficient conditions are then provided such that this error tends to zero as the sample size increases. Interestingly, in case of the so-called Exponential Tail estimator, these conditions lead to a subdivision of Gumbel maximum domain of attraction into three subsets. In constrast, the extrapolation error associated with Weissman estimator has a common behavior over the whole Fréchet maximum domain of attraction. First order equivalents of the extrapolation error are then derived and their accuracy is illustrated numerically.

(ii) In [53], We propose a new estimator for extreme quantiles under the log-generalized Weibull-tail model, introduced by Cees de Valk. This model relies on a new regular variation condition which, in some situations, permits to extrapolate further into the tails than the classical assumption in extreme-value theory. The asymptotic normality of the estimator is established and its finite sample properties are illustrated both on simulated and real datasets.

7.2.3. *Estimation of local intrinsic dimensionality with extreme-value methods*

Participant: Stéphane Girard.

Joint work with: L. Amsaleg (LinkMedia, Inria Rennes), O. Chelly (NII Japon), T. Furon (LinkMedia, Inria Rennes), M. Houle (NII Japon), K.-I. Kawarabayashi (NII Japon), M. Nett (Google).

This work is concerned with the estimation of a local measure of intrinsic dimensionality (ID). The local model can be regarded as an extension of Karger and Ruhl's expansion dimension to a statistical setting in which the distribution of distances to a query point is modeled in terms of a continuous random variable. This form of intrinsic dimensionality can be particularly useful in search, classification, outlier detection, and other contexts in machine learning, databases, and data mining, as it has been shown to be equivalent to a measure of the discriminative power of similarity functions. In [14], several estimators of local ID are proposed and analyzed based on extreme value theory, using maximum likelihood estimation, the method of moments, probability weighted moments, and regularly varying functions. An experimental evaluation is also provided, using both real and artificial data.

7.2.4. *Bayesian inference for copulas*

Participants: Julyan Arbel, Marta Crispino, Stéphane Girard.

We study in [58] a broad class of asymmetric copulas known as Liebscher copulas and defined as a combination of multiple—usually symmetric—copulas. The main thrust of this work is to provide new theoretical properties including exact tail dependence expressions and stability properties. A subclass of Liebscher copulas obtained by combining Fréchet copulas is studied in more details. We establish further dependence properties for copulas of this class and show that they are characterized by an arbitrary number of singular components. Furthermore, we introduce a novel iterative construction for general Liebscher copulas which *de facto* insures uniform margins, thus relaxing a constraint of Liebscher's original construction. Besides, we show that this iterative construction proves useful for inference by developing an Approximate Bayesian computation sampling scheme. This inferential procedure is demonstrated on simulated data.

In [22], we investigate the properties of a new transformation of copulas based on the co-copula and an univariate function. It is shown that several families in the copula literature can be interpreted as particular outputs of this transformation. Symmetry, association, ordering and dependence properties of the resulting copula are established.

7.2.5. Bayesian nonparametric clustering

Participant: Julyan Arbel.

Joint work with: Riccardo Corradin from Milano Bicocca, Michal Lewandowski from Bocconi University, Milan, Italy, Caroline Lawless from Université Paris-Dauphine, France.

For a long time, the Dirichlet process has been the gold standard discrete random measure in Bayesian nonparametrics. The Pitman–Yor process provides a simple and mathematically tractable generalization, allowing for a very flexible control of the clustering behaviour. Two commonly used representations of the Pitman–Yor process are the stick-breaking process and the Chinese restaurant process. The former is a constructive representation of the process which turns out very handy for practical implementation, while the latter describes the partition distribution induced. Obtaining one from the other is usually done indirectly with use of measure theory. In contrast, we propose in [66] an elementary proof of Pitman–Yor's Chinese Restaurant process from its stick-breaking representation.

In the discussion paper [56], we propose a simulation study to emphasise the difference between Variation of Information and Binder's loss functions in terms of number of clusters estimated by means of the use of the Markov chain Monte Carlo output only and a "greedy" method.

The chapter [47] is part of a book edited by Stéphane Girard and Julyan Arbel. It presents a Bayesian nonparametric approach to clustering, which is particularly relevant when the number of components in the clustering is unknown. The approach is illustrated with the Milky Way's globulars, that are clouds of stars orbiting in our galaxy. Clustering globulars is key for better understanding the Milky Way's history. We define the Dirichlet process and illustrate some alternative definitions such as the Chinese restaurant process, the Pólya Urn, the Ewens sampling formula, the stick-breaking representation through some simple *R* code. The Dirichlet process mixture model is presented, as well as the *R* package *BNPmix* implementing Markov chain Monte Carlo sampling. Inference for the clustering is done with the variation of information loss function.

7.2.6. Multi sensor fusion for acoustic surveillance and monitoring

Participants: Florence Forbes, Jean-Michel Bécu.

Joint work with: Pascal Vouagner and Christophe Thirard from ACOEM company.

In the context of the DGA-rapid WIFUZ project, we addressed the issue of determining the localization of shots from multiple measurements coming from multiple sensors. The WIFUZ project is a collaborative work between various partners: DGA, ACOEM and HIKOB companies and Inria. This project is at the intersection of data fusion, statistics, machine learning and acoustic signal processing. The general context is the surveillance and monitoring of a zone acoustic state from data acquired at a continuous rate by a set of sensors that are potentially mobile and of different nature. The overall objective is to develop a prototype for surveillance and monitoring that is able to combine multi sensor data coming from acoustic sensors (microphones and antennas) and optical sensors (infrared cameras) and to distribute the processing to multiple algorithmic blocs. As an illustration, the MISTIS contribution is to develop technical and scientific solutions as part of a collaborative protection approach, ideally used to guide the best coordinated response between the different vehicles of a military convoy. Indeed, in the case of an attack on a convoy, identifying the threatened vehicles and the origin of the threat is necessary to organize the best response from all members on the convoy. Thus it will be possible to react to the first contact (emergency detection) to provide the best answer for threatened vehicles (escape, lure) and for those not threatened (suppression fire, riposte fire). We developed statistical tools that make it possible to analyze this information (characterization of the threat) using fusion of acoustic and image data from a set of sensors located on various vehicles. We used Bayesian inversion and simulation techniques to recover multiple sources mimicking collaborative interaction between several vehicles.

7.2.7. Extraction and data analysis toward "industry of the future"

Participants: Florence Forbes, Hongliang Lu, Fatima Fofana, Jaime Eduardo Arias Almeida.

Joint work with: J. F. Cuccaro and J. C Trochet from Vi-Technology company.

Industry as we know it today will soon disappear. In the future, the machines which constitute the manufacturing process will communicate automatically as to optimize its performance as whole. Transmitted information essentially will be of statistical nature. In the context of VISION 4.0 project with Vi-Technology, the role of MISTIS is to identify what statistical methods might be useful for the printed circuits boards assembly industry. The topic of F. Fofana's internship was to extract and analyze data from two inspection machines of a industrial process making electronic cards. After a first extraction step in the SQL database, the goal was to enlighten the statistical links between these machines. Preliminary experiments and results on the Solder Paste Inspection (SPI) step, at the beginning of the line, helped identifying potentially relevant variables and measurements (eg related to stencil offsets) to identify future defects and discriminate between them. More generally, we have access to two databases at both ends (SPI and Component Inspection) of the assembly process. The goal is to improve our understanding of interactions in the assembly process, find out correlations between defects and physical measures, generate proactive alarms so as to detect departures from normality.

7.2.8. Change point detection for the analysis of dynamic single molecules

Participants: Florence Forbes, Theo Moins.

Joint work with: Virginie Stoppin-Mellet from Grenoble Institute of Neuroscience.

The objective of this study was to develop a statistical learning technique to analyze signals produced by molecules. The main difficulties are the noisy nature of the signals and the definition of a quality index to allow the elimination of poor-quality data and false positive signals. In collaboration with the GIN, we addressed the statistical analysis of intensity traces (2 month internship of Theo Moins, Ensimag 2A). Namely, the ImageJ Thunderstorm toolbox, which has been developed for the detection of single molecule in super resolution imaging, has been successfully used to detect immobile single molecules and generate time-dependent intensity traces. Then the R package Segmentor3IsBack, a fast segmentation algorithm based on 5 possible statistical models, proved efficient in the processing of the noisy intensity traces. This preliminary study led to a multidisciplinary project funded by the Grenoble data institute for 2 years in which we will also address additional challenges for the tracking of a large population of single molecules.

7.3. Graphical and Markov models

7.3.1. Fast Bayesian network structure learning using quasi-determinism screening

Participants: Thibaud Rahier, Stéphane Girard, Florence Forbes.

Joint work with: Sylvain Marié, Schneider Electric.

Learning the structure of Bayesian networks from data is a NP-Hard problem that involves an optimization task on a super-exponential sized space. In this work, we show that in most real life datasets, a number of the arcs contained in the final structure can be prescreened at low computational cost with a limited impact on the global graph score. We formalize the identification of these arcs via the notion of quasi-determinism, and propose an associated algorithm that reduces the structure learning to a subset of the original variables. We show, on diverse benchmark datasets, that this algorithm exhibits a significant decrease in computational time and complexity for only a little decrease in performance score. A first version of this work can be found in [71] and has been presented at the JFRB 2018 workshop [41].

7.3.2. Robust structure learning using multivariate t-distributions

Participants: Karina Ashurbekova, Florence Forbes.

Joint work with: Sophie Achard, senior researcher at CNRS, Gipsa-lab.

Structure learning is an active topic nowadays in different application areas, i.e. genetics, neuroscience. We addressed the issue of robust graph structure learning in continuous settings. We focused on sparse precision matrix estimation for its tractability and ability to reveal some measure of dependence between variables. For this purpose, we proposed to extract good features from existing methods, namely *tlasso* and CLIME procedures. The former is based on the observation that standard Gaussian modelling results in procedures that are too sensitive to outliers and proposes the use of t-distributions as an alternative. The latter is an alternative to the popular Lasso optimization principle which can handle some of its limitations. We then combined these ideas into a new procedure referred to as tCLIME that can be seen as a modified *tlasso* algorithm. Numerical performance was investigated using simulated data and reveals that tCLIME performs favorably compared to the other standard methods. This work was presented at the Journées de Statistiques de la Société Francaise de Statistique in Saclay, 2018, [39].

7.3.3. Structure learning via Hadamard product of correlation and partial correlation matrices

Participants: Karina Ashurbekova, Florence Forbes.

Joint work with: Sophie Achard, senior researcher at CNRS, Gipsa-lab.

Classical conditional independences or marginal independences may not be sufficient to express complex relationships. In this work we introduced a new structure learning procedure where an edge in the graph corresponds to a non zero of both correlation and partial correlation. A theoretical study was derived which shows the good properties of the proposed graph estimator, illustrated also on a synthetic example.

7.3.4. Spatial mixtures of multiple scaled t-distributions

Participants: Florence Forbes, Alexis Arnaud.

Joint work with: Steven Quinito Masnada, Inria Grenoble Rhone-Alpes

The goal is to implement an hidden Markov model version of our recently introduced mixtures of non standard multiple scaled t-distributions. The motivation for doing that is the application to multiparametric MRI data for lesion analysis. When dealing with MRI human data, spatial information is of primary importance. For our preliminary study on rat data [16], the results without spatial information were already quite smooth. The main anatomical structures can be identified. We suspect the reason is that the measured parameters already contain a lot of information about the underlying tissues. However, introducing spatial information is always useful and is our ongoing work. In the statistical framework we have developed (mixture models and EM algorithm), it is conceptually straightforward to introduce an additional Markov random field. In addition, when using a Markov random field it is easy to incorporate additional atlas information.

7.3.5. Spectral CT reconstruction with an explicit photon-counting detector model: a "one-step" approach

Participants: Florence Forbes, Pierre-Antoine Rodesch.

Joint work with: Veronique Rebuffel and Clarisse Fournier from CEA-LETI Grenoble.

In the context of Pierre-Antoine Rodesh's PhD thesis, we investigate new statistical and optimization methods for tomographic reconstruction from non standard detectors providing multiple energy signals. Recent developments in energy-discriminating Photon-Counting Detector (PCD) enable new horizons for spectral CT. With PCDs, new reconstruction methods take advantage of the spectral information measured through energy measurement bins. However PCDs have serious spectral distortion issues due to charge-sharing, fluorescence escape, pileup effect Spectral CT with PCDs can be decomposed into two problems: a noisy geometric inversion problem (as in standard CT) and an additional PCD spectral degradation problem. The aim of this study is to introduce a reconstruction method which solves both problems simultaneously: a one-step approach. An explicit linear detector model is used and characterized by a Detector Response Matrix (DRM). The algorithm reconstructs two basis material maps from energy-window transmission data. The results prove that the simultaneous inversion of both problems is well performed for simulation data. For comparison, we also perform a standard two-step approach: an advanced polynomial decomposition of measured sinograms combined with a filtered-back projection reconstruction. The results demonstrate the potential uses of this method for medical imaging or for non-destructive control in industry. Preliminary results have been presented at the SPIE medical imaging 2018 conference in Houston, USA [37].

7.3.6. Non parametric Bayesian priors for hidden Markov random fields

Participants: Florence Forbes, Julyan Arbel, Hongliang Lu.

Hidden Markov random field (HMRF) models are widely used for image segmentation or more generally for clustering data under spatial constraints. They can be seen as spatial extensions of independent mixture models. As for standard mixtures, one concern is the automatic selection of the proper number of components in the mixture, or equivalently the number of states in the hidden Markov field. A number of criteria exist to select this number automatically based on penalized likelihood (eg. AIC, BIC, ICL etc.) but they usually require to run several models for different number of classes to choose the best one. Other techniques (eg. reversible jump) use a fully Bayesian setting including a prior on the class number but at the cost of prohibitive computational times. In this work, we investigate alternatives based on the more recent field of Bayesian nonparametrics. In particular, Dirichlet process mixture models (DPMM) have emerged as promising candidates for clustering applications where the number of clusters is unknown. Most applications of DPMM involve observations which are supposed to be independent. For more complex tasks such as unsupervised image segmentation with spatial relationships or dependencies between the observations, DPMM are not satisfying. This work has been presented at the Joint Statistical Meeting in Vancouver Canada [29] and at the Journées de la Statistique in Saclay [40].

7.3.7. Hidden Markov models for the analysis of eye movements

Participants: Jean-Baptiste Durand, Brice Olivier.

This research theme is supported by a LabEx PERSYVAL-Lab project-team grant.
Joint work with: Anne Guérin-Dugué (GIPSA-lab) and Benoit Lemaire (Laboratoire de Psychologie et Neurocognition)

In the last years, GIPSA-lab has developed computational models of information search in web-like materials, using data from both eye-tracking and electroencephalograms (EEGs). These data were obtained from experiments, in which subjects had to decide whether a text was related or not to a target topic presented to them beforehand. In such tasks, reading process and decision making are closely related. Statistical analysis of such data aims at deciphering underlying dependency structures in these processes. Hidden Markov models (HMMs) have been used on eye movement series to infer phases in the reading process that can be interpreted as steps in the cognitive processes leading to decision. In HMMs, each phase is associated with a state of the Markov chain. The states are observed indirectly though eye-movements. Our approach was inspired by Simola et al. (2008), but we used hidden semi-Markov models for better characterization of phase length distributions [80]. The estimated HMM highlighted contrasted reading strategies (ie, state transitions), with both individual and document-related variability. However, the characteristics of eye movements within each phase tended to be poorly discriminated. As a result, high uncertainty in the phase changes arose, and it could be difficult to relate phases to known patterns in EEGs.

This is why, as part of Brice Olivier's PhD thesis, we are developed integrated models coupling EEG and eye movements within one single HMM for better identification of the phases. Here, the coupling incorporates some delay between the transitions in both (EEG and eye-movement) chains, since EEG patterns associated to cognitive processes occur lately with respect to eye-movement phases. Moreover, EEGs and scanpaths were recorded with different time resolutions, so that some resampling scheme had to be added into the model, for the sake of synchronizing both processes. An associated EM algorithm for maximum likelihood parameter estimation was derived.

New results were obtained in the standalone analysis of the eye-movements. A comparison between the effects of three types of texts was performed, considering texts either closely related, moderately related or unrelated to the target topic.

Our goal for this coming year is to implement and validate our coupled model for jointly analyzing eye-movements and EEGs in order to improve the discrimination of the reading strategies.

7.3.8. Lossy compression of tree structures

Participant: Jean-Baptiste Durand.

Joint work with: Christophe Godin and Romain Azaïs (Inria Mosaic)

The class of self-nested trees presents remarkable compression properties because of the systematic repetition of subtrees in their structure. The aim of our work is to achieve compression of any unordered tree by finding the nearest self-nested tree. Solving this optmization problem without more assumptions is conjectured to be an NP-complete or NP-hard problem. In [34], we firstly provided a better combinatorial characterization of this specific family of trees. In particular, we showed from both theoretical and practical viewpoints that complex queries can be quickly answered in self-nested trees compared to general trees. We also presented an approximation algorithm of a tree by a self-nested one that can be used in fast prediction of edit distance between two trees.

Our goal for this coming year is to apply this approach to quantify the degree of self-nestedness of several plant species and extend first results obtained on rice panicles stating that near self-nestedness is a fairly general pattern in plants.

7.3.9. Relations between structural characteristics in rose bush and visual sensory attributes for objective evaluation of the visual quality

Participant: Jean-Baptiste Durand.

Joint work with: Gilles Galopin (QUASAV, Agrocampus Ouest)

Within ornamental horticulture context, visual quality of plants is a critical criterion for consumers looking for immediate decorative effect products. Studying links between architecture and its phenotypic plasticity in response to growing conditions and the resulting plant visual appearance represents an interesting lever to propose a new approach for managing product quality from specialized crops. Objectives of the present study were to determine whether architectural components may be identified across different growing conditions (1) to study the architectural development of a shrub over time; and (2) to predict sensory attributes data characterizing multiple visual traits of the plants. The approach addressed in this study stands on the sensory profile method using a recurrent blooming modern rose bush presented in rotation using video stimuli. Plants were cultivated under a shading gradient in three distinct environments (natural conditions, under 55% and 75% shading nets). Architecture and video of the plants were recorded during three stages, from 5 to 15 months after plant multiplication. Predictive models of visual quality were obtained with regression and variable transformation to encompass non-linear relationships [21]. The proposed approach is a way to gain a better insight into the architecture of shrub plants together with their visual appearance to target processes of interest in order to optimize growing conditions or select the most fitting genotypes across breeding programs, with respect to contrasted consumer preferences.

As a perspective, dynamic traits issued from hidden-Markov-based growth models should be used for a better characterization of visual quality, as well as identification of reiterated complexes, which are believed to play a major role in rose bush structure.

7.3.10. Bayesian neural networks

Participants: Julyan Arbel, Mariia Vladimirova.

Joint work with: Pablo Mesejo from University of Granada, Spain.

We investigate in [45] and [44] deep Bayesian neural networks with Gaussian priors on the weights and ReLU-like nonlinearities, shedding light on novel sparsity-inducing mechanisms at the level of the units of the network, both pre- and post-nonlinearities. The main thrust of the paper is to establish that the units prior distribution becomes increasingly heavy-tailed with depth. We show that first layer units are Gaussian, second layer units are sub-Exponential, and we introduce sub-Weibull distributions to characterize the deeper layers units. Bayesian neural networks with Gaussian priors are well known to induce the weight decay penalty on the weights. In contrast, our result indicates a more elaborate regularisation scheme at the level of the units. This result provides new theoretical insight on deep Bayesian neural networks, underpinning their natural shrinkage properties and practical potential.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

CIFRE PhD with SCHNEIDER (2015-2018). F. Forbes and S. Girard are the advisors of a CIFRE PhD (T. Rahier) with Schneider Electric. The other advisor is S. Marié from Schneider Electric. The goal is to develop specific data mining techniques able to merge and to take advantage of both structured and unstructured (meta)data collected by a wide variety of Schneider Electric sensors to improve the quality of insights that can be produced. The total financial support for MISTIS is of 165 keuros.

PhD contract with EDF (2016-2018). S. Girard is the advisor of a PhD (A. Clément) with EDF. The goal is to investigate sensitivity analysis and extrapolation limits in extreme-value theory with application to extreme weather events. The financial support for MISTIS is of 140 keuros.

Contract with VALEO. S. Girard and Pascal Dkengne Sielenou are involved in a study with Valeo to assess the relevance of extreme-value theory in the calibration of sensors for autonomous cars. The financial support for MISTIS is of 100 keuros.

Contract with Andritz. F. Forbes and C. Braillon (SED) are involved in a study with Andritz to elaborate metrics based on image analysis to assess the quality of nonwaven tissues. The financial support for MISTIS is of 15 keuros.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. Grenoble Idex projects

MISTIS is involved in a transdiciplinary project **NeuroCoG** and in a newly accepted cross-disciplinary project (CDP) **Risk@UGA**. F. Forbes is also a member of the executive committee and responsible for the *Data Science for life sciences* work package in another project entitled **Grenoble Alpes Data Institute**.

- The main objective of the RISK@UGA project is to provide some innovative tools both for the management of risk and crises in areas that are made vulnerable because of strong interdependencies between human, natural or technological hazards, in synergy with the conclusions of Sendai conference. The project federates a hundred researchers from Human and Social Sciences, Information & System Sciences, Geosciences and Engineering Sciences, already strongly involved in the problems of risk assessment and management, in particular natural risks. The PhD thesis of Meryem Bousebata is one of the eleven PhDs funded by this project.
- The NeuroCoG project aims at understanding the biological, neurophysiological and functional bases of behavioral and cognitive processes in normal and pathological conditions, from cells to networks and from individual to social cognition. No decisive progress can be achieved in this area without an aspiring interdisciplinary approach. The interdisciplinary ambition of NeuroCoG is particularly strong, bringing together the best scientists, engineers and clinicians at the crossroads of experimental and life sciences, human and social sciences and information and communication sciences, to answer major questions on the workings of the brain and of cognition. One of the work package entitled InnobioPark is dedicated to Parkinson's Disease. The PhD thesis of Veronica Munoz Ramirez is one of the three PhDs in this work package.
- The Grenoble Alpes Data Institute aims at undertaking groundbreaking interdisciplinary research focusing on how data change science and society. It combines three fields of data-related research in a unique way: data science applied to spatial and environmental sciences, biology, and health sciences; data-driven research as a major tool in Social Sciences and Humanities; and studies about data governance, security and the protection of data and privacy. In this context, a 2-year multi-disciplinary projects has been granted in November 2018 to Mistis in collaboration with the Grenoble Institute of Neuroscience. The objective of this project is to develop a statistical learning technique that is able to solve a problem of tracking and analyzing a large population of single molecules. The main difficulties are: 1) the large number of observations to analyse, 2) the noisy nature of the signals, 3) the definition of a quality index to allow the elimination of poor-quality data and false positive signals. We also aim at providing a powerful, well-documented and open-source software, that will be user-friendly for non-specialists.

Also in the context of the Idex associated with the Université Grenoble Alpes, Alexandre Constantin was awarded half a PhD funding from IRS (Initiatives de Recherche Stratégique), 50 keuros.

9.1.2. Competitivity Clusters

The MINALOGIC VISION 4.0 project:MISTIS is involved in a three-year (2016-19) project. The project is led by VI-Technology, a world leader in Automated Optical Inspection (AOI) of a broad range of electronic components. The other partners are the G-Scop Lab in Grenoble and ACTIA company based in Toulouse. Vision 4.0 (in short Vi4.2) is one of the 8 projects labeled by Minalogic, the digital technology competitiveness cluster in Auvergne-Rhône-Alpes, that has been selected for the Industry 4.0 topic in 2016, as part of the 22nd call for projects of the FUI-Régions, for a total budget of the project of 3,4 Meuros.

Today, in the printed circuits boards (PCB) assembly industry, the assembly of electronic cards is a succession of ultra automated steps. Manufacturers, in constant quest for productivity, face sensitive and complex adjustments to reach ever higher levels of quality. Project VI4.2 proposes to build an innovative software solution to facilitate these adjustments, from images and measures obtained in automatic optical inspection (AOI). The idea is - from a centralized station for all the assembly line devices - to analyze and model the defects finely, to adjust each automatic machine, and to configure the interconnection logic between them to improve the quality. Transmitted information is essentially of statistical nature and the role of sc mistis is to identify which statistical methods might be useful to exploit at best the large amount of data registered by AOI machines. Preliminary experiments and results on the Solder Paste Inspection (SPI) step, at the beginning of the assembly line, helped determining candidate variables and measurements to identify future defects and to discriminate between them. More generally, the idea is to analyze two databases at both ends (SPI and Component Inspection) of the assembly process so as to improve our understanding of interactions in the assembly process, find out correlations between defects and physical measures and generate accordingly proactive alarms so as to detect as early as possible departures from normality.

9.1.3. Networks

MSTGA and AIGM INRA (French National Institute for Agricultural Research) networks: F. Forbes and J.B Durand are members of the INRA network called AIGM (ex MSTGA) network since 2006, http://carlit.toulouse.inra.fr/AIGM, on Algorithmic issues for Inference in Graphical Models. It is funded by INRA MIA and RNSC/ISC Paris. This network gathers researchers from different disciplines. MISTIS co-organized and hosted 2 of the network meetings in 2008 and 2015 in Grenoble. Thibaud Rahier and Brice Olivier participated as speakers to meetings of the network in 2018.

9.2. International Initiatives

9.2.1. Inria International Labs

International Laboratory for Research in Computer Science and Applied Mathematics
Associate Team involved in the International Lab:

9.2.1.1. SIMERG2E

Title: Statistical Inference for the Management of Extreme Risks, Genetics and Global Epidemiology

International Partner:

UGB (Senegal) - Abdou Kâ Diongue

Start year: 2018

See also: http://mistis.inrialpes.fr/simerge

SIMERG2E is built on the same two research themes as SIMERGE, with some adaptations to new applications: 1) Spatial extremes, application to management of extreme risks. We address the definition of new risk measures, the study of their properties in case of extreme events and their estimation from data and covariate information. Our goal is to obtain estimators accounting for possible variability, both in terms of space and time, which is of prime importance in many hydrological, agricultural and energy contexts. 2) Classification, application to genetics and global epidemiology. We address the challenge to build statistical models in order to test association between diseases and human host genetics in a context of genome-wide screening. Adequate models should allow to handle complexity in genomic data (correlation between genetic markers, high dimensionality) and additional statistical issues present in data collected from a family-based longitudinal survey (non-independence between individuals due to familial relationship and non-independence within individuals due to repeated measurements on a same person over time).

9.2.2. Inria International Partners

9.2.2.1. Informal International Partners

The context of our research is also the collaboration between MISTIS and a number of international partners such as the statistics department of University of Michigan, in Ann Arbor, USA, the statistics department of McGill University in Montreal, Canada, Université Gaston Berger in Senegal and Universities of Melbourne and Brisbane in Australia.

The main active international collaborations in 2018 are with:

- G. Stupfler, Nottingham University, UK.
- K. Qin, H. Nguyen and Kerrie Mengersen, D. Wraith resp. from Swinburne University and La Trobe university in Melbourne, Australia and Queensland University of Technology in Brisbane, Australia.
- E. Deme and S. Sylla from Gaston Berger university and IRD in Senegal.
- M. Houle from National Institute of Informatics, Tokyo, Japan.
- N. Wang and C-C. Tu from University of Michigan, Ann Arbor, USA.
- R. Steele, from McGill university, Montreal, Canada.
- Guillaume Kon Kam King, Stefano Favaro, Pierpaolo De Blasi, Collegio Carlo Alberto, Turin, Italy.
- Igor Prünster, Antonio Lijoi, and Riccardo Corradin Bocconi University, Milan, Italy.
- Bernardo Nipoti, Trinity College Dublin, Ireland.
- Yeh Whye Teh, Oxford University and DeepMind, UK.
- Stephen Walker, University of Texas at Austin, USA.

9.3. International Research Visitors

9.3.1. Visits of International Scientists

- Hien Nguyen, researcher at La Trobe University in Melbourne visited for a month in October 2018.
- Eric Marchand Professor at University of Sherbrook Canada, visited from March to June 2018.
- Riccardo Corradin, PhD student at Bocconi University, Milan, Italy visited for a month in March 2018.
- Aboubacrène Ag Ahmad, PhD student at Univ. Gaston Berger, Senegal visited from September 2018 until November 2018.

9.3.1.1. Internships

Caroline Lawless from University College Dublin visited for 2 months as part of her internship.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Florence Forbes, Stéphane Girard and Julyan Arbel organized the two-day workshop Bayesian learning theory for complex data modelling, on September 6-7 2018.

10.1.1.2. Member of the Organizing Committees

- Florence Forbes was a member of the scientific committee of the 50th journées de statistique of Société Francaise de Statistique (JDS 2018) organized in Saclay.
- Julyan Arbel co-organized the two-day workshop entitled Workshop sur la dynamique des communautés sur Twitter en période électorale : analyse par graphes aléatoires workshop on random graphs in Grenoble on April 26-27 2018.
- Julyan Arbel co-organized with Richard Nickl, Cambridge University, a session entitled Bayesian nonparametrics for stochastic processes at International Society for Bayesian Analysis (ISBA) World Meeting 2018 in Edinburgh.
- Jean-Baptiste Durand co-organized a three-day workshop on Models and Analysis of Eye Movements in Grenoble on June 6-8 2018 (https://eyemovements.sciencesconf.org/).

Seminars organization

- MISTIS participates in the weekly statistical seminar of Grenoble. Several lecturers have been invited in this context.
- Florence Forbes, Julyan Arbel and Marta Crispino are co-organizing a monthly reading group on Bayesian statistics.

10.1.2. Scientific Events Selection

10.1.2.1. Reviewer

In 2018, Florence Forbes, Stéphane Girard and Julyan Arbel have been a reviewer for *Journées de la Statistique* (JDS 2018). Additionally,

- In 2018, Julyan Arbel has been a reviewer for
  - Statistics Conferences: *Bayesian Young Statisticians Meeting proceedings* (BAYSM),
  - Machine Learning Conferences: *Conference on Neural Information Processing Systems* (NIPS), *International Conference on Learning Representations* (ICLR), *Symposium on Advances in Approximate Bayesian Inference* (AABI).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Stéphane Girard is Associate Editor of the *Statistics and Computing* journal since 2012 and Associate Editor of the *Journal of Multivariate Analysis* since 2016. He is also member of the Advisory Board of the *Dependence Modelling* journal since December 2014.
- Florence Forbes is Associate Editor of the journal Frontiers in ICT: Computer Image Analysis since its creation in Sept. 2014. She is also Associate Editor of the *Computational Statistics and Data Analysis* journal since May 2018.
- Julyan Arbel is Associate Editor of the *Bayesian Analysis* (BA) journal.

10.1.3.2. Reviewer - Reviewing Activities

- In 2018, Florence Forbes has been a reviewer for *Ecological Modelling* journal.
- In 2018, Stéphane Girard has been a reviewer for *Annals of the Institute of Statistical Mathematics, Statistics & Risk Modeling, Communications in Statistics - Theory and Methods, Extremes.*
- In 2018, Jean-Baptiste Durand has been a reviewer for *Behavior Research Methods* (BRM) and a guest editor for *PLOS Computational Biology* (PLOS Comput. Biol.).
- In 2018, Julyan Arbel has been a reviewer for: *Annals of Statistics* (AoS), *Bayesian Analysis* (BA), *Brazilian Journal of Probability and Statistics* (BJPS), *Computational Statistics & Data Analysis* (CSDA), *Electronic Journal of Statistics* (EJS), *Journal of Nonparametric Statistics* (JNS), *Scandinavian Journal of Statistics* (SJS), *Statistics and Probability Letters* (SPL).

10.1.4. Invited Talks

Florence Forbes has been invited to give talks at the following seminars and conferences:

- Data Science Seminar Series, December 2018 (Link) .
- 11th International Conference of Computational and Methodological Statistics (CMStat), University of Pisa, Italy, December 14-16, [60].
- NeuroCog Seminar Series, October 2018.
- La Trobe-Kyushu Joint Seminar on Mathematics for Industry, Melbourne, October 2018 (Link).
- Joint Statistical Meeting of the American Statistical Association, Vancouver Canada, July [29]
- Workshop on Bayesian nonparametrics, Bordeaux, France, July 2-4, [30]

Julyan Arbel has been invited to give talks at the following seminars and conferences:

- 11th International Conference of Computational and Methodological Statistics (CMStat), University of Pisa, Italy, December 14-16. Invited talk: Some distributional properties of Bayesian neural networks.
- Workshop on Bayesian nonparametrics, Bordeaux, France, July 2-4. Invited talk. Some distributional properties of Bayesian neural networks.
- Olympiades Académiques de Mathématiques, Grenoble. Talk: The mathematics of artificial intelligence.
- Trinity College Statistics Seminar, Dublin, Ireland, May 9. Invited talk: Bayesian graphs and neural networks.
- Journées statistiques de Rochebrune, Megève, France (26-30 March). Invited course: An introduction to Bayesian nonparametric statistics.
- R User group in Grenoble, France, February 8, 2018. Talk (with Alexis Arnaud): Good coding practice, coding style and R packages.
- Workshop on Statistical Methods for Post Genomic Data (SMPGD), Université de Montpellier, France, 11-12 January 2018. Invited talk: A Bayesian Nonparametric Approach to Ecological Risk Assessment.

Among the conferences listed in Section 10, [42], [43], [31], [33] were invited talks.

10.1.5. Scientific Expertise

Florence Forbes is Scientific Advisor since March 2015 for the Pixyl company.

10.1.6. Research Administration

- S. Girard is a member of the "Comité des Emplois Scientifiques" at Inria Grenoble Rhône-Alpes since 2015.
- Since 2015, S. Girard is a member of the INRA committee (CSS MBIA) in charge of evaluating INRA researchers once a year in the MBIA dept of INRA.
- S. Girard has been a reviewer of research projects for the Research Foundation Flanders (FWO), Belgium.
- Florence Forbes is a member of the "Comité Développement Technologique" for software development projects at Inria Grenoble Rhône-Alpes since 2015.
- Florence Forbes is a member of the "Comite d'organisation stratégique" of Inria Grenoble Rhône-Alpes since 2017.
- Florence Forbes is a member of the Executive Committee of the Grenoble data institute.
- Florence Forbes has been a member of the Selection committee for assistant professors at ENS Paris and a member of the Inria admission committee of junior researchers (CRCN) in June 2018.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Master : Stéphane Girard, *Statistique Inférentielle Avancée*, 18 ETD, M1 level, Ensimag. Grenoble-INP, France.
- Master : Stéphane Girard, *Data analysis, linear models and ANOVA*, 18 ETD, M1 level, MSIAM. UGA, France.
- Master and PhD course: Julyan Arbel, Bayesian statistics, Ensimag, Université Grenoble Alpes (UGA), 25 ETD.
- Master and PhD course: Julyan Arbel, Bayesian nonparametric statistics, Master Mathématiques Apprentissage et Sciences Humaines (M*A*S*H), Université Paris-Dauphine, 25 ETD.

- Master: Jean-Baptiste Durand, *Statistics and probability*, 192H, M1 and M2 levels, Ensimag Grenoble INP, France. Head of the MSIAM M2 program, in charge of the data science track.
- Jean-Baptiste Durand is a faculty member at Ensimag, Grenoble INP.

10.2.2. Supervision

- PhD defended: Clément Albert "Estimation des limites d'extrapolation par les lois de valeurs extrêmes. Application à des données environnementales", December 2018, Stéphane Girard, Université Grenoble Alpes.
- PhD defended: Thibaud Rahier "Réseaux Bayesiens pour la fusion de données statiques et temporelles", December 2018, Florence Forbes and Stéphane Girard, Université Grenoble Alpes.
- PhD defended: Pierre-Antoine Rodesch "Méthodes statistiques de reconstruction tomographique spectrale pour des systèmes à détection spectrométrique de rayons X", October 9, 2018, Florence Forbes, Université Grenoble Alpes.
- PhD defended: Alexis Arnaud "Analyse statistique d'IRM quantitatives par modèles de mélange : Application à la localisation et la caractérisation de tumeurs cérébrales", October 24, 2018, Florence Forbes and E. Barbier, Université Grenoble Alpes.
- PhD in progress: Karina Ashurbekova, "Robust Graphical Models" Florence Forbes and Sophie Achard, Université Grenoble Alpes, started on October 2016.
- PhD in progress: Veronica Munoz,"Extraction de signatures dans les données IRM de patients parkinsoniens de novo", Florence Forbes and Michel Dojat, Université Grenoble Alpes, started on October 2017.
- PhD in progress: Fabien Boux,"Développement de méthodes statistiques pour l'imagerie IRM fingerprinting", Florence Forbes and Emmanuel Barbier, Université Grenoble Alpes, started on October 2017.
- PhD in progress: Benoit Kugler, "Massive hyperspectral images analysis by inverse regression of physical models", Florence Forbes and Sylvain Douté, Université Grenoble Alpes, started on October 2018.
- PhD in progress: Chun-Chen Tu,"Gaussian mixture sub-clustering/reduction refinement of Non-linear high-to-low dimensional mapping", Florence Forbes and Naisyin Wang, University of Michigan, Ann Arbor.
- PhD in progress: Mariia Vladimirova, "Prior specification for Bayesian deep learning models and regularization implications", started on October 2018, Julyan Arbel and Jakob Verbeek.
- PhD in progress: Brice Olivier, "Joint analysis of eye-movements and EEGs using coupled hidden Markov and topic models", started on October 2015, Jean-Baptiste Durand and Anne Guérin-Dugué (Université Grenoble Alpes).
- PhD in progress: Aboubacrène Ag Ahmad "*A new location-scale model for heavy-tailed distributions*", started on September 2016, Séphane Girard and Alio Diop (Université Gaston Berger, Sénégal).
- PhD in progress: Meryem Bousebata "*Bayesian estimation of extreme risk measures: Implication for the insurance of natural disasters*", started on October 2018, Séphane Girard and Geffroy Enjolras (Université Grenoble Alpes).
- PhD in progress: Alexandre Constantin "*Analyse de séries temporelles massives d'images satellitaires: Applications à la cartographie des écosystèmes*", started on November 2018, Séphane Girard and Mathieu Fauvel (Université Grenoble Alpes).

10.2.3. Juries

- Julyan Arbel has been reviewer for the PhD thesis of Ilaria Bianchini "Modeling and computational aspects of dependent completely random measures in Bayesian nonparametric statistics", Politecnico di Milano, Italy.

- Stéphane Girard has been reviewer for the PhD thesis of Mor Absa Loum, “*Modèle de mélange et modèles linéaires généralisés, application aux données de co-infection*”, Univ. Paris-Saclay, France, et Gaston Berger, Sénégal.
- Stéphane Girard has been a member of the PhD committee of Antoine Usseglio Carleve, “*Estimation de mesures de risque pour les distributions elliptiques conditionnées*", Univ. Lyon, France.
- Florence Forbes has been reviewer for the PhD thesis of Amy Chan, University of Queensland, Brisbane and for the HDR thesis of Emilie Lebarbier, agroParisTech.
- Florence Forbes has been a member of the PhD committee of Israel Gebru Inria Grenoble, Marine Roux Gipsa-Lab, Grenoble and Jessica Sodjo, Bordeaux University.

10.3. Popularization

Florence Forbes was a speaker at the Paris Biotech Sante Forum on AI in life sciences, in November 2018, (Program).

10.3.1. Articles and contents

S. Girard and C. Albert have given an interview "When statistics help to predict disasters" for Citizen press: https://www.inria.fr/en/centre/grenoble/news/statistics-and-disasters

11. Bibliography

Major publications by the team in recent years

[1] C. AMBLARD, S. GIRARD. *Estimation procedures for a semiparametric family of bivariate copulas*, in "Journal of Computational and Graphical Statistics", 2005, vol. 14, n^{o} 2, p. 1–15

[2] J. BLANCHET, F. FORBES. *Triplet Markov fields for the supervised classification of complex structure data*, in "IEEE trans. on Pattern Analyis and Machine Intelligence", 2008, vol. 30(6), p. 1055–1067

[3] C. BOUVEYRON, S. GIRARD, C. SCHMID. *High dimensional data clustering*, in "Computational Statistics and Data Analysis", 2007, vol. 52, p. 502–519

[4] C. BOUVEYRON, S. GIRARD, C. SCHMID. *High dimensional discriminant analysis*, in "Communication in Statistics - Theory and Methods", 2007, vol. 36, n^{o} 14

[5] L. CHAARI, T. VINCENT, F. FORBES, M. DOJAT, P. CIUCIU. *Fast joint detection-estimation of evoked brain activity in event-related fMRI using a variational approach*, in "IEEE Transactions on Medical Imaging", May 2013, vol. 32, n^{o} 5, p. 821-837 [*DOI :* 10.1109/TMI.2012.2225636], http://hal.inria.fr/inserm-00753873

[6] A. DELEFORGE, F. FORBES, R. HORAUD. *High-Dimensional Regression with Gaussian Mixtures and Partially-Latent Response Variables*, in "Statistics and Computing", February 2014 [*DOI :* 10.1007/s11222-014-9461-5], https://hal.inria.fr/hal-00863468

[7] F. FORBES, G. FORT. *Combining Monte Carlo and Mean field like methods for inference in hidden Markov Random Fields*, in "IEEE trans. Image Processing", 2007, vol. 16, n^{o} 3, p. 824-837

[8] F. FORBES, D. WRAITH. *A new family of multivariate heavy-tailed distributions with variable marginal amounts of tailweights: Application to robust clustering*, in "Statistics and Computing", November 2014, vol. 24, n^{o} 6, p. 971-984 [*DOI :* 10.1007/s11222-013-9414-4], https://hal.inria.fr/hal-00823451

[9] S. GIRARD.*A Hill type estimate of the Weibull tail-coefficient*, in "Communication in Statistics - Theory and Methods", 2004, vol. 33, n^o^ 2, p. 205–234

[10] S. GIRARD, P. JACOB.*Extreme values and Haar series estimates of point process boundaries*, in "Scandinavian Journal of Statistics", 2003, vol. 30, n^o^ 2, p. 369–384

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] C. ALBERT.*Estimation of extrapolation limits based on extreme-value distributions. Application to environmental data*, Communauté Universite Grenoble Alpes, December 2018, https://hal.archives-ouvertes.fr/tel-01971408

[12] A. ARNAUD.*Statistical analysis of quantative MRI based on mixture models : Application to the localization and characterization of brain tumors*, Communauté Université Grenoble-Alpes, October 2018, https://hal.archives-ouvertes.fr/tel-01971217

[13] T. RAHIER.*Bayesian networks for static and temporal data fusion*, Communauté Université Grenoble-Alpes, December 2018, https://hal.archives-ouvertes.fr/tel-01971371

Articles in International Peer-Reviewed Journal

[14] L. AMSALEG, O. CHELLY, T. FURON, S. GIRARD, M. E. HOULE, K.-I. KAWARABAYASHI, M. NETT.*Extreme-value-theoretic estimation of local intrinsic dimensionality*, in "Data Mining and Knowledge Discovery", November 2018, vol. 32, n^o^ 6, p. 1768–1805 [*DOI :* 10.1007/s10618-018-0578-6], https://hal.archives-ouvertes.fr/hal-01864580

[15] J. ARBEL, P. DE BLASI, I. PRÜNSTER.*Stochastic approximations to the Pitman-Yor process*, in "Bayesian Analysis", 2018, p. 1-19 [*DOI :* 10.1214/18-BA1127], https://hal.archives-ouvertes.fr/hal-01950654

[16] A. ARNAUD, F. FORBES, N. COQUERY, N. COLLOMB, B. L. LEMASSON, E. L. BARBIER.*Fully Automatic Lesion Localization and Characterization: Application to Brain Tumors Using Multiparametric Quantitative MRI Data*, in "IEEE Transactions on Medical Imaging", July 2018, vol. 37, n^o^ 7, p. 1678-1689 [*DOI :* 10.1109/TMI.2018.2794918], https://hal.archives-ouvertes.fr/hal-01545548

[17] O. COMMOWICK, A. ISTACE, M. KAIN, B. LAURENT, F. LERAY, M. SIMON, S. CAMARASU-POP, P. GIRARD, R. AMELI, J.-C. FERRÉ, A. KERBRAT, T. TOURDIAS, F. CERVENANSKY, T. GLATARD, J. BEAUMONT, S. DOYLE, F. FORBES, J. KNIGHT, A. KHADEMI, A. MAHBOD, C. WANG, R. MCKINLEY, F. WAGNER, J. MUSCHELLI, E. SWEENEY, E. ROURA, X. LLADÓ, M. M. SANTOS, W. P. SANTOS, A. G. SILVA-FILHO, X. TOMAS-FERNANDEZ, H. URIEN, I. BLOCH, S. VALVERDE, M. CABEZAS, F. J. VERA-OLMOS, N. MALPICA, C. R. G. GUTTMANN, S. VUKUSIC, G. EDAN, M. DOJAT, M. STYNER, S. K. WARFIELD, F. COTTON, C. BARILLOT.*Objective Evaluation of Multiple Sclerosis Lesion Segmentation using a Data Management and Processing Infrastructure*, in "Scientific Reports", September 2018, vol. 8, 13650 [*DOI :* 10.1038/s41598-018-31911-7], https://www.hal.inserm.fr/inserm-01847873

[18] A. DAOUIA, S. GIRARD, G. STUPFLER.*Estimation of Tail Risk based on Extreme Expectiles*, in "Journal of the Royal Statistical Society: Series B", March 2018, vol. 80, n^o^ 2, p. 263–292 [*DOI :* 10.1111/RSSB.12254], https://hal.archives-ouvertes.fr/hal-01142130

[19] A. DAOUIA, S. GIRARD, G. STUPFLER.*Extreme M-quantiles as risk measures: From L1 to Lp optimization*, in "Bernoulli", 2018, forthcoming, https://hal.inria.fr/hal-01585215

[20] J. EL METHNI, L. GARDES, S. GIRARD.*Kernel estimation of extreme regression risk measures*, in "Electronic journal of statistics ", 2018, vol. 12, n^{o} 1, p. 359–398, https://hal.inria.fr/hal-01393519

[21] M. GARBEZ, R. SYMONEAUX, É. BELIN, Y. CARAGLIO, Y. CHÉNÉ, N. DONES, J.-B. DURAND, G. HUNAULT, D. RELION, M. SIGOGNE, D. ROUSSEAU, G. GALOPIN.*Ornamental plants architectural characteristics in relation to visual sensory attributes: a new approach on the rose bush for objective evaluation of the visual quality*, in "European Journal of Horticultural Science", 2018, vol. 83, n^{o} 3, p. 187-201 [*DOI :* 10.17660/EJHS.2018/83.3.8], https://hal.archives-ouvertes.fr/hal-01831318

[22] S. GIRARD.*Transformation of a copula using the associated co-copula*, in "Dependence Modeling", 2018, vol. 6, p. 298-308, https://hal.inria.fr/hal-01881969

[23] E. PERTHAME, F. FORBES, A. DELEFORGE.*Inverse regression approach to robust nonlinear high-to-low dimensional mapping*, in "Journal of Multivariate Analysis", January 2018, vol. 163, p. 1 - 14 [*DOI :* 10.1016/J.JMVA.2017.09.009], https://hal.inria.fr/hal-01347455

Invited Conferences

[24] J. ARBEL.*A Bayesian Nonparametric Approach to Ecological Risk Assessment*, in "SMPGD 2018 - Workshop on Statistical Methods for Post Genomic Data", Montpellier, France, January 2018, https://hal.archives-ouvertes.fr/hal-01950669

[25] J. ARBEL.*An introduction to Bayesian nonparametric statistics*, in "2018 - Journées statistiques de Rochebrune", Megève, France, March 2018, https://hal.archives-ouvertes.fr/hal-01950668

[26] J. ARBEL.*Bayesian neural network priors at the level of units*, in "CMStat 2018 - 11th International Conference of Computational and Methodological Statistics", Pisa, Italy, December 2018, https://hal.archives-ouvertes.fr/hal-01950661

[27] J. ARBEL.*Some distributional properties of Bayesian neural networks*, in "Workshop on Bayesian nonparametrics", Bordeaux, France, July 2018, https://hal.archives-ouvertes.fr/hal-01950667

[28] F. FORBES, A. ARNAUD, R. STEELE, B. LEMASSON, E. L. BARBIER.*Bayesian mixtures of multiple scale distributions*, in "CMStatistics 2018 - 11th International Conference of the ERCIM WG on Computational and Methodological Statistics", Pisa, Italy, December 2018, https://hal.archives-ouvertes.fr/hal-01941682

[29] F. FORBES, H. LU, J. ARBEL.*Non parametric Bayesian priors for hidden Markov random fields*, in "JSM 2018 - Joint Statistical Meeting", Vancouver, Canada, JSM Proceedings, Statistical Computing Section. Alexandria, VA: American Statistical Association, July 2018, p. 1-38, https://hal.archives-ouvertes.fr/hal-01941679

[30] F. FORBES, H. LU, J. ARBEL.*Non parametric Bayesian priors for hidden Markov random fields: application to image segmentation*, in "BNPSI 2018 : Workshop on Bayesian non parametrics for signal and image processing", Bordeaux, France, July 2018, https://hal.archives-ouvertes.fr/hal-01941687

[31] S. GIRARD.*Estimation of extreme regression risk measures*, in "2018 - Workshop Rare Events, Extremes and Machine Learning", Paris, France, May 2018, https://hal.inria.fr/hal-01800772

[32] T. RAHIER, S. MARIÉ, S. GIRARD, F. FORBES.*Screening strong pairwise relationships for fast Bayesian network structure learning*, in "2nd Italian-French Statistics Seminar - IFSS", Grenoble, France, September 2018, https://hal.archives-ouvertes.fr/hal-01941685

[33] G. STUPFLER, S. GIRARD.*Estimation of high-dimensional extreme conditional expectiles*, in "CMStatistics 2018 - 11th International Conference of the ERCIM WG on Computing and Statistics", Pisa, Italy, December 2018, https://hal.archives-ouvertes.fr/hal-01942210

International Conferences with Proceedings

[34] R. AZAÏS, J.-B. DURAND, C. GODIN.*Approximation of trees by self-nested trees*, in "ALENEX 2019 - Algorithm Engineering and Experiments", San Diego, United States, January 2019, p. 1-24, https://arxiv.org/abs/1810.10860 [*DOI :* 10.10860], https://hal.archives-ouvertes.fr/hal-01294013

[35] F. BOUX, F. FORBES, J. ARBEL, E. L. BARBIER.*Dictionary-Free MR Fingerprinting Parameter Estimation Via Inverse Regression*, in "Joint Annual Meeting ISMRM-ESMRMB 2018", Paris, France, Proceedings of Joint Annual Meeting ISMRM-ESMRMB 2018, June 2018, p. 1-2, https://hal.archives-ouvertes.fr/hal-01941630

[36] V. MUÑOZ RAMÍREZ, F. FORBES, J. ARBEL, A. ARNAUD, M. DOJAT.*Quantitative MRI characterization of brain abnormalities in de novo Parkinsonian patients*, in "IEEE International Symposium on Biomedical Imaging", Venice, Italy, Proceedings of IEEE International Symposium on Biomedical Imaging, April 2019, https://hal.archives-ouvertes.fr/hal-01970682

[37] P.-A. RODESCH, V. REBUFFEL, C. FOURNIER, F. FORBES, L. VERGER.*Spectral CT reconstruction with an explicit photon-counting detector model: a " one-step " approach*, in "SPIE Medical Imaging", Houston, United States, February 2018, vol. 10573, p. 1-4 [*DOI :* 10.1117/12.2285792], https://hal.inria.fr/hal-01652017

National Conferences with Proceeding

[38] C. ALBERT, A. DUTFOY, L. GARDES, S. GIRARD.*Un nouvel estimateur des quantiles extrêmes basé sur le modèle "log Weibull-tail" généralisé*, in "SFDS 2018 - 50èmes Journées de Statistique", Saclay, France, May 2018, p. 1-6, https://hal.inria.fr/hal-01807672

[39] K. ASHURBEKOVA, S. ACHARD, F. FORBES.*Robust structure learning using multivariate T-distributions*, in "50e Journées de la Statistique de la SFdS", Saclay, France, May 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01941643

[40] H. LU, J. ARBEL, F. FORBES.*Bayesian Nonparametric Priors for Hidden Markov Random Fields*, in "50e Journées de la Statistique de la SFdS", Saclay, France, Actes 50e Journées de la Statistique de la SFdS, May 2018, p. 1-5, https://hal.archives-ouvertes.fr/hal-01941638

[41] T. RAHIER, S. MARIÉ, S. GIRARD, F. FORBES.*Fast Bayesian Network Structure Learning using Quasi-Determinism Screening*, in "9èmes Journées Francophones sur les Réseaux Bayésiens et les Modèles Graphiques Probabilistes", Toulouse, France, Actes des 9èmes Journées Francophones sur les Réseaux Bayésiens et les Modèles Graphiques Probabilistes, June 2018, https://hal.archives-ouvertes.fr/hal-01941683

Conferences without Proceedings

[42] C. ALBERT, A. DUTFOY, S. GIRARD.*Extrapolation limits of extreme-value methods for return-levels estimation*, in "EGU General Assembly", Vienna, Austria, April 2018, https://hal.inria.fr/hal-01773660

[43] J. EL METHNI, L. GARDES, S. GIRARD.*Kernel estimation of extreme regression risk measures*, in "ICOR 2018 - 13th International Conference on Operations Research", La Havane, Cuba, March 2018, 1, https://hal.archives-ouvertes.fr/hal-01745322

[44] M. VLADIMIROVA, J. ARBEL, P. MESEJO.*Bayesian neural network priors at the level of units*, in "AABI 2018 - 1st Symposium on Advances in Approximate Bayesian Inference", Montréal, Canada, December 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01950659

[45] M. VLADIMIROVA, J. ARBEL, P. MESEJO.*Bayesian neural networks become heavier-tailed with depth*, in "NeurIPS 2018 - Thirty-second Conference on Neural Information Processing Systems", Montréal, Canada, December 2018, p. 1-7, https://hal.archives-ouvertes.fr/hal-01950658

[46] F. ZHENG, M. JALBERT, F. FORBES, S. BONNET, A. WOJTUSCISZYN, S. LABLANCHE, P.-Y. BENHAMOU.*Caractérisation de la variabilité glycémique journalière chez le patient avec diabète de type 1*, in "SFD 2018 - Congrès annuel de la Société Francophone du Diabète", Marseille, France, March 2019, https://hal.archives-ouvertes.fr/hal-01971621

Scientific Books (or Scientific Book chapters)

[47] J. ARBEL.*Clustering Milky Way's Globulars: a Bayesian Nonparametric Approach*, in "Statistics for Astrophysics: Bayesian Methodology", 2018, p. 113-137, https://hal.archives-ouvertes.fr/hal-01950656

[48] M. CLAUSEL, J.-B. DURAND.*Modèles génératifs*, in "Data Science. Cours et exercices", Eyrolles, August 2018, https://hal.inria.fr/hal-01961722

[49] F. FORBES.*Mixture Models for Image Analysis*, in "Handbook of Mixture Analysis", CRC press, December 2018, https://hal.archives-ouvertes.fr/hal-01970681

[50] C. MAGGIA, T. MISTRAL, S. DOYLE, F. FORBES, A. KRAINIK, D. GALANAUD, E. SCHMITT, S. KREMER, I. TROPRÈS, E. L. BARBIER, J.-F. PAYEN, M. DOJAT.*Traumatic Brain Lesion Quantification based on Mean Diffusivity Changes*, in "Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, BrainLes (MICCAI), ed Crimi Aea (Springer International Publishing AG)", 2018, https://hal.archives-ouvertes.fr/hal-01704679

[51] K. K. MENGERSEN, E. DUNCAN, J. ARBEL, C. ALSTON-KNOX, N. WHITE.*Applications in Industry*, in "Handbook of mixture analysis", CRC press, 2018, https://hal.archives-ouvertes.fr/hal-01963798

Other Publications

[52] A. AG AHMAD, A. DIOP, E. H. DEME, S. GIRARD.*A new location-scale model for conditional heavy-tailed distributions*, September 2018, IFSS 2018 - 2nd Italian-French Statistics Seminar, Poster, https://hal.archives-ouvertes.fr/hal-01942204

[53] C. ALBERT, A. DUTFOY, L. GARDES, S. GIRARD.*An extreme quantile estimator for the log-generalized Weibull-tail model*, September 2018, working paper or preprint, https://hal.inria.fr/hal-01783929

[54] C. ALBERT, A. DUTFOY, S. GIRARD.*Asymptotic behavior of the extrapolation error associated with the estimation of extreme quantiles*, July 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01692544

[55] J. ARBEL.*Bayesian Nonparametric Mixtures Why and How?*, September 2018, IFSS 2018 - 2nd Italian-French Statistics Seminar, Poster, https://hal.archives-ouvertes.fr/hal-01950664

[56] J. ARBEL, R. CORRADIN, M. LEWANDOWSKI.*Comment on Article by Wade and Ghahramani*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01950655

[57] J. ARBEL, R. CORRADIN, B. NIPOTI.*Dirichlet process mixtures under affine transformations of the data*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01950652

[58] J. ARBEL, M. CRISPINO, S. GIRARD.*Dependence properties and Bayesian inference for asymmetric multivariate copulas*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01963975

[59] J. ARBEL, S. FAVARO.*Approximating predictive probabilities of Gibbs-type priors*, January 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01693333

[60] A. ARNAUD, F. FORBES, R. STEELE, B. LEMASSON, E. L. BARBIER.*Bayesian mixtures of multiple scale distributions*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01953393

[61] P. BRUEL, S. QUINITO MASNADA, B. VIDEAU, A. LEGRAND, J.-M. VINCENT, A. GOLDMAN.*Autotuning under Tight Budget Constraints: A Transparent Design of Experiments Approach*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01953287

[62] A. DAOUIA, S. GIRARD, G. STUPFLER.*ExpectHill estimation, extreme risk and heavy tails*, August 2018, working paper or preprint, https://hal.inria.fr/hal-01856212

[63] A. DAOUIA, S. GIRARD, G. STUPFLER.*Tail expectile process and risk assessment*, July 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01744505

[64] D. FRAIX-BURNET, S. GIRARD, J. ARBEL, J.-B. MARQUETTE.*Statistics for Astrophysics: Bayesian Methodology*, 2018, notice, https://hal.archives-ouvertes.fr/hal-01952759

[65] L. GARDES, S. GIRARD, G. STUPFLER.*Beyond tail median and conditional tail expectation: extreme risk estimation using tail Lp -optimisation*, March 2018, working paper or preprint, https://hal.inria.fr/hal-01726328

[66] C. LAWLESS, J. ARBEL.*A simple proof of Pitman-Yor's Chinese restaurant process from its stick-breaking representation*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01950653

[67] C. LAWLESS, J. ARBEL.*Chinese restaurant process from stick-breaking for Pitman-Yor*, September 2018, 1, Bayesian learning theory for complex data modelling Workshop, Poster, https://hal.archives-ouvertes.fr/hal-01950662

[68] H. LU, J. ARBEL, F. FORBES.*Bayesian Nonparametric Priors for Hidden Markov Random Fields: Application to Image Segmentation*, September 2018, 1, IFSS 2018 - 2nd Italian-French Statistics Seminar, Poster, https://hal.archives-ouvertes.fr/hal-01950666

[69] A. MALKOVA, J. ARBEL, M. L. DELLE MONACHE.*DATASAFE: understanding Data Accidents for TrAffic SAFEty Acknowledgments*, September 2018, 1, Bayesian learning theory for complex data modelling Workshop, Poster, https://hal.archives-ouvertes.fr/hal-01950663

[70] O. MARCHAL, J. ARBEL.*Beta and Dirichlet sub-Gaussianity*, September 2018, Bayesian learning theory for complex data modelling Workshop, Poster, https://hal.archives-ouvertes.fr/hal-01950665

[71] T. RAHIER, S. MARIÉ, S. GIRARD, F. FORBES.*Fast Bayesian Network Structure Learning using Quasi-Determinism Screening*, March 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01691217

[72] C.-C. TU, F. FORBES, B. LEMASSON, N. WANG.*Structured Mixture of Linear Mappings in High Dimension* , February 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01700053

[73] M. VLADIMIROVA, J. ARBEL, P. MESEJO.*Bayesian neural network priors at the level of units*, November 2018, 1, Bayesian Statistics in the Big Data Era, Poster, https://hal.archives-ouvertes.fr/hal-01950660

[74] M. VLADIMIROVA, J. ARBEL, P. MESEJO.*Bayesian neural networks increasingly sparsify their units with depth*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01950657

References in notes

[75] C. BOUVEYRON.*Modélisation et classification des données de grande dimension. Application à l'analyse d'images*, Université Grenoble 1, septembre 2006, http://tel.archives-ouvertes.fr/tel-00109047

[76] P. EMBRECHTS, C. KLÜPPELBERG, T. MIKOSH.*Modelling Extremal Events*, Applications of Mathematics, Springer-Verlag, 1997, vol. 33

[77] F. FERRATY, P. VIEU.*Nonparametric Functional Data Analysis: Theory and Practice*, Springer Series in Statistics, Springer, 2006

[78] S. GIRARD.*Construction et apprentissage statistique de modèles auto-associatifs non-linéaires. Application à l'identification d'objets déformables en radiographie. Modélisation et classification*, Université de Cery-Pontoise, octobre 1996

[79] K. LI.*Sliced inverse regression for dimension reduction*, in "Journal of the American Statistical Association", 1991, vol. 86, p. 316–327

[80] B. OLIVIER, J.-B. DURAND, A. GUÉRIN-DUGUÉ, M. CLAUSEL.*Eye-tracking data analysis using hidden semi-Markovian models to identify and characterize reading strategies*, in "19th European Conference on Eye Movements (ECM 2017)", Wuppertal, Germany, August 2017, https://hal.inria.fr/hal-01671224

Project-Team MOEX

Evolving Knowledge

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

IN PARTNERSHIP WITH:
Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team MOEX

Creation of the Team: 2017 January 01, updated into Project-Team: 2017 November 01

Keywords:

Computer Science and Digital Science:

A3.2. - Knowledge
A3.2.1. - Knowledge bases
A3.2.2. - Knowledge extraction, cleaning
A3.2.4. - Semantic Web
A3.2.5. - Ontologies
A6.1.3. - Discrete Modeling (multi-agent, people centered)
A9. - Artificial intelligence
A9.1. - Knowledge

Other Research Topics and Application Domains:

B9. - Society and Knowledge
B9.8. - Reproducibility

1. Team, Visitors, External Collaborators

Research Scientist

Jérôme Euzenat [Team leader, Inria, Senior Researcher, HDR]

Faculty Members

Manuel Atencia Arcas [Univ Grenoble Alpes, Associate Professor]
Jérôme David [Univ Grenoble Alpes, Associate Professor]

PhD Students

Khadija Jradeh [Univ Grenoble Alpes, from Oct 2018]
Line Van Den Berg [Univ Grenoble Alpes, from Oct 2018]

Administrative Assistant

Alexandra Fitzgerald [Inria]

2. Overall Objectives

2.1. Overall Objectives

Human beings are apparently able to communicate knowledge. However, it is impossible for us to know if we share the same representation of knowledge.

mOeX addresses the evolution of knowledge representations in individuals and populations. We deal with software agents and formal knowledge representation. The ambition of the mOeX project is to answer, in particular, the following questions:

- How do agent populations *adapt* their knowledge representation to their environment and to other populations?
- How must this knowledge *evolve* when the environment changes and new populations are encountered?
- How can agents preserve knowledge *diversity* and is this diversity beneficial?

We study them chiefly in a well-controlled computer science context.

For that purpose, we combine knowledge representation and cultural evolution methods. The former provides formal models of knowledge; the latter provides a well-defined framework for studying situated evolution.

We consider knowledge as a culture and study the global properties of local adaptation operators applied by populations of agents by jointly:

- *experimentally* testing the properties of adaptation operators in various situations using experimental cultural evolution, and
- *theoretically* determining such properties by modelling how operators shape knowledge representation.

We aim at acquiring a precise understanding of knowledge evolution through the consideration of a wide range of situations, representations and adaptation operators.

In addition, we still investigate RDF data interlinking with link keys, a way to link entities in different data sets.

3. Research Program

3.1. Knowledge representation semantics

We work with semantically defined knowledge representation languages (like description logics, conceptual graphs and object-based languages). Their semantics is usually defined within model theory initially developed for logics.

We consider a language L as a set of syntactically defined expressions (often inductively defined by applying constructors over other expressions). A representation ($o \subseteq L$) is a set of such expressions. It may also be called an ontology. An interpretation function (I) is inductively defined over the structure of the language to a structure called the domain of interpretation (D). This expresses the construction of the "meaning" of an expression in function of its components. A formula is satisfied by an interpretation if it fulfills a condition (in general being interpreted over a particular subset of the domain). A model of a set of expressions is an interpretation satisfying all the expressions. A set of expressions is said consistent if it has at least one model, inconsistent otherwise. An expression (δ) is then a consequence of a set of expressions (o) if it is satisfied by all of their models (noted $o \models \delta$).

The languages dedicated to the semantic web (RDF and OWL) follow that approach. RDF is a knowledge representation language dedicated to the description of resources; OWL is designed for expressing ontologies: it describes concepts and relations that can be used within RDF.

A computer must determine if a particular expression (taken as a query, for instance) is the consequence of a set of axioms (a knowledge base). For that purpose, it uses programs, called provers, that can be based on the processing of a set of inference rules, on the construction of models or on procedural programming. These programs are able to deduce theorems (noted $o \vdash \delta$). They are said to be sound if they only find theorems which are indeed consequences and to be complete if they find all the consequences as theorems.

3.2. Data interlinking with link keys

Vast amounts of RDF data are made available on the web by various institutions providing overlapping information. To be fully exploited, different representations of the same object across various data sets, often using different ontologies, have to be identified. When different vocabularies are used for describing data, it is necessary to identify the concepts they define. This task is called ontology matching and its result is an alignment A, i.e., a set of correspondences $\langle e, r, e' \rangle$ relating entities e and e' of two different ontologies by a particular relation r (which may be equivalence, subsumption, disjointness, etc.) [3].

At the data level, data interlinking is the process of generating links identifying the same resource described in two data sets. Parallel to ontology matching, from two datasets (d and d') it generates a link set, L made of pairs of resource identifier.

We have introduced link keys [3], [1] which extend database keys in a way which is more adapted to RDF and deals with two data sets instead of a single relation. More precisely, a link key is a structure $\langle K^{eq}, K^{in}, C\rangle$ such that:

- K^{eq} and K^{in} are sets of pairs of property expressions;
- C is a pair of class expressions (or a correspondence).

Such a link key holds if and only if for any pair of resources belonging to the classes in correspondence such that the values of their property in K^{eq} are pairwise equal and the values of those in K^{in} pairwise intersect, the resources are the same. Link keys can then be used for finding equal individuals across two data sets and generating the corresponding owl:sameAs links. Link keys take into account the non functionality of RDF data and have to deal with non literal values. In particular, they may use arbitrary properties and class expressions. This renders their discovery and use difficult.

3.3. Experimental cultural knowledge evolution

Cultural evolution applies an idealised version of the theory of evolution to culture. Cultural evolution experiments are performed through multi-agent simulation: a society of agents adapts its culture through a precisely defined protocol [15]: agents perform repeatedly and randomly a specific task, called game, and their evolution is monitored. This aims at discovering experimentally the states that agents reach and the properties of these states.

Experimental cultural evolution has been successfully and convincingly applied to the evolution of natural languages [14], [16]. Agents play *language games* and adjust their vocabulary and grammar as soon as they are not able to communicate properly, i.e., they misuse a term or they do not behave in the expected way. It showed its capacity to model various such games in a systematic framework and to provide convincing explanations of linguistic phenomena. Such experiments have shown how agents can agree on a colour coding system or a grammatical case system.

We adapt this experimental strategy to knowledge representation [2]. Agents use their, shared or private, knowledge to play games and, in case of failure, they use adaptation operators to modify this knowledge. We monitor the evolution of agent knowledge with respect to its ability to perform the game (success rate) and with respect to the properties satisfied by the resulting knowledge itself. Such properties may, for instance, be:

- Agents converge to a common knowledge representation (a convergence property).
- Agents converge towards different but compatible (logically consistent) knowledge (a logical epistemic property), or towards closer knowledge (a metric epistemic property).
- That under the threat of a changing environment, agents which have operators that preserve diverse knowledge recover faster from the changes than those which have operators that converge towards a single representation (a differential property under environment change).

Our goal is to determine which operators are suitable for achieving desired properties in the context of a particular game.

4. New Results

4.1. Cultural knowledge evolution

Our cultural knowledge evolution work currently focusses on alignment evolution.

Agents may use ontology alignments to communicate when they represent knowledge with different ontologies: alignments help reclassifying objects from one ontology to the other. Such alignments may be provided by dedicated algorithms [7], but their accuracy is far from satisfying. Yet agents have to proceed. They can take advantage of their experience in order to evolve alignments: upon communication failure, they will adapt the alignments to avoid reproducing the same mistake.

We performed such repair experiments [2] and revealed that, by playing simple interaction games, agents can effectively repair random networks of ontologies or even create new alignments.

4.1.1. Strengthening modality for cultural alignment repair

Participants: Jérôme Euzenat [Correspondent], Iris Lohja.

Our previous work on cultural alignment repair achieved 100% precision for all adaptation operators, i.e., all the correspondences in the alignments were correct, but were still missing some correspondences, and did not achieve 100% recall. We had conjectured that this was due to a phenomenon called reverse shadowing [2], avoiding to find specific correspondences.

This year we introduced a new adaptation modality, strengthening, to test this hypothesis. The strengthening modality replaces a successful correspondence by one of its subsumed correspondences covering the current instance. This modality is different from those developed so far, because it leads agents to adapt their alignment when the game played has been a success (previously, it was always when a failure occurred). We defined three alternative definitions of this modality depending on if the agent chooses the most general, most specific or a random such correspondence.

The strengthening modality has been implemented in our *Lazy lavender* software. We experimentally showed that it was not interferring with the other modalities as soon as the *add* operator was used. This means that all properties of the previous adaptation operators are preserved. Moreover, as expected, recall was greatly increased, to the point that some operators achieve 99% F-measure. However, the agents still do not reach 100% recall.

4.1.2. Experiment reproducibility through container technology

Participants: Jérôme Euzenat [Correspondent], Bilal Lahmami.

Performing experiments and reporting them requires care in order for others to be able to repeat them.

We experimented with container technology in order to embed our experiments and offer to others to run them easily. To that extent, we developed scripts associated to the *Lazy lavender* software to specify, run, and analyse experiments. In particular, these scripts are able to generate a Docker container specification that can perform experiments in the same conditions or with updated software. The documentation of the experiments on our Wiki platform (https://gforge.inria.fr/plugins/mediawiki/wiki/lazylav/index.php/Lazy_Lavender) is also eased by this process.

4.2. Link keys

Link keys (§3.2) are explored following two directions:

- Extracting link keys;
- Reasoning with link keys.

4.2.1. Link key extraction with relational concept analysis

Participants: Manuel Atencia, Jérôme David [Correspondent], Jérôme Euzenat.

We have further investigated link key extraction using relational concept analysis and the associated prototype implementation [8]. In particular, we showed that that link keys extracted by formal concept analysis are equivalent to an extension of those which were extracted by our former algorithm [1]

4.2.2. Link key extraction under ontological constraints

Participants: Jérôme David [Correspondent], Jérôme Euzenat, Khadija Jradeh.

We investigated the use of link keys taking advantage of ontologies. This can be carried out in two different directions: exploiting the ontologies under which data sets are published, and extracting link keys using ontology constructors for combining attribute and class names.

Following the first approach, we extended our existing algorithms to extract link keys involving inverse ($^{-1}$), union ($\sqcup$), intersection ($\sqcap$) and paths ($\circ$) of properties. This helps providing link keys when it is not possible otherwise (without inverse, there is no possible correspondence if one data set is using parents and the other is using children). We showed how the paths could be normalised to reduce the search space. Extracting link keys under these conditions required to introduce better indexing techniques to avoid unnecessary link key generation and even looping.

We implemented this method and evaluated it by running experiments on two real data sets, this resulted in finding the correct link keys that were not found without them.

4.2.3. Tableau method for $\mathcal{ALC}$+Link key reasoning

Participants: Manuel Atencia [Correspondent], Jérôme Euzenat, Khadija Jradeh.

Link keys can also be thought of as axioms in a description logic. We further worked on the tableau method designed for the $\mathcal{ALC}$ description logic to support reasoning with link keys.

4.3. Semantic web queries

4.3.1. Evaluation of query transformations without data

Participants: Jérôme David, Jérôme Euzenat [Correspondent].

Query transformations are ubiquitous in semantic web query processing. For any situation in which transformations are not proved correct by construction, the quality of these transformations has to be evaluated. Usual evaluation measures are either overly syntactic and not very informative —the result being: correct or incorrect— or dependent from the evaluation sources. Moreover, both approaches do not necessarily yield the same result. We proposed to ground the evaluation on query containment [4]. This allows for a data-independent evaluation that is more informative than the usual syntactic evaluation. In addition, such evaluation modalities may take into account ontologies, alignments or different query languages as soon as they are relevant to query evaluation [6].

5. Partnerships and Cooperations

5.1. National Initiatives

5.1.1. ANR Elker

Program: ANR-PRC

Project acronym: ELKER

Project title: Extending link keys: extraction and reasoning

Duration: October 2017 - September 2021

Coordinator: LIG/Manuel Atencia

Participants: Manuel Atencia Arcas, Jérôme David, Jérôme Euzenat

Other partners: Inria Lorraine, Université de Vincennes

Abstract: The goal of ELKER is to extend the foundations and algorithms of link keys (see §3.2) in two complementary ways: extracting link keys automatically from datasets and reasoning with link keys.

5.1.2. Framework agreement Ministère de la culture et de la communication

Program: Framework agreement Inria-Ministère de la culture et de la communication

Project acronym: GINCO V3

Project title: Outil d'aide à l'alignement pour l'élaboration du graphe culture

Duration: November 2017 - December 2018

Coordinator: Jérôme David

Participants: Jérôme David, Jérôme Euzenat, Manuel Atencia Arcas

Abstract: The GINCO V3 project aims at extending the GINCO tool with ontology alignment capabilities.

Program: Framework agreement Inria-Ministère de la culture et de la communication

Project acronym: FNE

Project title: Algorithmes d'aide à la définition de clés de liage et d'alignement d'autorités

Duration: November 2017 - December 2018

Coordinator: Jérôme David

Participants: Jérôme David, Manuel Atencia Arcas, Jérôme Euzenat

Other partners: Bibliothèque nationale de France

Abstract: The goal of the FNE cooperation is to evaluate the suitability of link key extraction algorithms to matching authorities from BnF, ABES and the ministry of Culture and to improve such algorithms if necessary.

6. Dissemination

6.1. Promoting Scientific Activities

6.1.1. Scientific Events Organisation

6.1.1.1. Member of the Organizing Committees

- Jérôme David had been organiser of the workshop Symbolic methods for data-interlinking of the 21st EKAW, Nancy (FR), 2018 (with Miguel Couceiro)
- Jérôme Euzenat had been organiser of the 13th Ontology matching workshop of the 18th ISWC, Monterey (CA US), 2018 (with Pavel Shvaiko, Ernesto Jiménez Ruiz, Michele Cheatham and Oktie Hassanzadeh)

6.1.2. Scientific Events Selection

6.1.2.1. Chair of Conference Program Committees

- Jérôme Euzenat had been program chairman of the "French national artificial intelligence conference (CNIA)" [10], Nancy (FR), 2018.
- Manuel Atencia had been workshop and tutorial chairman (with Marieke van Erp) of the 21st EKAW, Nancy (FR), 2018.

6.1.2.2. Member of the Conference Program Committees

- Jérôme David and Jérôme Euzenat had been programme committee members of the "International joint conference on artificial intelligence (IJCAI)" 2018
- Jérôme David and Jérôme Euzenat had been programme committee member of the "National conference on artificial intelligence (AAAI)" 2018

- Manuel Atencia and Jérôme Euzenat had been programme committee members of the “International semantic web conference (ISWC)” 2018
- Jérôme David and Jérôme Euzenat had been programme committee members of the “European Semantic Web Conference (ESWC)” 2017
- Manuel Atencia, Jérôme David and Jérôme Euzenat had been programme committee member of the “Web Conference (WWW)” 2018
- Jérôme David had been programme committee member of the “Pacific Rim Knowledge Acquisition Workshop (PKAW)” 2018
- Jérôme Euzenat had been programme committee member of the “International Conference on Semantic Systems (Semantics)” 2018
- Jérôme Euzenat had been programme committee member of the “International conference on knowledge engineering and knowledge management (EKAW)” 2018
- Jérôme David had been programme committee member of the “French national artificial intelligence conference (CNIA)” 2018.
- Jérôme Euzenat had been programme committee member of the “French fundamental artificial intelligence days (JAIF)” 2018
- Manuel Atencia and Jérôme David had been programme committee members of the ISWC “Ontology matching” workshop (OM) 2018
- Manuel Atencia and Jérôme David had been programme committee members of the “French Extraction and gestion des connaissances conference (EGC)” 2018
- Jérôme David had been programme committee member of the “29es Journées francophones d’ingénierie des connaissances (IC)” 2018

6.1.3. Journal

6.1.3.1. Member of the Editorial Boards

- Jérôme Euzenat is member of the editorial board of *Journal of web semantics* (area editor), *Journal on data semantics* and the *Semantic web journal*.
- Jérôme Euzenat had been guest editor of a special issue of *Semantic web journal* on “semantic technologies and interoperability in the build environment” (with Álvaro Sicilia, Pieter Pauwels, Leandro Madrazo, and María Poveda-Villalón).

6.1.3.2. Reviewer - Reviewing Activities

- Manuel Atencia had been reviewer for *Journal of web semantics* and *Applied ontology*.
- Jérôme David had been reviewer for *Information system journal* and *Ingénierie des systèmes d’information* (special issue on “Impact des Open et/ou Linked Data dans les systèmes d’information”).
- Jérôme Euzenat had been reviewer for *ACM transactions on the web*, and *International journal on metadata, semantics and ontologies*.

6.1.4. Invited Talks

- “Towards cultural knowledge evolution: experiments with alignments repair”, Seminar LIP6, Paris (FR), 2018-10-15 (Jérôme Euzenat)
- “Enhancing Link Keys: Extraction and Reasoning”, BNF workshop on “Données liées et données à lier: quels outils pour quels alignements?”, Paris (FR), 2018-07-10 (Manuel Atencia)
- “Link key extraction with a variation of relational concept analysis”, EKAW workshop on symbolic methods for data interlinking, Nancy (FR), 2018-11-12 (Jérôme Euzenat)

6.1.5. Leadership within the Scientific Community

- Jérôme Euzenat is member of the scientific council of the CNRS GDR “Intelligence artificielle”.

6.1.6. Scientific Expertise

- Manuel Atencia had been evaluator for CAPES-COFECUB projects (BR)
- Jérôme Euzenat had been member of the scientific evaluation committee “CE23: Data, knowledge, big data, multimedia content – artificial intelligence” of the French national research agency (ANR)
- Manuel Atencia had been evaluator for the French national research agency (ANR)

6.2. Teaching - Supervision - Juries

6.2.1. Teaching

- Jérôme David is coordinator of the Master “Mathematiques et informatiques appliquées aux sciences humaines et sociales” (Univ. Grenoble Alpes)
- Manuel Atencia is co-responsible of the Master 2nd year “Mathematiques et informatiques appliquées aux sciences humaines et sociales” (Univ. Grenoble Alpes)
- Jérôme Euzenat had been, with Danielle Ziébelin, coordinator of the “AI and the web” option of the M2R in computer science and applied mathematics (Univ. Grenoble Alpes)

6.2.1.1. Lectures

Licence: Jérôme David, Algorithmique et programmation par objets, 70h, L2 MIASHS, UGA, France

Licence: Jérôme David, Introduction à Python, Licence ESSIG, 24h, UGA, France

Licence: Jérôme David, Système, L3 MIASHS, 18h, UGA, France

Licence: Manuel Atencia, Technologies du web, LP ESSIG, 18h, UGA, France

Licence: Manuel Atencia, Introduction aux technologies du Web, 60h, L3 MIASHS, UGA, France

Master: Jérôme David, Programmation Java 2, 30h, M1 MIASHS, UGA, France

Master: Jérôme David, JavaEE, 30h, M2 MIASHS, UGA, France

Master: Jérôme David, Développement Web Mobile, 30h, M2 MIASHS, UGA, France

Master: Jérôme David, Web sémantique, 3h, M2 MIASHS, UGA, France

Master: Manuel Atencia, Formats de données du web, 15h, M1 MIASHS, UGA, France

Master: Manuel Atencia, Introduction à la programmation web, 30h, M1 MIASHS, UGA, France

Master: Manuel Atencia, Intelligence artificielle, 7.5h, M1 MIASHS, UGA, France

Master: Manuel Atencia, Web sémantique, 27h, M2 MIASHS, UGA, France

Master: Manuel Atencia, Semantic web: from XML to OWL, 22.5h, M2R MoSIG, UGA, France

Master: Jérôme David, Stage de programmation, 10h, M2 MIASHS, UGA, France

6.2.2. Supervision

- PhD in progress : Nacira Abbas, “Link key extraction and relational concept analysis”, 2018-10-01 (Jérôme David and Amedeo Napoli)
- PhD in progress : Khadija Jradeh, “Reasoning with link keys”, 2018-10-01 (Manuel Atencia and Chan Le Duc)
- PhD in progress : Line van den Berg, “Knowledge Evolution in Agent Populations”, 2018-10-01 (Manuel Atencia and Jérôme Euzenat)
- MSc: Iris Lohja, “Improving semantic recall of ontology alignments in cultural knowledge evolution”, M2R Informatics, Univ. Grenoble Alpes, June 2018 (Manuel Atencia and Jérôme Euzenat)
- MSc: Khadija Jradeh, “Link key extraction under ontological constraints”, M2R Informatics, Univ. Grenoble Alpes, June 2018 (Jérôme David and Jérôme Euzenat)

6.2.3. PhD panels

- Jérôme David had been panel member of the computer science PhD of Valentina Beratta (IMT Mines Alès) “Évaluation de la véracité des données: améliorer la découverte de la vérité en utilisant des connaissances a priori” supervised by Sylvie Ranwez and Isabelle Mougenot
- Jérôme Euzenat had been panel member of the computer science PhD of Paula Chocrón (Universitat Autònoma de Barcelona) “A pragmatic approach to translation: vocabulary alignment through multiagent interaction and observation” supervised by Marco Schorlemmer
- Jérôme Euzenat had been panel chair of the computer science PhD of Louis Jachiet (Université Grenoble-Alpes) “On the foundations for the compilation of web data queries: optimization and distributed evaluation of SPARQL” supervised by Nabil Layaïda and Pierre Genevès
- Jérôme Euzenat had been reviewer of the computer science habilitation of Chan Le Duc (Université Paris 8) “Raisonnement et révision pour des ontologies en logique de description”
- Jérôme Euzenat is member of the PhD supervision committee of Élodie Thiéblin (Université de Toulouse 2-Jean Jaurès) supervised by Ollivier Haemmerlé and Cássia Trojahn dos Santos

6.3. Popularization

6.3.1. Interventions

- Jérôme Euzenat gave a talk on “Artificial intelligence: a broad view” to BEST Spring course, Grenoble (FR), 2018-04-16
- Jérôme Euzenat gave a talk on “mOeX: évolution de la connaissance” Inria “Mon projet en 180 secondes”, Grenoble (FR), 2018-05-29
- Jérôme Euzenat gave a talk on “Evolving knowledge: different facets of Artificial Intelligence” to the French-American Doctoral Exchange (FADEx) seminar, Grenoble (FR), 2018-06-27

7. Bibliography

Major publications by the team in recent years

[1] M. ATENCIA, J. DAVID, J. EUZENAT.*Data interlinking through robust linkkey extraction*, in "Proc. 21st european conference on artificial intelligence (ECAI), Praha (CZ)", Amsterdam (NL), T. SCHAUB, G. FRIEDRICH, B. O'SULLIVAN (editors), IOS press, 2014, p. 15-20, ftp://ftp.inrialpes.fr/pub/exmo/publications/atencia2014b.pdf

[2] J. EUZENAT.*Interaction-based ontology alignment repair with expansion and relaxation*, in "Proc. 26th International Joint Conference on Artificial Intelligence (IJCAI), Melbourne (VIC AU)", 2017, p. 185–191

[3] J. EUZENAT, P. SHVAIKO.*Ontology matching*, 2nd, Springer-Verlag, Heidelberg (DE), 2013, http://book.ontologymatching.org

Publications of the year

Articles in International Peer-Reviewed Journal

[4] M. CHEKOL, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA.*SPARQL Query Containment under Schema*, in "Journal on Data Semantics", April 2018, vol. 7, n^{o} 3, p. 133-154 [*DOI :* 10.1007/s13740-018-0087-1], https://hal.inria.fr/hal-01767887

Articles in Non Peer-Reviewed Journal

[5] J. EUZENAT.*De la langue à la connaissance: approche expérimentale de l'évolution culturelle*, in "Bulletin de l'Association Française pour l'Intelligence Artificielle", April 2018, vol. 100, p. 9-12, https://hal.inria.fr/hal-01953227

International Conferences with Proceedings

[6] J. DAVID, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA.*Evaluation of Query Transformations without Data*, in "WWW 2018 - Companion of The Web Conference", Lyon, France, ACM Press, April 2018, p. 1599-1602 [*DOI :* 10.1145/3184558.3191617], https://hal.inria.fr/hal-01891182

[7] J. DA SILVA, K. REVOREDO, F. ARAUJO BAIÃO, J. EUZENAT.*Interactive ontology matching: using expert feedback to select attribute mappings*, in "13th ISWC workshop on ontology matching (OM)", Monterey, United States, No commercial editor., October 2018, p. 25-36, https://hal.archives-ouvertes.fr/hal-01964686

National Conferences with Proceeding

[8] J. DAVID, J. EUZENAT, J. VIZZINI.*Linkky: Extraction de clés de liage par une adaptation de l'analyse relationnelle de concepts*, in "IC 2018 - 29es Journées Francophones d'Ingénierie des Connaissances", Nancy, France, S. RANWEZ (editor), AFIA, July 2018, p. 271-274, https://hal.archives-ouvertes.fr/hal-01839642

Books or Proceedings Editing

[9] K. ADRIAN, J. EUZENAT, D. GROMANN (editors). *Proc. 1st JOMO workshop on Interaction-Based Knowledge Sharing (WINKS)*, No commercial editor., Bozen-Bolzano, Italy, 2018, p. 1-42, https://hal.archives-ouvertes.fr/hal-01951964

[10] J. EUZENAT, F. SCHWARZENTRUBER (editors). *Actes de la Conférence Nationale en Intelligence Artificielle et des Rencontres Jeunes Chercheurs en Intelligence Artificielle (CNIA+RJCIA)*, No commercial editor., Nancy, France, 2018, p. 1-133, https://hal.archives-ouvertes.fr/hal-01838695

[11] P. SHVAIKO, J. EUZENAT, E. JIMÉNEZ-RUIZ, M. CHEATHAM, O. HASSANZADEH (editors). *Ontology Matching: OM-2018: Proceedings of the ISWC Workshop*, No commercial editor., Monterey, United States, 2018, p. 1-227, https://hal.archives-ouvertes.fr/hal-01964687

[12] A. SICILIA, P. PAUWELS, L. MADRAZO, M. POVEDA VILLALÓN, J. EUZENAT (editors). *Semantic web journal: Special Issue on Semantic Technologies and Interoperability in the Build Environment*, IOS Press, 2018, vol. 9, n^o 6, p. 729-855, https://hal.inria.fr/hal-01953272

Other Publications

[13] P. PAUWELS, M. POVEDA VILLALÓN, A. SICILIA, J. EUZENAT.*Semantic technologies and interoperability in the built environment*, IOS Press, 2018, p. 731-734, Editorial, Semantic web journal 9(6):731-734 [*DOI :* 10.3233/SW-180321], https://hal.archives-ouvertes.fr/hal-01951968

References in notes

[14] L. STEELS (editor). *Experiments in cultural language evolution*, John Benjamins, Amsterdam (NL), 2012

[15] R. AXELROD.*The dissemination of culture: a model with local convergence and global polarization*, in "Journal of conflict resolution", 1997, vol. 41, n^{o} 2, p. 203–226

[16] M. SPRANGER.*The evolution of grounded spatial language*, Language science press, Berlin (DE), 2016

Project-Team MORPHEO

Capture and Analysis of Shapes in Motion

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

IN PARTNERSHIP WITH:
Institut polytechnique de Grenoble

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Vision, perception and multimedia interpretation

Table of contents

Project-Team MORPHEO

Creation of the Team: 2011 March 01, updated into Project-Team: 2014 January 01

Keywords:

Computer Science and Digital Science:

A5.1.8. - 3D User Interfaces
A5.4. - Computer vision
A5.4.4. - 3D and spatio-temporal reconstruction
A5.4.5. - Object tracking and motion analysis
A5.5.1. - Geometrical modeling
A5.5.4. - Animation
A5.6. - Virtual reality, augmented reality
A6.2.8. - Computational geometry and meshes

Other Research Topics and Application Domains:

B2.6.3. - Biological Imaging
B2.8. - Sports, performance, motor skills
B9.2.2. - Cinema, Television
B9.2.3. - Video games
B9.4. - Sports

1. Team, Visitors, External Collaborators

Research Scientists

Edmond Boyer [Team leader, Inria, Senior Researcher, HDR]
Julien Pansiot [Inria, Starting Research Position]
Stefanie Wuhrer [Inria, Researcher, HDR]

Faculty Members

Jean Sébastien Franco [Institut polytechnique de Grenoble, Associate Professor]
Sergi Pujades Rocamora [Univ Grenoble Alpes, Associate Professor, from Dec 2018]

External Collaborators

Franck Hétroy-Wheeler [Univ de Strasbourg, HDR]
Sergi Pujades Rocamora [Max Planck Institute for Perceiveing Systems, from Sep 2018 until Nov 2018]
Jinlong Yang [Facebook, from Oct 2018]

Technical Staff

Eymeric Amselem [Inria, from Jul 2018]
Laurence Boissieux [Inria]
Tomas Svaton [Inria, from Apr 2018]

PhD Students

Matthieu Armando [Inria]
Jean Basset [Inria, from Sep 2018]
Victoria Fernandez Abrevaya [Univ Grenoble Alpes]
Claude Goubet [Inria]
Roman Klokov [Inria]
Vincent Leroy [Inria]
Di Meng [Inria, from Oct 2018]

Abdullah Haroon Rasheed [Inria]
Romain Rombourg [Institut polytechnique de Grenoble]
Nitika Verma [Univ Grenoble Alpes]
Jinlong Yang [Inria, until Sep 2018]

Post-Doctoral Fellow

Adnane Boukhayma [Inria, until Feb 2018]

Administrative Assistant

Nathalie Gillot [Inria]

2. Overall Objectives

2.1. Overall Objectives

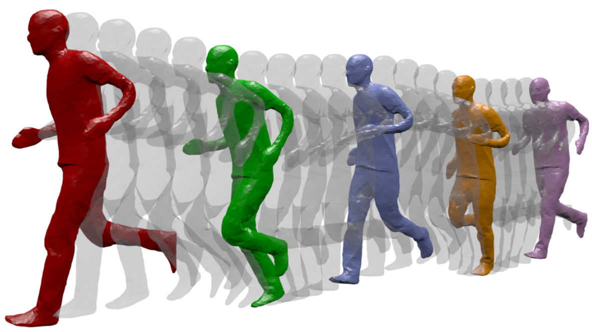


Figure 1. Dynamic Geometry Modeling

MORPHEO's ambition is to perceive and to interpret shapes that move using multiple camera systems. Departing from standard motion capture systems, based on markers, that provide only sparse information on moving shapes, multiple camera systems allow dense information on both shapes and their motion to be recovered from visual cues. Such ability to perceive shapes in motion brings a rich domain for research investigations on how to model, understand and animate real dynamic shapes, and finds applications, for instance, in gait analysis, bio-metric and bio-mechanical analysis, animation, games and, more insistently in recent years, in the virtual and augmented reality domain. The MORPHEO team particularly focuses on three different axes within the overall theme of 3D dynamic scene vision or 4D vision:

1. Shape and appearance models: how to build precise geometric and photometric models of shapes, including human bodies but not limited to, given temporal sequences.
2. Dynamic shape vision: how to register and track moving shapes, build pose spaces and animate captured shapes.
3. Inside shape vision: how to capture and model inside parts of moving shapes using combined color and X-ray imaging.

The strategy developed by MORPHEO to address the mentioned challenges is based on methodological tools that include in particular geometry, Bayesian inference and numerical optimization. Following the evolution in computer vision, our strategy has also evolved towards data driven approaches, as they have proved to be beneficial on different components of 3D vision solutions. Thus, our methodology include now machine learning tools whose potential in 4D vision are still to be fully investigated.

3. Research Program

3.1. Shape and Appearance Modeling

Standard acquisition platforms, including commercial solutions proposed by companies such as Microsoft, 3dMD or 4DViews, now give access to precise 3D models with geometry, e.g. meshes, and appearance information, e.g. textures. Still, state-of-the-art solutions are limited in many respects: They generally consider limited contexts and close setups with typically at most a few meter side lengths. As a result, many dynamic scenes, even a body running sequence, are still challenging situations; They also seldom exploit time redundancy; Additionally, data driven strategies are yet to be fully investigated in the field. The MORPHEO team builds on the Kinovis platform for data acquisition and has addressed these issues with, in particular, contributions on time integration, in order to increase the resolution for both for shapes and appearances, on representations, as well as on exploiting recent machine learning tools when modeling dynamic scenes. Our originality lies, for a large part, in the larger scale of the dynamic scenes we consider as well as in the time super resolution strategy we investigate. Another particularity of our research is a strong experimental foundation with the multiple camera Kinovis platforms.

3.2. Dynamic Shape Vision

Dynamic Shape Vision refers to research themes that consider the motion of dynamic shapes, with e.g. shapes in different poses, or the deformation between different shapes, with e.g. different human bodies. This includes for instance shape tracking, shape registration, all these themes being covered by MORPHEO. While progress has been made over the last decade in this domain, challenges remain, in particular due to the required essential task of shape correspondence that is still difficult to perform robustly. Strategies in this domain can be roughly classified into two categories: (i) data driven approaches that learn shape spaces and estimate shapes and their variations through space parameterizations; (ii) model based approaches that use more or less constrained prior models on shape evolutions, e.g. locally rigid structures, to recover correspondences. The MORPHEO team is substantially involved in the second category that leaves more flexibility for shapes that can be modeled, an important feature with the Kinovis platform. The team is anyway also considering the first category with faces and body under clothes modeling, classes of shapes that are more likely to evolve in spaces with reasonable dimensions. The originality of MORPHEO in this axis is to go beyond static shape poses and to consider also the dynamics of shape over several frames when modeling moving shapes, this in particular with shape tracking, animation and, more recently, with face registration.

3.3. Inside Shape Vision

Another research axis is concerned with the ability to perceive inside moving shapes. This is a more recent research theme in the MORPHEO team that has gained importance. It was originally the research associated to the Kinovis platform installed in the Grenoble Hospitals. This platform is equipped with two X-ray cameras and ten color cameras, enabling therefore simultaneous vision of inside and outside shapes. We believe this opens a new domain of investigation at the interface between computer vision and medical imaging. Interesting issues in this domain include the links between the outside surface of a shape and its inner parts, especially with the human body. These links are likely to help understanding and modeling human motions. Until now, numerous dynamic shape models, especially in the computer graphic domain, consist of a surface, typically a mesh, bound to a skeletal structure that is never observed in practice but that help anyway parameterizing human motion. Learning more accurate relationships using observations can therefore significantly impact the domain.

3.4. Shape Animation

3D animation is a crucial part of digital media production with numerous applications, in particular in the game and motion picture industry. Recent evolutions in computer animation consider real videos for both the creation and the animation of characters. The advantage of this strategy is twofold: it reduces the creation cost and increases realism by considering only real data. Furthermore, it allows to create new motions, for real characters, by recombining recorded elementary movements. In addition to enable new media contents to be produced, it also allows to automatically extend moving shape datasets with fully controllable new motions. This ability appears to be of great importance with the recent advent of deep learning techniques and the associated need for large learning datasets. In this research direction, we investigate how to create new dynamic scenes using recorded events.

4. Application Domains

4.1. 4D modeling

Modeling shapes that evolve over time, analyzing and interpreting their motion has been a subject of increasing interest of many research communities including the computer vision, the computer graphics and the medical imaging communities. Recent evolutions in acquisition technologies including 3D depth cameras (Time-of-Flight and Kinect), multi-camera systems, marker based motion capture systems, ultrasound and CT scanners have made those communities consider capturing the real scene and their dynamics, create 4D spatio-temporal models, analyze and interpret them. A number of applications including dense motion capture, dynamic shape modeling and animation, temporally consistent 3D reconstruction, motion analysis and interpretation have therefore emerged.

4.2. Shape Analysis

Most existing shape analysis tools are local, in the sense that they give local insight about an object's geometry or purpose. The use of both geometry and motion cues makes it possible to recover more global information, in order to get extensive knowledge about a shape. For instance, motion can help to decompose a 3D model of a character into semantically significant parts, such as legs, arms, torso and head. Possible applications of such high-level shape understanding include accurate feature computation, comparison between models to detect defects or medical pathologies, and the design of new biometric models.

4.3. Human Motion Analysis

The recovery of dense motion information enables the combined analysis of shapes and their motions. Typical examples include the estimation of mean shapes given a set of 3D models or the identification of abnormal deformations of a shape given its typical evolutions. The interest arises in several application domains where temporal surface deformations need to be captured and analyzed. It includes human body analyses for which potential applications are anyway numerous and important, from the identification of pathologies to the design of new prostheses.

4.4. Virtual and Augmented Reality

This domain has actually seen new devices emerged that enable now full 3D visualization, for instance the HTC Vive, the Microsoft Hololens and the Magic leap one. These devices create a need for adapted animated 3D contents that can either be generated or captured. We believe that captured 4D models will gain interest in this context since they provide realistic visual information on moving shapes that tend to avoid negative perception effects such as the uncanny valley effect. Besides 3D visualization devices, many recent applications also rely on everyday devices, such as mobile phones, to display augmented reality contents with free viewpoint ability. In this case, 3D and 4D contents are also expected.

5. Highlights of the Year

5.1. Highlights of the Year

MORPHEO created holograms for an augmented reality application developed for the clothing retailer Zara. This application enables the brand's customers to enjoy a virtual and interactive shopping experience via their smartphones in one of the 120 stores across the world taking part in this experiment. Last January, all of the holograms presented in the Zara AR application were captured using the Kinovis 4D platform. The challenge with regard to the acquisition of the 12 sequences created was to accurately reproduce the models in sweeping movements and with complex clothing effects due to the materials and styles chosen.

6. New Software and Platforms

6.1. 4D repository

KEYWORDS: 4D - Dynamic scene
FUNCTIONAL DESCRIPTION: This website hosts dynamic mesh sequences reconstructed from images captured using a multi-camera set up. Such mesh-sequences offer a new promising vision of virtual reality, by capturing real actors and their interactions. The texture information is trivially mapped to the reconstructed geometry, by back-projecting from the images. These sequences can be seen from arbitrary viewing angles as the user navigates in 4D (3D geometry + time) . Different sequences of human / non-human interaction can be browsed and downloaded from the data section.

- Contact: Edmond Boyer
- URL: http://4drepository.inrialpes.fr/

6.2. Lucy Viewer

KEYWORDS: Data visualization - 4D - Multi-Cameras
SCIENTIFIC DESCRIPTION: Lucy Viewer is an interactive viewing software for 4D models, i.e, dynamic three-dimensional scenes that evolve over time. Each 4D model is a sequence of meshes with associated texture information, in terms of images captured from multiple cameras at each frame. Such data is available from the 4D repository website hosted by Inria Grenoble.

With Lucy Viewer, the user can use the mouse to zoom in onto the 4D models, zoom out, rotate, translate and view from an arbitrary angle as the 4D sequence is being played. The texture information is read from the images at each frame in the sequence and applied onto the meshes. This helps the user visualize the 3D scene in a realistic manner. The user can also freeze the motion at a particular frame and inspect a mesh in detail. Lucy Viewer lets the user to also select a subset of cameras from which to apply texture information onto the meshes. The supported formats are meshes in .OFF format and associated images in .PNG or .JPG format.
FUNCTIONAL DESCRIPTION: Lucy Viewer is an interactive viewing software for 4D models, i.e, dynamic three-dimensional scenes that evolve over time. Each 4D model is a sequence of meshes with associated texture information, in terms of images captured from multiple cameras at each frame.

- Participants: Edmond Boyer, Jean-Sébastien Franco, Matthieu Armando and EYMERIC AMSELEM
- Contact: Edmond Boyer
- URL: https://kinovis.inria.fr/lucyviewer/

6.3. Shape Tracking

FUNCTIONAL DESCRIPTION: We are developing a software suite to track shapes over temporal sequences. The motivation is to provide temporally coherent 4D Models, i.e. 3D models and their evolutions over time , as required by motion related applications such as motion analysis. This software takes as input a temporal sequence of 3D models in addition to a template and estimate the template deformations over the sequence that fit the observed 3D models.

- Contact: Edmond Boyer

6.4. QuickCSG V2

KEYWORDS: 3D modeling - CAD - 3D reconstruction - Geometric algorithms
SCIENTIFIC DESCRIPTION: See the technical report "QuickCSG: Arbitrary and Faster Boolean Combinations of N Solids", Douze, Franco, Raffin.

The extension of the algorithm to self-intersecting meshes is described in "QuickCSG with self-intersections", a document inside the package.
FUNCTIONAL DESCRIPTION: QuickCSG is a library and command-line application that computes Boolean operations between polyhedra. The basic algorithm is described in the research report "QuickCSG: Arbitrary and Faster Boolean Combinations of N Solids", Douze, Franco, Raffin. The input and output polyhedra are defined as indexed meshes. In version 2, that was developed in the context of a software transfer contract, the meshes can be self-intersecting, in which case the inside and outside are defined by the non-zero winding rule. The operation can be any arbitrary Boolean function, including one that is defined as a CSG tree. The focus of QuickCSG is speed. Robustness to degeneracies is obtained by carefully applied random perturbations.

- Authors: Matthys Douze, Jean-Sébastien Franco and Bruno Raffin
- Contact: Jean-Sébastien Franco
- URL: https://kinovis.inria.fr/quickcsg/

6.5. CVTGenerator

KEYWORDS: Mesh - Centroidal Voronoi tessellation - Implicit surface
FUNCTIONAL DESCRIPTION: CVTGenerator is a program to build Centroidal Voronoi Tessellations of any 3D meshes and implicit surfaces.

- Partner: INP Grenoble
- Contact: Li Wang
- URL: http://cvt.gforge.inria.fr/

6.6. Kinovis Platform

Kinovis (http://kinovis.inrialpes.fr/) is a multi-camera acquisition project that was selected within the call for proposals ”Equipements d’Excellence” of the program "Investissement d’Avenir" funded by the French government. The project involves 2 institutes: the Inria Grenoble Rhône-Alpes, the Université Joseph Fourier and 4 laboratories: the LJK (laboratoire Jean Kuntzmann - applied mathematics), the LIG (Laboratoire d’informatique de Grenoble - Computer Science), the Gipsa lab (Signal, Speech and Image processing) and the LADAF (Grenoble Hospitals - Anatomy). The Kinovis environment is composed of 2 complementary platforms (see Figure 2). A first platform located at Inria Grenoble with a 10mx10m acquisition surface is equipped with 68 color cameras and 20 IR motion capture (mocap) cameras. It is the evolution of the Grimage platform towards the production of better models of more complex dynamic scenes. A second platform located at Grenoble Hospitals (CHU), within the LADAF anatomy laboratory, is equipped with 10 color and 2 X-ray cameras to enable combined analysis of internal and external shape structures, such as skeletons and their surrounding bodies. Both platforms have already demonstrated their potential through a range of projects lead by the team and externally. Members of Morpheo are highly involved in this project. Edmond Boyer is coordinating this project, and Julien Pansiot is managing the technical resources of both platforms.

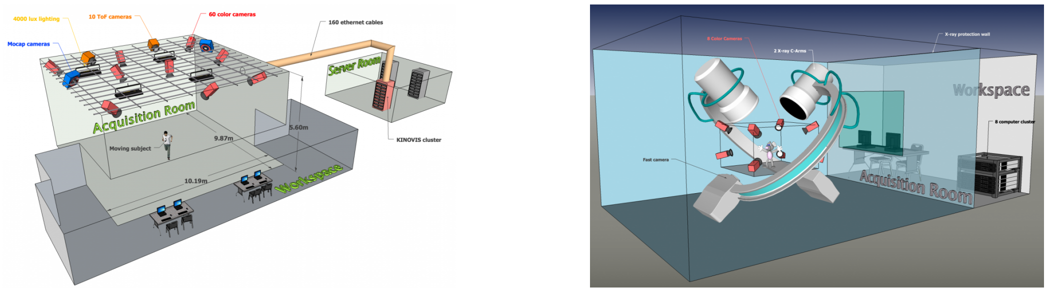


Figure 2. Kinovis platforms: on the left the Inria platform; on the right Grenoble Hospital platform.

7. New Results

7.1. Surface Motion Capture Animation Synthesis

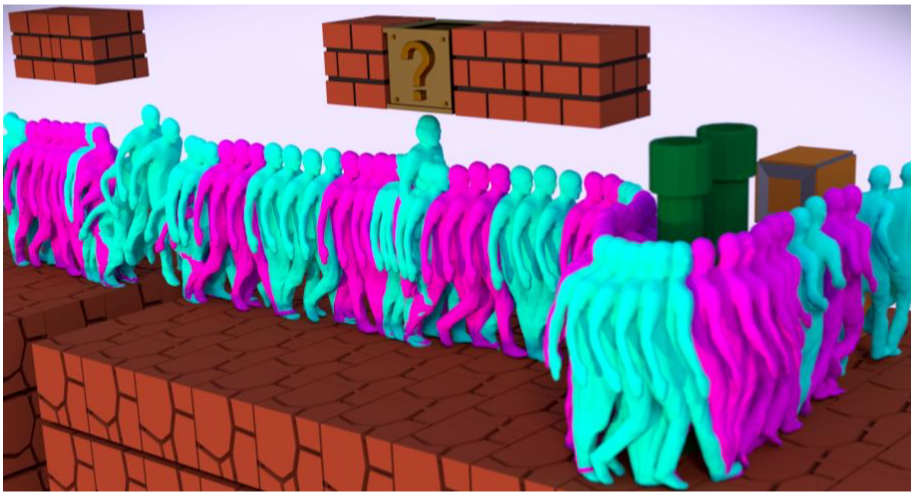


Figure 3.

We propose to generate novel animations from a set of elementary examples of video-based surface motion capture, under user-specified constraints. 4D surface capture animation is motivated by the increasing demand from media production for highly realistic 3D content. To this aim, data driven strategies that consider video-based information can produce animation with real shapes, kinematics and appearances. Our animations rely on the combination and the interpolation of textured 3D mesh data, which requires examining two aspects: (1) Shape geometry and (2) appearance. First, we propose an animation synthesis structure for the shape geometry, the Essential graph, that outperforms standard Motion graphs in optimality with respect to quantitative criteria, and we extend optimized interpolated transition algorithms to mesh data. Second, we propose a compact view-independent representation for the shape appearance. This representation encodes subject appearance changes due to viewpoint and illumination, and due to inaccuracies in geometric modelling independently. Besides providing compact representations, such decompositions allow for additional applications such as interpolation for animation (see figure 3).

This result was published in a prominent computer graphics journal, IEEE Transactions on Visualization and Computer Graphics [2].

7.2. A Multilinear Tongue Model Derived from Speech Related MRI Data of the Human Vocal Tract

We present a multilinear statistical model of the human tongue that captures anatomical and tongue pose related shape variations separately. The model is derived from 3D magnetic resonance imaging data of 11 speakers sustaining speech related vocal tract configurations. To extract model parameters, we use a minimally supervised method based on an image segmentation approach and a template fitting technique. Furthermore, we use image denoising to deal with possibly corrupt data, palate surface information reconstruction to handle palatal tongue contacts, and a bootstrap strategy to refine the obtained shapes. Our evaluation shows that, by limiting the degrees of freedom for the anatomical and speech related variations, to 5 and 4, respectively, we obtain a model that can reliably register unknown data while avoiding overfitting effects. Furthermore, we show that it can be used to generate plausible tongue animation by tracking sparse motion capture data.

This result was published in Computer Speech and Language 51 [3].

7.3. CBCT of a Moving Sample from X-rays and Multiple Videos

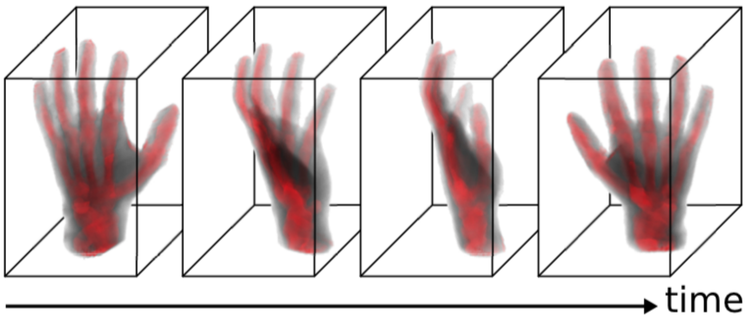


Figure 4. Dense volumetric attenuation reconstruction from a rigidly moving sample captured by a single planar X-ray imaging device and a surface motion capture system. Higher attenuation (here bone structure) is highlighted in red.

We consider dense volumetric modeling of moving samples such as body parts. Most dense modeling methods consider samples observed with a moving X-ray device and cannot easily handle moving samples. We propose instead a novel method to observe shape motion from a fixed X-ray device and to build dense in-depth attenuation information. This yields a low-cost, low-dose 3D imaging solution, taking benefit of equipment widely available in clinical environments. Our first innovation is to combine a video-based surface motion capture system with a single low-cost/low-dose fixed planar X-ray device, in order to retrieve the sample motion and attenuation information with minimal radiation exposure. Our second innovation is to rely on Bayesian inference to solve for a dense attenuation volume given planar radioscopic images of a moving sample. This approach enables multiple sources of noise to be considered and takes advantage of very limited prior information to solve an otherwise ill-posed problem. Results show that the proposed strategy is able to reconstruct dense volumetric attenuation models from a very limited number of radiographic views over time on synthetic and in-situ data, as illustrated in Figure 4.

This result was published in a prominent medical journal, IEEE Transactions on Medical Imaging [4].

7.4. Automatic camera calibration using multiple sets of pairwise correspondences

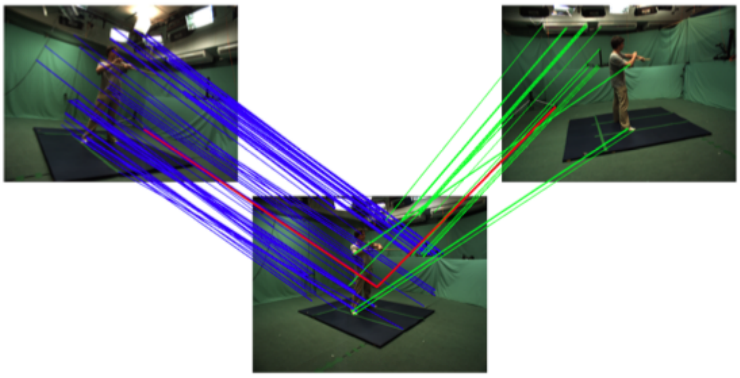


Figure 5. Correspondences extracted from SIFT features. Given the wide baseline between the views there is a single reliable triple correspondence (red) while there are many reliable pairwise correspondences (blue and green).

We propose a new method to add an uncalibrated node into a network of calibrated cameras using only pairwise point correspondences (see figure 5). While previous methods perform this task using triple correspondences, these are often difficult to establish when there is limited overlap between different views. In such challenging cases we must rely on pairwise correspondences and our solution becomes more advantageous. Our method includes an 11-point minimal solution for the intrinsic and extrinsic calibration of a camera from pairwise correspondences with other two calibrated cameras, and a new inlier selection framework that extends the traditional RANSAC family of algorithms to sampling across multiple datasets. Our method is validated on different application scenarios where a lack of triple correspondences might occur: addition of a new node to a camera network; calibration and motion estimation of a moving camera inside a camera network; and addition of views with limited overlap to a Structure-from-Motion model.

This result was published in a prominent medical journal, IEEE Transactions on Pattern Analysis and Machine Intelligence [5].

7.5. Multilinear Autoencoder for 3D Face Model Learning

Generative models have proved to be useful tools to represent 3D human faces and their statistical variations (see figure 6). With the increase of 3D scan databases available for training, a growing challenge lies in the ability to learn generative face models that effectively encode shape variations with respect to desired attributes, such as identity and expression, given datasets that can be diverse. This paper addresses this challenge by proposing a framework that learns a generative 3D face model using an autoencoder architecture, allowing hence for weakly supervised training. The main contribution is to combine a convolutional neural network-based en-coder with a multilinear model-based decoder, taking therefore advantage of both the convolutional network robustness to corrupted and incomplete data, and of the multilinear model capacity to effectively model and decouple shape variations. Given a set of 3D face scans with annotation labels for the desired attributes, e.g. identities and expressions, our method learns an expressive multilinear model that decouples shape changes due to the different factors. Experimental results demonstrate that the proposed

Figure 6. Shape variations caused by different expressions of the same subject.

method outperforms recent approaches when learning multilinear face models from incomplete training data, particularly in terms of space decoupling, and that it is capable of learning from an order of magnitude more data than previous methods.

This result was published in IEEE Winter Conference on Applications of Computer Vision [6].

7.6. Spatiotemporal Modeling for Efficient Registration of Dynamic 3D Faces

We consider the registration of temporal sequences of 3D face scans. Face registration plays a central role in face analysis applications, for instance recognition or transfer tasks, among others. We propose an automatic approach that can register large sets of dynamic face scans without the need for landmarks or highly specialized acquisition setups. This allows for extended versatility among registered face shapes and deformations by enabling to leverage multiple datasets, a fundamental property when e.g. building statistical face models. Our approach is built upon a regression-based static registration method, which is improved by spatiotemporal modeling to exploit redundancies over both space and time. We experimentally demonstrate that accurate registrations can be obtained for varying data robustly and efficiently by applying our method to three standard dynamic face datasets.

This work has been published in 3D Vision 2018 [7].

7.7. Shape Reconstruction Using Volume Sweeping and Learned Photoconsistency

The rise of virtual and augmented reality fuels an increased need for contents suitable to these new technologies including 3D contents obtained from real scenes (see figure 7). We consider in this paper the problem of 3D shape reconstruction from multi-view RGB images. We investigate the ability of learning-based strategies to effectively benefit the reconstruction of arbitrary shapes with improved precision and robustness. We especially target real life performance capture, containing complex surface details that are difficult to recover with existing approaches. A key step in the multi-view reconstruction pipeline lies in the search for matching features between viewpoints in order to infer depth information. We propose to cast the matching on a 3D receptive field along viewing lines and to learn a multi-view photoconsistency measure for that purpose. The intuition is that deep networks have the ability to learn local photometric configurations in a broad way, even with respect to different orientations along various viewing lines of the same surface point. Our results demonstrate this ability, showing that a CNN, trained on a standard static dataset, can help recover surface details on dynamic scenes that are not perceived by traditional 2D feature based methods. Our evaluation also shows that our solution compares on par to state of the art reconstruction pipelines on standard evaluation datasets, while yielding significantly better results and generalization with realistic performance capture data.

This work has been published in the European Conference on Computer Vision 2018 [9] and Reconnaissance des Formes, Image, Apprentissage et Perception 2018 [8].

7.8. FeaStNet: Feature-Steered Graph Convolutions for 3D Shape Analysis

Figure 7. Challenging scene captured with Kinovis. (left) one input image, (center) reconstructions obtained with our previous work based on classical 2D features, (right) proposed solution. Our results validate the key improvement of a CNN-learned disparity to MVS for performance capture scenarios. Results particularly improve in noisy, very low contrast and low textured regions such as the arm, the leg or even the black skirt folds.

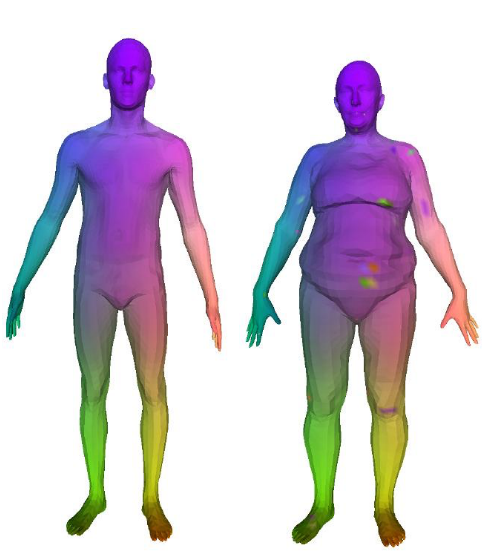


Figure 8. Two examples of texture transfer from a reference shape in neutral pose (left) using shape correspondences predicted by FeaStNet (multi-scale architecture, without refinement).

Convolutional neural networks (CNNs) have massively impacted visual recognition in 2D images, and are now ubiquitous in state-of-the-art approaches. CNNs do not easily extend, however, to data that are not represented by regular grids, such as 3D shape meshes or other graph-structured data, to which traditional local convolution operators do not directly apply. To address this problem, we propose a novel graph-convolution operator to establish correspondences between filter weights and graph neighborhoods with arbitrary connectivity. The key novelty of our approach is that these correspondences are dynamically computed from features learned by the network, rather than relying on predefined static coordinates over the graph as in previous work. We obtain excellent experimental results that significantly improve over previous state-of-the-art shape correspondence results (see figure 8). This shows that our approach can learn effective shape representations from raw input coordinates, without relying on shape descriptors.

This work has been published in the IEEE Conference on Computer Vision and Pattern Recognition 2018 [11].

7.9. Analyzing Clothing Layer Deformation Statistics of 3D Human Motions

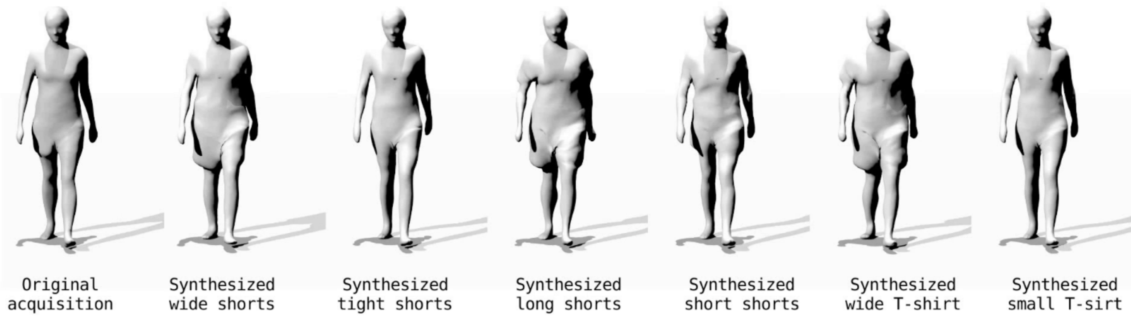


Figure 9. Examples of clothing re-synthesis based on our clothing layer regression model.

Recent capture technologies and methods allow not only to retrieve 3D model sequence of moving people in clothing, but also to separate and extract the underlying body geometry and motion component and separate the clothing as a geometric layer. So far this clothing layer has only been used as raw offsets for individual applications such as retargeting a different body capture sequence with the clothing layer of another sequence, with limited scope, e.g. using identical or similar motions. The structured, semantics and motion-correlated nature of the information contained in this layer has yet to be fully understood and exploited. To this purpose we propose a comprehensive analysis of the statistics of this layer with a simple two-component model, based on PCA subspace reduction of the layer information on one hand, and a generic parameter regression model using neural networks on the other hand, designed to regress from any semantic parameter whose variation is observed in a training set, to the layer parameterization space. We show that this model not only allows to reproduce previous motion retargeting works, but generalizes the data generation capabilities of the method to other semantic parameters such as clothing variation and size (see figure 9), or physical material parameters with synthetically generated training sequence, paving the way for many kinds of capture data-driven creation and augmentation applications.

This work has been published in the European Conference on Computer Vision 2018 [12].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

A collaboration with the French Start up Holooh started in 2017 and was pursued in 2018. Holooh aims at producing high quality holograms for VR and AR applications, especially for the fashion and music domains. Holooh's objective is to set up a multi-camera studio in Paris for that purpose. Edmond Boyer is involved in the collaboration.

8.2. Bilateral Grants with Industry

The Morpheo Inria team and Microsoft research set up a collaboration on the capture and modelling of moving shapes using multiple videos. Two PhD proposals will be part of this collaboration with the objective to make contributions on 4D Modeling. The PhDs will take place at Inria Grenoble Rhône-Alpes and will involve regular visits and stays at Microsoft in Redmond (USA) and Cambridge (UK). At Microsoft, Steve Sullivan, Andrew Fitzgibbon and Marta Wilczkowiak will be participating to the project.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

9.1.1.1. ANR PRCE CaMoPi – Capture and Modelling of the Shod Foot in Motion

The main objective of the CaMoPi project is to capture and model dynamic aspects of the human foot with and without shoes. To this purpose, video and X-ray imagery will be combined to generate novel types of data from which major breakthroughs in foot motion modelling are expected. Given the complexity of the internal foot structure, little is known about the exact motion of its inner structure and the relationship with the shoe. Hence the current state-of-the art shoe conception process still relies largely on ad-hoc know-how. This project aims at better understanding the inner mechanisms of the shod foot in motion in order to rationalise and therefore speed up and improve shoe design in terms of comfort, performance, and cost. This requires the development of capture technologies that do not yet exist in order to provide full dense models of the foot in motion. To reach its goals, the CaMoPi consortium comprises complementary expertise from academic partners : Inria (combined video and X-ray capture and modeling) and Mines St Etienne (finite element modeling), as well as industrial : CTC Lyon (shoe conception and manufacturing, dissemination). The project has effectively started in October 2017 with Claude Goubet's recruitment as a PhD candidate followed by Tomas Svaton as an engineer in April 2018.

9.1.1.2. ANR project Achmov – Accurate Human Modeling in Videos

The technological advancements made over the past decade now allow the acquisition of vast amounts of visual information through the use of image capturing devices like digital cameras or camcorders. A central subject of interest in video are the humans, their motions, actions or expressions, the way they collaborate and communicate. The goal of ACHMOV is to extract detailed representations of multiple interacting humans in real-world environments in an integrated fashion through a synergy between detection, figure-ground segmentation and body part labeling, accurate 3D geometric methods for kinematic and shape modeling, and large-scale statistical learning techniques. By integrating the complementary expertise of two teams (one French, MORPHEO and one Romanian, CLVP), with solid prior track records in the field, there are considerable opportunities to move towards processing complex real world scenes of multiple interacting people, and be able to extract rich semantic representations with high fidelity. This would enable interpretation, recognition and synthesis at unprecedented levels of accuracy and in considerably more realistic setups than currently considered. This project has funded the work of two soon to defend PhD students Vincent Leroy and Jinlong Yang, and ended during the year 2018.

9.1.2. Competitivity Clusters

9.1.2.1. FUI project Creamove

Creamove is a collaboration between the Morpheo team of the Inria Grenoble Rhône-Alpes, the 4D View Solution company specialized in multi-camera acquisition systems, the SIP company specialized in multi-media and interactive applications and a choreographer. The objective is to develop new interactive and artistic applications where humans can interact in 3D with virtual characters built from real videos. Dancer performances will be pre-recorded in 3D and used on-line to design new movement sequences based on inputs coming from human bodies captured in real time. Website: http://www.creamove.fr.

9.1.2.2. FUI24 SPINE-PDCA

The goal of the SPINE-PDCA project is to develop a unique medical platform that will streamline the medical procedure and achieve all the steps of a minimally invasive surgery intervention with great precision through a complete integration of two complementary systems for pre-operative planning (EOS platform from EOS IMAGING) and imaging/intra-operative navigation (SGV3D system from SURGIVISIO). Innovative low-dose tracking and reconstruction algorithms will be developed by Inria, and collaboration with two hospitals (APHP Trousseau and CHU Grenoble) will ensure clinical feasibility. The medical need is particularly strong in the field of spinal deformity surgery which can, in case of incorrect positioning of the implants, result in serious musculoskeletal injury, a high repeat rate (10 to 40% of implants are poorly positioned in spine surgery) and important care costs. In paediatric surgery (e. g. idiopathic scoliosis), the rate of exposure to X-rays is an additional major consideration in choosing the surgical approach to engage. For these interventions, advanced linkage between planning, navigation and postoperative verification is essential to ensure accurate patient assessment, appropriate surgical procedure and outcome consistent with clinical objectives. The project has effectively started in October 2018 with Di Meng's recruitment as a PhD candidate.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. Reviewer

- Jean-Sébastien Franco reviewed for 3DV, CVPR and ECCV 2018, and the CVPR 2018 Humans Workshop.
- Sergi Pujades reviewed for 3DV.
- Julien Pansiot reviewed for ECCV 2018 and CVPR 2019.
- Stefanie Wuhrer reviewed for CVPR, ECCV, and SIGGRAPH.
- Edmond Boyer reviewed for 3DV, CVPR and RFIAP 2018.
- Edmond Boyer was area chair for BMVC and ECCV 2018.

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

- Edmond Boyer is associate editor of the International Journal of Computer Vision (Springer).

10.1.2.2. Reviewer - Reviewing Activities

- Stefanie Wuhrer reviewed for IJCV and PAMI.
- Edmond Boyer reviewed for IJCV and PAMI.
- Sergi Pujades reviewed for the Journal of Imaging (MDPI), Sensors and IEEE TVCG.
- Jean-Sébastien Franco reviewed for IEEE Computer Graphics and Applications.

10.1.3. Invited Talks

- Edmond Boyer gave invited talks at: Microsoft Cambridge (visit), Institut Descarte Paris (conférence maths et mouvement), New York University (visit) and Montpellier University (Module Image).

10.1.4. Scientific Expertise

- Jean-Sebastien Franco reviewed for the Euregio Science Fund (EGTC) - Interregional Project Networks (IPN) in 2018.
- Jean-Sebastien Franco was a member of the recruiting commitee of Universite Grenoble Alpes for an Assistant Professor position in 2018.
- Jean-Sebastien Franco was a member of the Ensimag Engineering school - Grenoble INP steering commitee (Conseil d'École) in 2018.
- Jean-Sebastien Franco was a member of the recruiting commitee of Ensimag - Grenoble INP Engineering school for temporary research and teaching associates (ATER) in 2018.
- Sergi Pujades reviewed for the DFG (German ANR).

10.1.5. Research Administration

- Edmond Boyer is auditor for the Computer Vision European Association.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence: Jean-Sébastien Franco, Introduction to Imperative Programming, 57h, Ensimag 1st year, Grenoble INP.

Licence: Jean-Sébastien Franco, C Programming project, 27h, Ensimag 3rd year, Grenoble INP.

Master: Jean-Sébastien Franco, Supervision of the 2nd year program (300 students), 36h, Ensimag 2nd year, Grenoble INP.

Master: Jean-Sébastien Franco, Introduction to Computer Graphics, 44h, Ensimag 2nd year, Grenoble INP.

Master: Jean-Sébastien Franco, Introduction to Computer Vision, 27h, Ensimag 3rd year, Grenoble INP.

Master: Jean-Sébastien Franco, End of study project tutoring (PFE), 12h, Ensimag 3rd year, Grenoble INP.

Master: Edmond Boyer, 3D Modeling, 18h, M2R Mosig GVR, Grenoble INP.

Master: Edmond Boyer, Introduction to Visual Computing, 30h, M1 MoSig, Université Grenoble Alpes.

Master: Stefanie Wuhrer, 3D Graphics, 13.5h, M1 MoSig and MSIAM, Université Grenoble Alpes.

10.2.2. Supervision

PhD in progress: Victoria Fernandez Abrevaya, 3D Dynamic Human Motion Representations, Université Grenoble Alpes (France), started 01/10/2016, supervised by Edmond Boyer and Stefanie Wuhrer.

PhD in progress: Jean Basset, Learning Morphologically Plausible Pose Transfer, Université Grenoble Alpes (France), started 01/10/2018, supervised by Edmond Boyer, Franck Multon and Stefanie Wuhrer.

PhD in progress: Abdullah Haroon Rasheed, Cloth Modeling and Simulation, Université Grenoble Alpes (France), started 01/11/2017, supervised by Florence Bertails-Descoubes, Jean-Sébastien Franco and Stefanie Wuhrer.

PhD in progress: Jinlong Yang, Learning shape spaces of dressed 3D human models in motion, Université Grenoble Alpes (France), started 01/10/2015, supervised by Franck Hétroy-Wheeler and Stefanie Wuhrer.

HdR : Stefanie Wuhrer, Deformation Models for Human Shape Analysis, Université Grenoble Alpes (France), September 2018

PhD in progress: Claude Goubet, Dense 4D Modelling of Moving Shapes, Université Grenoble Alpes (France), started 02/10/2017, supervised by Edmond Boyer and Julien Pansiot

PhD in progress: Di Meng, Deep learning for low-dose CBCT reconstruction and registration, Université Grenoble Alpes (France), started 01/10/2018, supervised by Edmond Boyer and Julien Pansiot.

PhD in progress: Matthieu Armando, Temporal Integration for Shape and Appearance Modeling, Université Grenoble Alpes (France), started 01/01/2018, supervised by Edmond Boyer and Jean-Sébastien Franco.

PhD in progress: Vincent Leroy, 4D shape reconstruction from photoconsistency cues, Université Grenoble Alpes, started 01/10/2015, supervised by Edmond Boyer and Jean-Sébastien Franco.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] S. WUHRER.*Deformation Models for Human Shape Analysis*, Université Grenoble Alpes (France), September 2018, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01894716

Articles in International Peer-Reviewed Journal

[2] A. BOUKHAYMA, E. BOYER.*Surface Motion Capture Animation Synthesis*, in "IEEE Transactions on Visualization and Computer Graphics", 2018, p. 1-14 [*DOI :* 10.1109/TVCG.2018.2831233], https://hal.inria.fr/hal-01781164

[3] A. HEWER, S. WUHRER, I. STEINER, K. RICHMOND.*A Multilinear Tongue Model Derived from Speech Related MRI Data of the Human Vocal Tract*, in "Computer Speech and Language", September 2018, vol. 51, p. 68-92, https://arxiv.org/abs/1612.05005 [*DOI :* 10.1016/J.CSL.2018.02.001], https://hal.archives-ouvertes.fr/hal-01418460

[4] J. PANSIOT, E. BOYER.*CBCT of a Moving Sample from X-rays and Multiple Videos*, in "IEEE Transactions on Medical Imaging", August 2018, p. 1-11 [*DOI :* 10.1109/TMI.2018.2865228], https://hal.inria.fr/hal-01857487

[5] F. VASCONCELOS, J. P. BARRETO, E. BOYER.*Automatic camera calibration using multiple sets of pairwise correspondences*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", April 2018, vol. 40, n^o 4, p. 791-803 [*DOI :* 10.1109/TPAMI.2017.2699648], https://hal.inria.fr/hal-01675686

International Conferences with Proceedings

[6] V. FERNÁNDEZ ABREVAYA, S. WUHRER, E. BOYER.*Multilinear Autoencoder for 3D Face Model Learning*, in "WACV 2018 - IEEE Winter Conference on Applications of Computer Vision", Lake Tahoe, NV/CA, United States, IEEE, March 2018, p. 1-9 [*DOI :* 10.1109/WACV.2018.00007], https://hal.archives-ouvertes.fr/hal-01700934

[7] V. FERNÁNDEZ ABREVAYA, S. WUHRER, E. BOYER.*Spatiotemporal Modeling for Efficient Registration of Dynamic 3D Faces*, in "3DV 2018 - 6th International Conference on 3D Vision", Verona, Italy, September 2018, p. 1-10, https://hal.inria.fr/hal-01855955

[8] V. LEROY, J.-S. FRANCO, E. BOYER.*Apprentissage de la Cohérence Photométrique pour la Reconstruction de Formes Multi-Vues*, in "RFIAP 2018 - Reconnaissance des Formes, Image, Apprentissage et Perception", Marne la Vallée, France, June 2018, https://hal.archives-ouvertes.fr/hal-01857627

[9] V. LEROY, J.-S. FRANCO, E. BOYER.*Shape Reconstruction Using Volume Sweeping and Learned Photo-consistency*, in "European Conference on Computer Vision", Munich, Germany, September 2018, https://hal.archives-ouvertes.fr/hal-01849286

[10] A.-H. RASHEED, V. ROMERO, F. BERTAILS-DESCOUBES, A. LAZARUS, S. WUHRER, J.-S. FRANCO.*Estimating friction in cloth, using simulation and machine learning*, in "American Physical Society March Meeting 2019", Boston, United States, March 2019, https://hal.inria.fr/hal-01982257

[11] N. VERMA, E. BOYER, J. VERBEEK.*FeaStNet: Feature-Steered Graph Convolutions for 3D Shape Analysis*, in "CVPR - IEEE Conference on Computer Vision & Pattern Recognition", Salt Lake City, United States, June 2018, p. 1-9, https://arxiv.org/abs/1706.05206 , https://hal.inria.fr/hal-01540389

[12] J. YANG, J.-S. FRANCO, F. HÉTROY-WHEELER, S. WUHRER.*Analyzing Clothing Layer Deformation Statistics of 3D Human Motions*, in "ECCV 2018 - European Conference on Computer Vision", Munich, Germany, September 2018, p. 1-17, https://hal.inria.fr/hal-01763706

Team MOSAIC

MOrphogenesis Simulation and Analysis In siliCo

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Computational Biology

Table of contents

Team MOSAIC

Creation of the Team: 2018 January 01

Keywords:

Computer Science and Digital Science:

A3.4. - Machine learning and statistics
A6.1. - Methods in mathematical modeling
A6.2. - Scientific computing, Numerical Analysis & Optimization
A6.3. - Computation-data interaction
A6.5. - Mathematical modeling for physical sciences
A7.1. - Algorithms
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.3. - Geometry, Topology
A8.7. - Graph theory
A9.2. - Machine learning

Other Research Topics and Application Domains:

B1.1.2. - Molecular and cellular biology
B1.1.3. - Developmental biology
B1.1.7. - Bioinformatics
B1.1.8. - Mathematical biology
B1.1.9. - Biomechanics and anatomy
B1.1.10. - Systems and synthetic biology
B1.1.11. - Plant Biology
B3.5. - Agronomy
B9.1.2. - Serious games
B9.5.1. - Computer science
B9.5.2. - Mathematics
B9.5.5. - Mechanics
B9.5.6. - Data science

1. Team, Visitors, External Collaborators

Research Scientists

Olivier Ali [Inria, Researcher]
Romain Azaïs [Inria, Researcher, from Mar 2018]
Christophe Godin [Inria, Senior Researcher, Team leader, HDR]

External Collaborators

Patrick Lemaire [CNRS, Biologist, HDR]
François Parcy [CNRS, Biologist, HDR]
Jan Traas [INRA, Biologist, HDR]
Samuel Teva Vernoux [CNRS, Biologist, HDR]
Arezki Boudaoud [ENS de Lyon, Physicist, HDR]

Frédéric Boudon [CIRAD, Computer scientist]
Ibrahim Cheddadi [INRA, Applied mathematician]
Emmanuel Faure [CNRS, Computer scientist]

Technical Staff

Guillaume Cerutti [INRA, Research Engineer, from Sep 2018]
Jonathan Legrand [CNRS, Research Engineer, 50%]
Florian Gacon [Inria, from Oct 2018]

PhD Students

Katia Mirande [Inria, from Nov 2018]
Hadrien Oliveri [Inria]
Anne Schneider [INRA]

Post-Doctoral Fellow

Bruno Leggio [Inria]

Visiting Scientist

Farah Djelloul [Université de Sidi Bel Abbes, Jul 2018]

Administrative Assistant

Sylvie Boyer [Inria]

2. Overall Objectives

2.1. Overall Objectives

Our general aim in MOSAIC is to identify key principles of organism development in close collaboration with biologists by constructing a new generation of models based on explicit mathematical and computational representations of forms. For this we will develop a dual modeling approach where conceptual models will be used to identify self-organizing principles and realistic models will be used to test non-trivial genetic and physical hypotheses *in silico* and assess them against observations. This will contribute to extend the domain of systems biology to developmental systems and help interpret where possible the vast amount of geometric, molecular and physical data collected on growing forms. The main originality of the project lies in its integrated approach: we want to face the complexity of living organisms by developing an integrated view of form development, relying on the study of the interaction between coupled processes.

While our approach will mainly focus on plant development at different scales, the MOSAIC project will also consider the morphogenesis of model animal systems, such as ascidians [0], to cross-fertilize the approaches and to open the possibility to identify abstractions and principles that are relevant to morphogenesis of living forms in general. Our work will focus on how physical and chemical processes interact within the medium defined by the form and feedback on its development. We will seek to integrate both mechanistic and stochastic components in our models to account for biological variability in shape development. In the long run, the team's results are expected to contribute to set up a new vision of morphogenesis in biology, at the origin of a new physics of living matter, and based on a more mechanistic understanding of the link between genes, forms and their environment.

To achieve the team's objectives, we will develop over the next 12 years a project focused on the definition of a consistent mathematical framework to formalize form growth and on the development of corresponding computational algorithms. The mathematical framework will extend classical dynamical systems to dynamical systems with a dynamical state-structure, i.e. to dynamical systems whose state is represented as a graph of components that may change in time. A similar approach was successfully developed in the last two decades in the restricted context of branching organisms and plant development. We now want to extend it to more general forms, and address the diversity of associated new and stimulating computational challenges. For this, we will organize our research program into three main research axes.

[0] A large class of marine animals (also called sea-squirt) in the phylum of Tunicates that is close to vertebrates, shares a particularly well conserved developmental program and that is a good model to study the development of chordates.

3. Research Program

3.1. Axis1: Representation of biological organisms and their forms in silico

The modeling of organism development requires a formalization of the concept of form, *i.e.* a mathematical definition of what is a form and how it can change in time, together with the development of efficient algorithms to construct corresponding computational representations from observations, to manipulate them and associate local molecular and physical information with them. Our aim is threefold. First, we will develop new computational structures that make it possible to represent complex forms efficiently in space and time. For branching forms, the challenge will be to reduce the computational burden of the current tree-like representations that usually stems from their exponential increase in size during growth. For tissue structures, we will seek to develop models that integrate seamlessly continuous representations of the cell geometry and discrete representations of their adjacency network in dynamical and adaptive framework. Second, we will explore the use of machine learning strategies to set up robust and adaptive strategies to construct form representations in computers from imaging protocols. Finally, we will develop the notion of digital atlases of development, by mapping patterns of molecular (gene activity, hormones concentrations, cell polarity, ...) and physical (stress, mechanical properties, turgidity, ...) expressions observed at different stages of development on models representing average form development and by providing tools to manipulate and explore these digital atlases.

3.2. Axis2: Data-driven models of form development

Our aim in this second research axis will be to develop models of physiological patterning and bio-physical growth to simulate the development of 3D biological forms in a realistic way. Models of key processes participating to different aspects of morphogenesis (signaling, transport, molecular regulation, cell division, etc.) will be developed and tested *in silico* on 3D data structures reconstructed from digitized forms. The way these component-based models scale-up at more abstract levels where forms can be considered as continuums will also be investigated. Altogether, this will lead us to design first highly integrated models of form development, combining models of different processes in one computational structure representing the form, and to analyze how these processes interact in the course of development to build up the form. The simulation results will be assessed by quantitative comparison with actual form development. From a computational point of view, as branching or organ forms are often represented by large and complex data-structures, we aim to develop optimized data structures and algorithms to achieve satisfactory compromises between accuracy and efficiency.

3.3. Axis3: Plasticity and robustness of forms

In this research axis, building on the insights gained from axes 1 and 2 on the mechanisms driving form development, we aim to explore the mechanistic origin of form plasticity and robustness. At the ontogenetic scale, we will study the ability of specific developmental mechanisms to buffer, or even to exploit, biological noise during morphogenesis. For plants, we will develop models capturing morphogenetic reactions to specific environmental changes (such as water stress or pruning), and their ability to modulate or even to reallocate growth in an opportunistic manner.

At the phylogenetic scale, we will investigate new connections that can be drawn from the use of a better understanding of form development mechanisms in the evolution of forms. In animals, we will use ascidians as a model organism to investigate how the variability of certain genomes relates to the variability of their forms. In plants, models of the genetic regulation of form development will be used to test hypotheses on the evolution of regulatory gene networks of key morphogenetic mechanisms such as branching. We believe that a better mechanistic understanding of developmental processes should shed new light on old evo-devo questions related to the evolution of biological forms, such as understanding the origin of *developmental constraints*[0] how the internal rules that govern form development, such as chemical interactions and physical constraints, may channel form changes so that selection is limited in the phenotype it can achieve?

[0] Raff, R. A. (1996). The Shape of Life: Genes, Development, and the Evolution of Form. Univ. Chicago Press.

3.4. Key modeling challenges

During the project lifetime, we will address several computational challenges related to the modeling of living forms and transversal to our main research axes. During the first phase of the project, we concentrate on 4 key challenges.

3.4.1. A new paradigm for modeling tree structures in biology

There is an ubiquitous presence of tree data in biology: plant structures, tree-like organs in animals (lungs, kidney vasculature), corals, sponges, but also phylogenic trees, cell lineage trees, *etc*. To represent, analyze and simulate these data, a huge variety of algorithms have been developed. For a majority, their computational time and space complexity is proportional to the size of the trees. In dealing with massive amounts of data, like trees in a plant orchard or cell lineages in tissues containing several thousands of cells, this level of complexity is often intractable. Here, our idea is to make use of a new class of tree structures, that can be efficiently compressed and that can be used to approximate any tree, to cut-down the complexity of usual algorithms on trees.

3.4.2. Efficient computational mechanical models of growing tissues

The ability to simulate efficiently physical forces that drive form development and their consequences in biological tissues is a critical issue of the MOSAIC project. Our aim is thus to design efficient algorithms to compute mechanical stresses within data-structures representing forms as the growth simulation proceeds. The challenge consists of computing the distribution of stresses and corresponding tissue deformations throughout data-structures containing thousands of 3D cells in close to interactive time. For this we will develop new strategies to simulate mechanics based on approaches originally developed in computer graphics to simulate in real time the deformation of natural objects. In particular, we will study how meshless and isogeometric variational methods can be adapted to the simulation of a population of growing and dividing cells.

3.4.3. Realistic integrated digital models

Most of the models developed in MOSAIC correspond to specific parts of real morphogenetic systems, avoiding the overwhelming complexity of real systems. However, as these models will be developed on computational structures representing the detailed geometry of an organ or an organism, it will be possible to assemble several of these sub-models within one single model, to figure out missing components, and to test potential interactions between the model sub-components as the form develops.

Throughout the project, we will thus develop two digital models, one plant and one animal, aimed at integrating various aspects of form development in a single simulation system. The development of these digital models will be made using an agile development strategy, in which the models are created and get functional at a very early stage, and become subsequently refined progressively.

3.4.4. Development of a computational environment for the simulation of biological form development

To support and integrate the software components of the team, we aim to develop a computational environment dedicated to the interactive simulation of biological form development. This environment will be built to support the paradigm of dynamical systems with dynamical structures. In brief, the form is represented at any time by a central data-structure that contains any topological, geometric, genetic and physiological information. The computational environment will provide in a user-friendly manner tools to up-load forms, to create them, to program their development, to analyze, visualize them and interact with them in 3D+time.

4. Highlights of the Year

4.1. Highlights of the Year

The year 2018 was marked by the following events:

- **Creation of the team.** The team MOSAIC started in January 2018 at the Inria Grenoble Rhône-Alpes Research Center and is part of the laboratoire de reproduction des plantes (RDP research unit) at ENS de Lyon campus. Romain Azaïs joint the team in March 2018 and Guillaume Cerutti was hired as an Inra research engineer in September 2018.
- **Edition of *Statistical Inference for Piecewise-deterministic Markov Processes*.** Piecewise-deterministic Markov processes form a class of stochastic models with a sizeable scope of applications. Such processes are defined by a deterministic motion punctuated by random jumps at random times, and offer simple yet challenging models to study. The issue of statistical estimation of the parameters ruling the jump mechanism is far from trivial. Responding to new developments in the field as well as to current research interests and needs, the book “Statistical Inference for Piecewise-deterministic Markov Processes” edited by Romain Azaïs and Florian Bouguet [10] gather 7 chapters by different authors on the topic. The idea for this book stemmed from a workshop organized in Nancy in the 2016-17 winter.
- **Invited talk at the Jacques Monod conference in Roscoff**. Christophe Godin was invited in Sep 2018 at the prestigious Jacques Monod series of international conferences in Roscoff, France, to present an overview of the current research on phyllotaxis. The talk was entitled *Phyllotaxis at the era of molecular and computational biology: the revival of an old enigma* and prepared with Teva Vernoux.
- **First prototype of the software platform Gnomon**. A first, fully functional, prototype of the Gnomon software platform, dedicated to the modeling and simulation of plant and animal morphogenesis, was developed during a series of intensive coding sessions in Lyon and Sophia-Antipolis. This new concept of platform dedicated to the study of morphogenesis was presented in November 2018 to a panel of modelers and biologists at the RDP lab, who will contribute next year to the further testing and refining the platform. This prototype is a clear milestone and results from a strong collaboration between the Inria software engineering group from Sophia-Antipolis (who provides the software architecture kernel - DTK)) and the Mosaic team and is supported by Inria (Action de Developpement Tecnologique, ADT).

5. New Software and Platforms

5.1. treex

KEYWORDS: Graph algorithmics - Data structures - Combinatorics
SCIENTIFIC DESCRIPTION: Trees form an expanded family of combinatorial objects that offers a wide range of application fields, especially in biology, from plant modeling to blood vessels network analysis through study of lineages. Consequently, it is crucial for the team to develop numerical tools and algorithms for processing tree data, in particular to answer questions about the representation of biological organisms and their forms in silico.

treex is a Python 3 library dedicated to the manipulation of tree objects, whatever they are ordered or not, with or without quantitative or qualitative labels.
FUNCTIONAL DESCRIPTION: treex is a Python library for manipulating rooted trees. The trees can be ordered or not, with or without labels on their vertices.

The package provides a data structure for rooted trees as well as the following main functionalities: - Random generation algorithms - DAG compression for ordered or not, labeled or not, trees - Approximation algorithms for unordered trees - Edit distance for unordered labeled trees - Computation of coding processes (Harris path, Lukasiewicz walk and height process) - Visualization algorithms in Matplotlib or in LaTeX

Representations of trees. With treex, we aim to propose all the standard representations of trees as well as the one-to-one correspondences between them. Main coding processes (Harris path, Lukasiewicz walk, and height process), DAG representation, doubly-chained tree structures have been already coded. Standard exploration algorithms and editing methods have also been developed. Through these generic tools, treex enables the manipulation of trees from various application contexts.

Easy-to-use. We think that treex must be user-friendly to be adopted by collaborators from fields of biology. To this end, we develop high-level algorithms and provide an extensive documentation (with sphinx) as well as a simple installation method through conda.

Algorithms. The current version of treex provides edit distance algorithms, approximation algorithms (that can be used to control the complexity of the edit distance algorithms), and visualization algorithms (with an interface for TEX / LATEX and Matplotlib). A first statistical learning module is in progress.

RELEASE FUNCTIONAL DESCRIPTION: The first release of treex happened in the late 2018 after an intensive work to ease both installation and handling. A publication on treex is planned for the next early year. In addition, treex will be integrated in Gnomon in the following months.

NEWS OF THE YEAR: The first release of treex happened in the late 2018 after an intensive work to ease both installation and handling. A publication on treex is planned for the next early year. In addition, treex will be integrated in Gnomon in the following months.

- Participants: Romain Azais, Guillaume Cerutti, Didier Gemmerle and FLORIAN INGELS
- Contact: Romain Azais
- URL: https://gitlab.inria.fr/azais/treex

5.2. Gnomon

KEYWORDS: 4D - Modelization and numerical simulations - Finite element modelling - Computational biology - Data visualization

SCIENTIFIC DESCRIPTION: Gnomon is a user-friendly computer platform developed by the Mosaic team for seamless simulation of form development in silico. It is intended to be a major tool for the team members to develop, integrate and share their models, algorithms and tools. In Gnomon, a developing form is represented at any time by a central data-structure that contains topological, geometric, genetic and physiological information and that represents the state of the growing form. Flexible components (plugins) make it possible to up-load or to create such data-structures, to program their development, to analyze, visualize them and interact with them in 3D+time.

FUNCTIONAL DESCRIPTION: Gnomon is developed based on the past experience of the team with the OpenAlea platform, but moving towards a more scalable software engineering solution based on the dtk kernel developed by the group of software engineers (SED) from the Sophia-Antipolis Inria Center. Gnomon development uses extensively an agile methodology and emphasizes three main aspects:

Deployable and extensible software architecture. The Gnomon platform is intended to become a perennial common resource for the members of the team as well as a tool to easily diffuse our methods to collaborators. It is a plugin-based architecture, relying on the dtk meta-platform developed by the SED at Inria Sophia-Antipolis. dtk eases the conception of a high-level C++ environment inside which added-value components are injected autonomously by the team members as Python plugins wrapping our scientific libraries. A significant effort is put on the packaging and deployability of this software (using conda), adding up unit testing, continuous integration and cross-platform installation.

Exploration of forms. The environment will provide tools to create and visualize forms, and explore them in space and time. Building on the algorithmic resources developed by the team for image sequences of multicellular tissues, user-friendly interfaces are being designed for the exploration of such structures. This gives the user the possibility to reconstruct computational representations from experimental data in an intuitive way, and to explore these spatio-temporal data in an interactive and visual manner.

Integrated form simulation framework. Within a general framework for the modeling of dynamical systems that the team is developing, a core component is the mechanical simulation engine that will handle the resolution of physical equations controlling form development. We want the simulation framework to be integrated within the Gnomon platform in a nearly transparent way for the user. To achieve this, we develop a high-level interface for a generic differential equation solver based on the *fenics* FEM library. Mapping the general concepts of morphogenesis modeling to this engine will allow the user to specify behavior rules of the system at high-level and easily design simulation scenarios directly in the Gnomon application.

Gnomon project organization:

- Project leader: Christophe Godin
- Software development coordinator: Guillaume Cerutti
- DTK backend coordinator: Thibaud Kloczko
- Plugin coordinators: Jonathan Legrand (TimageTK), treex (Romain Azais), Olivier Ali (Mechanics), Frédéric Boudon (L-Systems).
- Diffusion to end-users: Teva Vernoux

This work is part of the *Gnomon* ADT project supported by the Inria centers of Grenoble Rhône-Alpes and Sophia-Antipolis Méditerranée.

RELEASE FUNCTIONAL DESCRIPTION: A first, fully functional, prototype of the Gnomon software platform, dedicated to the modeling and simulation of plant and animal morphogenesis, was developed during a series of intensive coding sessions in Lyon and Sophia-Antipolis. This new concept of platform dedicated to the study of morphogenesis was presented in November 2018 to a panel of modelers and biologists at the RDP lab, who will contribute next year to the further testing and refining the platform. This prototype is a clear milestone and results from a strong collaboration between the Inria software engineering group from Sophia-Antipolis (who provides the software architecture kernel - DTK)) and the Mosaic team and is supported by Inria (Action de Developpement Tecnologique, ADT).

NEWS OF THE YEAR: A first, fully functional, prototype of the Gnomon software platform, dedicated to the modeling and simulation of plant and animal morphogenesis, was developed during a series of intensive coding sessions in Lyon and Sophia-Antipolis. This new concept of platform dedicated to the study of morphogenesis was presented in November 2018 to a panel of modelers and biologists at the RDP lab, who will contribute next year to the further testing and refining the platform. This prototype is a clear milestone and results from a strong collaboration between the Inria software engineering group from Sophia-Antipolis (who provides the software architecture kernel - DTK)) and the Mosaic team and is supported by Inria (Action de Developpement Tecnologique, ADT).

- Participants: Olivier Ali, Frédéric Boudon, Guillaume Cerutti, FLORIAN GACON, Christophe Godin, Jonathan Legrand and Grégoire Malandain
- Contact: Christophe Godin

5.3. TimageTK: a Python package for image processing of multicellular architectures

Participants: Frédéric Boudon [External Collaborator], Guillaume Cerutti, Christophe Godin, Jonathan Legrand, Grégoire Malandain [External Collaborator].

- Related Research Works: **RA1** (*Representations of forms in silico*) & **RA2** (*Data-driven models*) & **RA3** (*Plasticity and robustness of forms*)
- Related Key Modeling Challenges: **KMC3** (*Realistic integrated digital models*)

Processing images of multicellular tissue architectures in plants and animals present difficult computational challenges, notably when dealing with heterogeneous data sources, temporal data. As for now, only limited computational tools exit to analyze these types of images efficiently. Based on our initial experience with the development of MARS-ALT, a pipeline for segmenting and tracking cell lineages, we have recently redesigned our software in order to develop a new high-level Python package named Tissue image ToolKit (TimageTK).

C/Python library. TimageTK is written in Python and is largely built on top of a C library (VT) developed by the Morpheme Inria team. Part of the C library was developed for MARS-ALT software by the Morpheme team. TimageTK provides high-level wrapping of these algorithms with additional functionalities directly written in Python, such as cell tracking.

Well documented, high-level image processing. TimageTK offer high-level methods, where few parameters are required to tune algotithms. They offer a safer use of function through well documented class methods. For experts in image processing, low-level functions and wrappings of C-functions are still accessible.

Easy deployment. TimageTK has simple and robust installation procedure based on conda (Conda is a package and environment management system that runs on Windows, macOS and Linux and allows to get rid of most installation issues, such as compilation or dependency errors). For now only macOSX and Linux version are packaged (x64 architecture) using Inria continuous integration tools.

Continuous integration. Using Inria continuous integration tools, TimageTK is regularly released through conda packaging mechanism. We also make use of these resources to regularly generate (sphinx) and publish updates in the documentation.

Compatibility with Gnomon. TimageTK is based on data structures representing images that are fully compatible with Gnomon. This makes it possible to use TimageTK as a plugin of the Gnomon software.

6. New Results

6.1. Dynamical characterization of morphogenesis at cellular scale

Participants: Guillaume Cerutti, Emmanuel Faure [External Collaborator], Christophe Godin, Bruno Leggio, Jonathan Legrand, Patrick Lemaire [External Collaborator], Grégoire Malandain [External Collaborator], Jan Traas [External Collaborator].

- Research Axes: **RA1** (*Representation of biological organisms and their forms in silico*) & **RA3** (*Plasticity & robustness of forms*)
- Key Modeling Challenges: **KMC3** (*Realistic integrated digital models*)

The modeling of morphogenesis requires to explore the interconnection of different spatial and temporal scales of developing organisms. Non-trivial questions such as whether the observed robustness of morphogenesis is rooted in some highly conserved properties at the cellular level or whether it emerges as a macroscopic phenomenon, necessitate precise, quantitative analyses of complex 3D dynamic structures. The study of dynamical properties at the cellular scale poses at the same time key technical challenges and fundamental theoretical questions. An example of the former category is how to characterize and follow the change of shape of cells within tissues and of tissues within organs, and how to couple this change with, for instance, gene expression dynamics; an illustration of the latter is how to define cell-scale variability of morphogenesis within and between species. Our team has produced this year several results in this context:

Cells spatio-temporal properties and patterns characterization. Over the past few years, we have achieved quantitative characterization of some of the cells physical properties, such as volumes or curvatures, in a developing tissue. Together with cell lineaging, it also enabled the quantification of temporal properties at cellular scale such as volumetric growth rate or strain patterns. To ease-up the analysis and to structure the previously described data, we have implemented a dedicated spatio-temporal graph structure, formalizing the cell network and its change in time.

To further characterize the tissue development, we developed clustering methods to identify cellular patterns based on a selection of quantified cell properties, including topology. Since such data are highly structured, both in time and space, we developed two complementary approaches:

1. spatial oriented: this approach use the cell neighborhood and a selection of cell descriptors to create pairwise distance maps latter clustered by a distance-based method, such as Ward's hierarchical clustering.
2. temporal oriented: this approach uses the lineage forest and a selection of cell descriptors to infer cell identities using Hidden Markov Tree (HMT) models.

Both approaches allow later characterization of the detected cluster or groups of cells based on their properties and should be published during the first half of 2019.

Atlases. One fundamental requirement to understand morphogenesis is the creation of atlases of different properties and different species. This year we have started creating two morphogenetic atlases: the atlas of gene expression patterns in the *Arabidopsis thaliana* flower development and the atlas of early embryonic development of the ascidian *Phallusia mammillata*.

Phallusia mammillata embryos develop with an invariant cellular lineage and with a relatively low number of cells ($\sim$ 700) up until the end of neurulation. This allows the creation of atlases with cellular resolution. Developing embryos from in-vitro fertilised dechorionated eggs have been injected with mRNA to fluorescently label their cell membranes and imaged by light-sheet microscopy for several hours of development. Automated image reconstruction through the segmentation pipeline ASTEC allowed to collect a large number of wild-type and mutated development with single-cell resolution and with a temporal resolution of two minutes. Based on this amount of data and on the invariant early ascidian lineage, we started curating an atlas of wild-type cellular, tissue and embryonic properties. Each cell, classified by its unique name, is identified in each wild-type embryo and analyzed through the dedicated computational pipelines. The result of this work provides a comprehensive view on the variability (in time, within and between embryos) of properties such as cell volume, cell surface, cell and tissue shape, cell topology, length of cell cycle, cell position within its tissue and globally within the embryo, orientation of cell's cleavage plane. This cellular networks have been coupled via cell names with genetic data coming from the the ascidian genetic database (ANISEED) and a specific tool, Morphonet, has been developed to explore these morphodynamic atlases seamlessly within a web-browser (paper in revision).

On the other hand, developing digital atlases of organism or organs development is a complex challenge for organisms presenting a strong variability in the cellular layout. Indeed contrary to *C. Elegans* or *P. mammillata*, for instance, that posses a very strict cell lineage, the development of most organisms or organs is under the influence of robust genetic patterns but without a unique cellular layout. In that respect, proposing a cell-based atlas of flower development for instance is not straightforward and specific methods have been developed to choose a representative examples of the developing *A. thaliana* flower. Using this representative flower we have generated an atlas in which we have introduced manually the expression patterns of 27 genes. The knowledge generated by the creation of this atlas makes it possible to have a first quantitative (correlative) view on the relation between gene activity and growth.

Both these works should result in publications in 2019.

Robustness of ascidian embryonic development. The image segmentation pipeline ASTEC developed by the team allows the 3D dynamic reconstruction of early ascidian embryogenesis at cellular resolution. Based on the high-quality wild-type data of our ascidian morphogenetic atlas and on ANISEED, we investigated the robustness of ascidian embryonic development and established a model to explain its origin. Thanks to the image-analysis pipelines we developed, we could extract relevant information from data and to perform cell-to-cell comparisons between different embryos of the same species (Phallusia mammillata). Since embryos developing from dechorionated eggs are left-right symmetric, we assessed the degree of cell-level variability between two embryos with different genomes (genetic variability) by comparing it to the intrinsic left-right variability in cellular properties within each embryo (stochastic variability). We showed that the same degree of variability is observed within and between embryos, demonstrating how ascidian embryonic development is highly canalised, and that the high reproducibility of shapes observed during embryogenesis is rooted in the robustness of cellular geometry and topology. Based on these observations, we studied the dynamics of embryonic patterning by developing a quantitative mathematical model for cellular fate-restriction events based on kinetic equations describing biochemical signalling. This model suggests that the robustness of cell topology and geometry is necessary for cell-cell biochemical interactions to give rise to the correct fate restriction events, a phenomenon which might represent a strong evolutive constraint to cell-scale variability in ascidians.

These results gave rise to a work which is currently under review and published as a preprint [16].

Digital reconstruction of developing *Arabidopsis* ovule. The ovule is a relatively simple organ, with limited developmental variability, which makes it an excellent case study for the computational modeling of organ development. In order to test various hypotheses of cellular growth, we reconstructed a first 4D digital tissue structure of a developing ovule as a triangulated cellular complex. It can be used as an input for FEM-based simulations, and will allow to compare quantitatively the results of growth models with actual ovule development.

This work was part of the *Imago* project.

6.2. Reconstruction of macroscopic forms from images and characterization of their variability

Participants: Guillaume Cerutti, Christophe Godin, Jonathan Legrand, Katia Mirande.

- Research Axes: **RA1** (*Representations of forms in silico*) & **RA3** (*Plasticity & robustness of forms*)
- Key Modeling Challenges: **KMC3** (*Realistic integrated digital models*)

To study the variability of macroscopic forms resulting from development, it is necessary to both develop digital reconstruction methods, typically based on image acquisitions, and statistical tools to define notions of distance or average between these forms. The automatic inference of computational representations of forms or organ traits from images of different types is therefore an essential step, for which the use of prior knowledge can be very beneficial. Realistic synthetic models of forms can guide the reconstruction algorithms and/or assess their performances. Computational representations of forms can then be used to analyze how forms vary at the scale of a population, of a species or between species, with potential applications in species identification and genetic or environmental robustness estimation.

Automatized characterization of 3D plant architecture. The digital reconstruction of branching and organ forms and the quantification of phenotypic traits (lengths of internodes, angles between organs, leaf shapes) is of great interest for the analysis of plant morphology at population scale. We develop an automated processing pipeline that involves the 3D reconstruction of plant architecture from RGB image acquisitions performed by a robot, and the segmentation of the reconstructed plant into organs. To provide validation data for the pipeline, we designed a generative model of *Arabidopsis thaliana* simulating the development of the plant architecture at organ scale. This model was used to develop the method for the measurement of angles of organs and test its accuracy. In a second phase, the model will be used to generate training data for machine learning techniques introduced in the reconstruction methods.

This work is part of the *ROMI* project.

Identification of plant species from morphological traits. The description of morphological traits of the various organs of the plants (leaves, bark, flowers and fruits) is essential for the characterization of a phenotype, and is highly relevant in the context of species or variety identification. In the context of tree species identification from RGB images of their organs, we study methods to represent the morphological charateristics of the plant organs, and the way to combine those different sources of information to enhance the classification performance. We demonstrated that botany-inspired descriptors of bark improves tree species classification based on leaves [13]. We also explore the possibility of using deep learning techniques to train a system to extract botanically relevant information from images [3].

This work is part of the *ReVeRIES* ANR project, in which the team is not directly involved.

This work has led to a publication in *Ecological Informatics* and to a participation at the *International Worskhop on Image Analysis Methods for the Plant Sciences* in Nottingham in January 2018.

6.3. Analysis of tree data

Participants: Romain Azaïs, Christophe Godin, Florian Ingels, Clément Legrand.

- Related Research Axes: **RW1** (*Representations of forms in silico*)
- Related Key Modeling Challenges: **KMC1** (*A new paradigm for modeling tree structures in biology*)

Tree-structured data naturally appear at different scales and in various fields of biology where plants and blood vessels may be described by trees. In the team, we aim to investigate *a new paradigm for modeling tree structures in biology* in particular to solve complex problems related to the *representation of biological organisms and their forms in silico*.

In 2018, we investigated the following questions linked to the analysis of tree data. (i) How to control the complexity of the algorithms used to solve queries on tree structures? For example, computing the edit distance matrix of a dataset of large trees is numerically expensive. (ii) How to estimate the parameters within a stochastic model of trees? And finally, (iii) how to develop statistical learning algorithms adapted to tree data? In general, trees do not admit a Euclidean representation, while most of classification algorithms are only adapted to Euclidean data. Consequently, we need to study methods that are specific to tree data.

Approximation of trees by self-nested trees. Complex queries on tree structures (*e.g.*, computation of edit distance, finding common substructures, compression) are required to handle tree objects. A critical question is to control the complexity of the algorithms implemented to solve these queries. One way to address this issue is to approximate the original trees by simplified structures that achieve good algorithmic properties. One can expect good algorithmic properties from structures that present a high level of redundancy in their substructures. Indeed, one can take account these repetitions to avoid redundant computations on the whole structure. In the team, we think that the class of self-nested trees, that are the most compressed trees by DAG compression scheme, is a good candidate to be such an approximation class.

In [7], we have proved the algorithmic efficiency of self-nested trees through different questions (compression, evaluation of recursive functions, evaluation of edit distance) and studied their combinatorics. In particular, we have established that self-nested trees are roughly exponentially less frequent than general trees. This combinatorics can be an asset in exhaustive search problems. Nevertheless, this result also says that one can not always take advantage of the remarkable algorithmic properties of self-nested trees when working with general trees. Consequently, our aim is to investigate how general trees can be approximated by simplified trees in the class of self-nested trees from both theoretical and numerical perspectives.

We conjecture that the problem of optimal approximation by a self-nested tree is NP-hard. Despite a substantial work in 2018 (internship of Clément Legrand), this remains an open question. Consequently, we have developed a suboptimal approximation algorithm based on the *height profile* of a tree that can be used to very rapidly predict the edit distance between two trees, which is a usual but costly operation for comparing tree data in computational biology [7]. Another algorithm based on the simulation of Gibbs measures on the space of trees is currently under development. This work should result in a publication next year.

Statistical inference. The main objective of statistical inference is to retrieve the unknown parameters of a stochastic model from observations. A Galton-Watson tree is the genealogical tree of a population starting from one initial ancestor in which each individual gives birth to a random number of children according to the same probability distribution, independently of each other. In a recent work [12], we have focused on Galton-Watson trees conditional on their number of nodes. Several main classes of random trees can be seen as conditioned Galton-Watson trees. For instance, an ordered tree picked uniformly at random in the set of all ordered trees of a given size is a conditioned Galton-Watson tree with offspring distribution the geometric law with parameter $1/2$. Statistical methods were developed for conditioned Galton-Watson trees in [19]. We have introduced new estimators and stated their consistency. Our techniques improve the existing results both theoretically and numerically. A simulation study shows the good behavior of our procedure on finite-sample sizes and from missing or noisy data.

In a very different context, a substantial work has been made on statistical inference for piecewise-deterministic processes [2], [9], [8].

Kernel methods for tree data. In statistical learning, one aims to build a decision rule of a qualitative variable Y as a function of a feature X (typically a vector of $\mathbb{R}^d$) from a training dataset $(X_i, Y_i)_{1 \leq i \leq n}$. We assume that X is a tree, ordered or not, with or without labels. This framework is quite original since the state space of X is not endowed with a canonical inner product. Kernel methods are particularly adapted to this setting since they enable to transform the raw data into a Hilbert space. In this context, the main issue is related to the construction of a *good kernel*. A kernel function adapted to trees is the subtree kernel introduced [24]. While the literature has never been focused on the weight function involved in the subtree kernel, we have shown that this function is crucial in prediction problems. We have proposed a new algorithm for computing the subtree kernel. It has been designed to allow learning the weight function directly from the data. On some difficult datasets, the prediction error is dramatically decreased from $> 50\%$ to 3%.

This work is part of the *ROMI* project, that aims to develop an open and lightweight robotics platform for microfarms. This project requires to investigate advanced analysis and modeling techniques for plant structures. A main issue that arises in this context is to predict a feature of the plant (species, health status, etc) from its topology.

Invited talk on tree structures and algorithms Christophe Godin gave a invited talk entitled *Can we manipulate tree forms like numbers?* that was prepared with Romain Azaïs at the workshop on *Mathematics for Developmental Biology* organized at the Banff International Research Station for Mathematical Innovation and Discovery, organized by P. Prusinkiewicz and E. Mjolsness (Banff, Canada, December 2017).

Abstract: Tree-forms are ubiquitous in nature and recent observation technologies make it increasingly easy to capture their details, as well as the dynamics of their development, in 3 dimensions with unprecedented accuracy. These massive and complex structural data raise new conceptual and computational issues related to their analysis and to the quantification of their variability. Mathematical and computational techniques that usually successfully apply to traditional scalar or vectorial datasets fail to apply to such structural objects: How to define the average form of a set of tree-forms? How to compare and classify tree-forms? Can we solve efficiently optimization problems in tree-form spaces? How to approximate tree-forms? Can their intrinsic exponential computational curse be circumvented? In this talk, we presented a recent work to approach these questions from a new perspective, in which tree-forms show properties similar to that of numbers or real functions: they can be decomposed, approximated, averaged, and transformed in dual spaces where specific computations can be carried out more efficiently. We will discuss how these first results can be applied to the analysis and simulation of tree-forms in developmental biology (https://www.birs.ca/events/2017/5-day-workshops/17w5164).

6.4. Mechanics of tissue morphogenesis

Participants: Olivier Ali, Arezki Boudaoud [External Collaborator], Guillaume Cerutti, Ibrahim Cheddadi [External Collaborator], Christophe Godin, Bruno Leggio, Jonathan Legrand, Hadrien Oliveri, Jan Traas [External Collaborator].

- Research Axes: **RA2** (*Data-driven models*) & **RA3** (*Plasticity & robustness of forms*)
- Key Modeling Challenges: **KMC2** (*Efficient computational mechanical models of growing tissues*) & **KMC3** (*Realistic integrated digital models*)

As deformations supporting morphogenesis require the production of mechanical work within tissues, the ability to simulate accurately the mechanical behavior of growing living tissues is a critical issue of the MOSAIC project. From a macroscopic perspective, tissues mechanics can be formalized through the framework of continuum mechanics. However, the fact that they are composed, at the microscopic level, by active building blocks out of equilibrium (namely cells) offers genuine modeling challenges and opportunities. This section describes the team's efforts on integrating cellular behaviors such as mechano-sensitivity, intercellular fluxes of materials and cell division into a macroscopic mechanical picture of morphogenesis.

Mechanical influence of inner tissues. Mechanical stress patterns within plant tissues emerge from the balance between inner-pressure-induced forces and the elastic response of the cell wall [0] over the entire tissue. Being able to derive, from a specific cellular architecture, the corresponding pattern of stresses within a tissue is crucial for the study of morphogenesis. It requires a precise description of the tissue as a network of connected cells and the ability to run numerical simulations of force balance on such heterogeneous structures.

[0] A thick protective exoskeleton surrounding plant cells

To that end, we developed numerical methods to generate finite element meshes from: i) 3D microscopic images with sub-cellular resolution (referred to as *bio-inspired* structures) and ii) 3D cellularized geometrical volumes (referred to as *artificial* structures). Combined with a FEM-based simulation framework previously developed within the team [20], we generated quantitative maps of stress distributions in multilayered reconstructed tissues. The combined analysis of stress patterns on *bio-inspired* and *artificial* structures showed how mechanical stresses experienced by cells convey geometrical information to cells about the global shape of the tissue as well as the local shape of cells.

This work was part of the *Morphogenetics* IPL and Jan Traas ERC grant *Morphodynamics*.

This work is currently under review in the *Bulletin of Mathematical Biology* and has been presented at the *19th International Conference of System Biology* held in Lyon at fall.

Shape regulation. Reproducible and robust morphogenesis requires growth coordination of thousands of cells. How such coordination can be "implemented" in living organism is a core question for *RA3*. One identified mechanism in plant to coordinate growth rely on cells mechano-sensitivity [0]. Combined with the geometrical dependency of mechanical stress (*c.f.* previous subsection), this suggests the existence of a feedback mechanism that regulates tissue shape changes. We have been investigating closely the consequences of such a mechanism.

To that end, we first modeled the bio-molecular pathway relating mechanical stress experienced by cells to actual modification of their mechanical properties (*e.g.* cell wall stiffness). This work enabled us to describe plant tissues as an active material featuring large-scale properties, such as stress stiffening [0], emerging from sub-cellular dynamics. This work has been published in *Journal of Mathematical Biology* [5].

In parallel, we modeled the influence of cell wall elasticity (value, orientation) on the growth dynamics of tissues. This was done in the context of plant organogenesis, in close collaboration with biologists investigating the effect of cell-wall-related mutations on plant organ initiation. Our modeling approach was based on our previously developed *strain-based growth* model [20]. This joint study has been published in *Development* earlier this year [1].

We then studied how initial spherical symmetry (e.g. dome-shaped primordia) can be potentially broken during development in such active tissues and lead to elongated or flat shapes. For this, we integrated the *stress feedback* model with the *strain-based growth* model to investigate how their interplay could influence the morphogenesis of 3D cellularized structures. In particular, we showed that a stress-based feedback mechanism can maintain the typical plant growth modes (*i.e.* axial elongation or 2D flat expansion) and amplify asymmetries. This computational approach to symmetry breaking in growing tissues has been developed in parallel to experimental investigations addressing the shape evolution of sepals [0].

This work was part of the *Mophogenetics* IPL and Jan Traas ERC grant *Morphodynamics*.

The whole story has been presented at the *9th International Plant Biomechanics Conference* in Montreal this summer. A journal article combining both our modeling approach and experimental work in the context of symmetry breaking during plant organogenesis is currently being written.

Influence of water fluxes on plant morphogenesis. Since pressure appears as the "engine" behind growth-related deformation in plants, its regulation by cells is a major control mechanism of morphogenesis. We developed 2D computational models to investigate the morphological consequences of the interplay between cell expansion, water fluxes between cells and tissue mechanics. This interdisciplinary work, combining experiments and modeling, addresses the influence of turgor pressure heterogeneities on relative growth rate between cells. We showed that the coupling between fluxes and mechanics allows us to predict observed morphological heterogeneities without any *ad hoc* assumption. It also reveals the existence of a putative inhibitory action of organ growth on growth in immediatemy neighboring regions, due to the hydraulic coupling between cells during growth.

[0] the ability to probe mechanical stress around them and to modify accordingly their growth behavior
[0] the ability of the tissue to re-enforce itself in the directions of high mechanical solicitations
[0] leaf-like organs surrounding and protecting flowers

This work was part of the Agropolis fundation project *MecaFruit3D* and Arezki Boudaoud's ERC *PhyMorph*.

Two papers report the results of this work (one currently under review in *Nature Physics* [21] and a second one that is about to be submitted. These results have also been presented last summer at the *9th International Plant Biomechanics Conference* in Montreal.

Influence of dividing cells on tissue mechanics during morphogenesis in ascidians. The control of cell division orientation is of prime importance for patterning and shape emergence, especially in animal embryos where the first developmental stages happen at constant volume. In recent years, the Hetwig's rule appeared as a physical model accounting for orientation of cell division. Within animal tissues it has been shown that the coupling of externally induced strain and Hertwig's rule leads to the orientation of cell divisions with the main stress direction.

We investigated through modeling the consequences, in a multicellular context, of such stress-based regulation of cell division orientation. To that end, we developed a theoretical standpoint on the many-body energetic thermodynamics of cell divisions in the presence of external anisotropic stress. We showed that Hertwig's rule emerges as a limiting-case behavior and how anisotropic mechanical stresses can provide important cues to guide cell divisions. Our model accounts for the division pattern observed in the epidermis of the embryo of ascidian *Phallusia mammillata*, including those reproducible observed deviations from Hertwig's rule which have so far eluded explanation.

This work was part of the *Digem* project.

This work has been presented in two national conferences: the *IBC Scientific Days* and the *Cell Cycle Days* both held in Montpellier. A paper is currently being written.

Automatic quantification of adhesion defects in microscopy images. Direct measurements of mechanical stresses experienced by living tissues are not yet feasible. To circumvent this limitation, we developed an indirect method based on measurements of cracks in tissues: Our biologist colleagues developed cell-adhesion mutants in which strong connections between epithelial cells are impaired. As a consequence, mechanical stresses within the tissue produce cracks. Distribution and orientation of these cracks can be related to the main directions of the mechanical forces at play. We developed a 2D image analysis pipeline to detect and quantify these cracks in microscopy projections of epithelia, and deduce the magnitude and orientation of tensions in organs and tissues. This tool has been used to evidence new mechanical signaling mechanisms in *Arabidopsis*.

This analysis pipeline has been published in [6] and used by collaborators in the analysis performed in [26].

6.5. Signaling and transport for tissue patterning

Participants: Romain Azaïs, Guillaume Cerutti, Christophe Godin, Bruno Leggio, Jonathan Legrand, Teva Vernoux [External Collaborator].

- Research Axes: **RA1** (*Representations of forms in silico*) & **RA2** (*Data-driven models*)
- Key Modelling Challenges: **KMC3** (*Realistic integrated digital models*)

One central mechanism in the shaping of biological forms is the definition of regions with different genetic identities or physiological properties through bio-chemical processes operating at cellular level. Such patterning of the tissue is often controlled by the action of molecular signals for which active or passive transport mechanisms determine patterning spatial precision. The shoot apical meristem (SAM) of flowering plants is a remarkable example of such finely controlled system where the dynamic interplay between the hormone auxin and the polarization of efflux carriers PIN1 during growth governs the rhythmic patterning of organs, and the consequent emergence of phyllotaxis. Using *Arabidopsis thaliana* as a model system, we developed an integrated view of the meristem as a self-organizing dynamical form by reconstructing the dynamics of physiological processes from living tissues, and by proposing computational models integrating transport and signaling to study tissue patterning *in silico*.

Automatic quantification of auxin transport polarities. Time-lapse imaging of living SAM tissues marked with various fluorescent proteins allows monitoring the dynamics of cell-level molecular processes. Using a co-visualization of functional fluorescent auxin transporter (PIN1-GFP) with a dye staining of cell walls with propidium iodide (PI), we developed a method to quantify in 3D the polarization of auxin transport for every anticlinal wall of the first layer of cells. The digitally reconstructed network evidenced an overall stable convergence of PIN1 polarities towards the center of the meristem, with local front lines matching dynamic accumulations of auxin [15]. It also showed that the apparent crescent shape often thought to indicate polarities in cells might sometimes be misleading, and opens the way for a new view of how auxin transport is regulated.

Temporal auxin signaling in meristem organ patterning. Morphogenetic signals such as auxin define spatial distributions that are thought to control tissue patterning, but it has been proposed in animals that they also carry temporal information in their dynamics. A recent model developed by our group has postulated the existence of a stochastic mechanism to explain disturbed phyllotaxis patterns. This model assumes that organ initiation results from a temporal integration of a morphogenetic signal that buffers molecular noise [22]. Using a quantitative analysis of the dynamics of auxin distribution and response, we provide evidence that organ initiation in the SAM is indeed dependent on the temporal integration of the auxin signal [15]. The duration of cell exposition to auxin is used to differentiate temporally sites of organ initiation, and provide robustness to the rhythmic organ patterning.

Computational models of integrated transport and signaling. To interpret these new observations of auxin signaling and transport in the meristem, we investigate theoretical and computational models to study dynamic auxin distributions and the consequent organ patterning at the level of the meristem. Building on existing models of auxin transport [23], [25], we investigate different competing hypotheses on the auxin-PIN interplay, through numerical simulations based on rate equations for molecular transport and efflux carrier polarization. Quantitative comparisons with in vivo observations will provide cues on how the system responses are linked to memory effects and information exchanges between auxin and PINs.

These works were part of the *BioSensors* HFSP project and are carried out in the *Phyllo* ENS-Lyon project and gave rise to a journal article submitted for publication. These results have been presented at the *International Worskhop on Image Analysis Methods for the Plant Sciences* in Nottingham in January 2018 and in several invited talks given by Teva Vernoux and Christophe Godin.

6.6. Regulation of branching mechanisms in plants

Participants: Romain Azaïs, Frédéric Boudon [External Collaborator], Christophe Godin.

- Research Axes: **RA2** (*Data-driven models*) & **RA3** (*Plasticity & robustness of forms*)
- Key Modelling Challenges: **KMC3** (*Realistic integrated digital models*)

Branching in plants results from the development of apical meristems that recursively produce lateral meristems. These meristems may be more or less differentiated with respect to the apical meristem from which they originate, potentially leading to different types of lateral branches or organs. They also can undergo a more or less long period of inactivation, due to systemic regulation. The understanding of branching systems morphogenesis in plants thus relies on the analysis of the regulatory mechanisms that control both meristem differentiation and inactivation.

Analysis of the diversity of inflorescence architecture in different rice species. Rice is a major cereal for world food security and understanding the genetic and environmental determinants of its branching habits is a timely scientific challenge. The domestication, i.e., the empirical selection by humans, of rice began 10 000 years ago in Asia and 3 000 years ago in Africa. It thus provides a short-term model of the processes of evolution of plants.

Hélène Adam and Stéphane Jouannic from the group Evo-Devo de l'Inflorescence of UMR DIADE at IRD (Montpellier) have collected for years on the different continents an outstanding database of panicle-type inflorescence phenotypes in Asian and African, cultivated and wild, rice species. Classical statistical analysis based on the extraction of characteristic traits for each individual branching system were able to separate wild species from cultivated ones, but could not discriminate between wild species, suggesting that the entire branching structure should be used for classification methods to operate. For this, we are currently developing statistical methods on tree structures (see section 6.3) that should allow us to achieve better discrimination between panicles, based on their branching topology in addition to geometric traits. By coupling the quantitative study of the panicles to genomic analyses carried out by the IRD group, we should be able to highlight which regulation pathways have been selected or altered during the domestication process.

The role of sugars in apical dominance. The outgrowth of axillary buds is a key process in plant branching and which is often shown to be suppressed by the presence of auxin in nodal stems. However, local auxin levels are not always sufficient to explain bud outgrowth inhibition. Recent studies have also identified a contribution of sugar deprivation to this phenomenon. Whether sugars act independently of auxin or other hormones auxin regulates is unknown. Auxin has been shown to induce a decrease of cytokinin levels and to upregulate strigolactone biosynthesis in nodes. Based on rose and pea experiments, both in vitro and in planta, with our collaborators Jessica Bertheloot, Soulaiman Sakr from Institut de Recherche en Horticulture et Semences (IRHS) in Angers, we have shown that sucrose and auxin act antagonistically, dose-dependently, and non-linearly to modulate bud outgrowth. The Angers group provided experimental evidence that sucrose represses bud response to strigolactones but does not markedly affect the action of auxin on cytokinin levels. Using a modeling approach, we tested the ability of this complex regulatory network to explain the observed phenotypes. The computational model can account for various combinations of sucrose and hormones on bud outgrowth in a quantitative manner and makes it possible to express bud outgrowth delay as a simple function of auxin and sucrose levels in the stem. These results provide a simple auxin-sucrose-cytokinin-strigolactone network that accounts for plant adaptation to growing conditions. A paper relating this work is currently under review.

The fractal nature of plants. Inflorescence branching systems are complex and diverse. They result from the interaction between meristem growth and gene regulatory networks that control the flowering transition during morphogenesis. To study these systems, we focused on cauliflower mutants, in which the meristem repeatedly fails in making a complete transition to the flower and for which a complete mechanistic explanation is still lacking.

In collaboration with Eugenio Azpeitia and François Parcy's group in Grenoble, we have developed a first model of the control of floral initiation by genes, refining previous networks from the literature so that they can integrate our hypotheses about the emergence of cauliflower phenotypes. The complete network was validated by multiple analyses, including sensitivity analyses, stable state analysis, mutant analysis, among others. It was then coupled with an architectural model of plant development using L-systems. The coupled model was used to study how changes in gene dynamics and expression could impact in different ways the architectural properties of plants. The model was then used to study how changes in certain parameters could generate different curd morphologies, including the normal and the fractal-like Romanesco. A paper reporting this work is currently being written.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

Participants: Frédéric Boudon [External Collaborator], Christophe Godin.

We started a collaboration with A.M.R. a start-up whose aim is to develop a web application to create social networks for project management. This application makes use of plant representations at different levels for which the expertize of the Mosaic group was required. In 2018, we hosted two internships during 6 months in co-supervision with Guillaume Asselot (founder of A.M.R.) to work on plant models for the web application. Guillaume Asselot is seeking to raise new funds to pursue the collaboration in the coming years.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. ENS de Lyon projets Emergents - Phyllo (2018 - 2019)

Participants: Christophe Godin, Bruno Leggio, Teva Vernoux [External Collaborator].

The aim in this project is to develop a model of phyllotaxis that would be compatible with the recent detailed and quantitative observations made by our group of the distribution of auxin in space and time at the SAM. In particular the work will seek at using the new quantitative data to estimate the parameters of the stochastic model previously developed of organ patterning.

8.1.2. IDEX Lyon Impulsion - MecaField (2019 - 2020

Participants: Christophe Godin, Teva Vernoux [External Collaborator].

In a previous work, we have shown that the coupling of mechanical and hydraulical descriptions in a 2D model of multicellular tissue growth induces the emergence of remarkable phenomena at tissue level. In particular, we have shown that the growth of an organ may induce a lateral inhibition surrounding the organ that prevents other organs to grow in its vicinity. The goal of this project is to estimate the hydraulic and mechanical parameters of such a model from confocal images of a growing SAM and to compare observations with the order of magnitude of the predicted inhibitory zones and of their amplitude at cellular resolution.

8.2. National Initiatives

8.2.1. Inria ADT - Gnomon

Participants: Olivier Ali, Romain Azaïs, Guillaume Cerutti, Florian Gacon, Christophe Godin, Jonathan Legrand, Grégoire Malandain [External Collaborator], Teva Vernoux [External Collaborator].

Gnomon is a user-friendly computer platform developed by the Mosaic team for seamless simulation of form development in silico. It is intended to be a major tool for the team members to develop, integrate and share their models, algorithms and tools. Flexible components (plugins) make it possible to up-load or to create such data-structures, to program their development, to analyze, visualize them and interact with them in 3D+time.

Based on the past experience of the team with the OpenAlea platform, the goal of this ADT is to develop a more scalable software engineering solution based on the dtk kernel developed by the group of software engineers (SED) from the Sophia-Antipolis Inria Center.

Partners:

- SED Sophia Antipolis Inria Research Centre
- Morpheme Inria projec-team, Sophia Antipolis, France

8.2.2. Inria IPL - Naviscope

Participants: Guillaume Cerutti, Emmanuel Faure [External Collaborator], Christophe Godin, Jonathan Legrand, Grégoire Malandain [External Collaborator].

In this project, we plan to develop original and cutting-edge visualization and navigation methods to assist scientists, enabling semi-automatic analysis, manipulation, and investigation of temporal series of multi-valued volumetric images, with a strong focus on live cell imaging and microscopy application domains. We will build Naviscope upon the strength of scientific visualization and machine learning methods in order to provide systems capable to assist the scientist to obtain a better understanding of massive amounts of information. Such systems will be able to recognize and highlight the most informative regions of the dataset by reducing the amount of information displayed and guiding the observer attention. Finally, we will overcome the technological challenge of gathering up the software developed in each team to provide a unique original tool for users in biological imaging, and potentially in medical imaging.

8.2.3. ANR - ReVeRIES

Participant: Guillaume Cerutti.

The aim of ReVeRIES (Reconnaissance de Végétaux Récréative, Interactive et Educative sur Smartphone) is to make use of mobile technologies to transmit general knowledge and identification skills on the plant world to an urban audience who has little to no botanical background. Following the work of the ReVeS project and the development of the Folia mobile application, a major objective is to recognize automatically the species of trees and shrubs encountered in France using photographs of their leaves, fruits, flowers and barks, while providing the user the botanical vocabulary and the keys to learn how to identify species.

Partners:

- EVS Laboratoire Environnement Ville et Société, Saint-Etienne
- IRHS Institut de Recherches en Horticulture et Semences, Angers
- LIRIS Laboratoire d'Informatique en Image et Système d'Information, Lyon
- LISTIC Laboratoire d'Informatique, Système, Traitement de l'Information et de la Connaissance, Annecy
- LIUM Laboratoire d'Informatique de l'Université du Maine, Le Mans

8.2.4. ANR - Imago (2016 - 2019)

Participants: Guillaume Cerutti, Christophe Godin, Jonathan Legrand.

The goal of this project is to investigate the role of ovule growth constraints on germ cell fate establishment. This project is motivated by recent findings from the partners' groups suggesting that disturbances in cell divisions and expansion in early (pre-meiotic) ovules are sufficient to induce ectopic germ cells. These observations suggest novel routes to engineer apomixis in plants but remains poorly understood. Recent developments in high-resolution 3D imaging, image processing, and modeling offer a powerful combination of approaches to investigate this question. IMAGO proposes to elucidate patterning rules governing ovule growth, and their contribution to female germ cell fate acquisition. We use a combination of high-resolution static and real-time 3D imaging, quantitative image processing, cell-based growth models and functional approaches to (1) define cellular growth patterns in the ovule primordium using quantitative imaging (2) test patterning rules in silico by cell-based growth models (3) validate patterning rules in vivo using genetic, pharmacological and mechanical perturbations.

Partners:

- UMR DIADE, IRD, Montpellier, France
- Department of Plant and Microbial Biology, Zurich, Swisszerland
- RDP, ENS de Lyon, France

8.2.5. ANR DigEM (2015 - 2019)

Participants: Christophe Godin, Bruno Leggio, Patrick Lemaire [External Collaborator], Grégoire Malandain [External Collaborator].

In this project, we will use advanced ligh-sheet imaging of live embryos to quantitatively describe embryonic morphogenesis in ascidians, a class of animals that undergo very rapid genomic divergence, yet show an extraordinary stasis of embryonic morphologies, based on invariant early cell lineages shared by all studied species. The global aims of the proposal, which will bridge micro- and macroevolutionary scales of analysis, are: i) to provide a global systems-level description at cellular resolution of an animal embryonic program; ii) to use this description to characterize intra-specific and inter-specific patterns of morphogenetic variations; iii) to analyze possible molecular mechanisms explaining the unusual robustness of this program to environmental and genetic perturbations. To achieve these aims, we will combine advanced live light-sheet microscopy, computational biology, functional gene assays and evolutionary approaches.

Partners:

- UMR CRBM, CNRS Montpellier, France
- Morpheme Inria projec-team, Sophia Antipolis, France

8.2.6. ERA-CAPS Genes2shape (2018 - 2021)

Participants: Olivier Ali, Guillaume Cerutti, Christophe Godin, Bruno Leggio, Jan Traas [External Collaborator].

This project is aimed at understanding how molecular regulation integrates with mechanics to control overall plant shape, an unresolved problem with wide implications for both fundamental and applied biology. We will address this issue in the Arabidopsis flower, which, besides their obvious importance as reproductive structures, are amongst the best characterised systems in plant developmental biology. From a mechanistic point of view, it is widely accepted that regulatory molecular networks interfere with the properties of the structural cellular elements (cell wall, cytoskeleton) to induce particular growth patterns. How this occurs and how this is coordinated in space is not known. To obtain a mechanistic understanding of such a complex process, information from multiple scales, from molecular networks to physical properties and geometry have to be combined into a single picture. An integrated tool to do so is currently not available. Building on our complementary experience in interdisciplinary research on plant development, we will therefore develop a tool, called the "Computable Flower" that permits (i) integration of data on geometry, gene expression and biomechanics and (ii) the user to explore, interpret and generate hypotheses based on data supported by mechanistic modelling approaches. The tool therefore provides an integrated description in the form of a 3D dynamic template of the growing flower bud.

Partners:

- University of Cambridge (Sainsbury Lab.)
- California Institute of Technology
- MaxPlanck Institutes of Molecular Plant Physiology

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

Program: H2020

Project acronym: ROMI

Project title: RObotics for MIcrofarms

Duration: November 2017 - October 2021

Coordinator: Sony

Other partners: Iaac, (Spain), FEI (France), Inria (France), CNRS (France), UBER (Germany), Chatelain (France)

Abstract: All over Europe, young farmers are starting small market farms and direct sales businesses. These farms can be found both in rural, peri-urban and urban areas. They grow a large variety of crops (up to 100 different varieties of vegetables per year) on small surfaces (0.01 to 5 ha) using organic farming practices. These farms have proven to be highly productive, sustainable and economically viable. However, a lot of work is done manually, resulting in physically challenging work conditions. ROMI will develop an open and lightweight robotics platform for these microfarms. We will assist these farms in weed reduction and crop monitoring. This will reduce manual labour and increase the productivity through advanced planning tools. Thanks to ROMI's weeding robot, farmers will save 25 percents of their time. This land robot will also acquire detailed information on sample plants and will be coupled with a drone that acquires more global information at crop level. Together, they will produce an integrated, multi-scale picture of the crop development that will help the farmer monitor the crops to increase efficient harvesting. For this, ROMI will have to adapt and extend state-of-the-art land-based and air-borne monitoring tools to handle small fields with complex layouts and mixed crops. To achieve this, we will: (i) develop and bring to the market an affordable, multi-purpose, land-based robot, (ii) develop a weeding app for this robot that is adapted for organic microfarms, (iii) apply advanced 3D plant analysis and modelling techniques to in-field data acquisition, (iv) integrate these analysis techniques in the robot for detailed plant monitoring, (iv) integrate these techniques also in the aerial drone N-E-R-O for multi-scale crop monitoring, (v) extend the robot with novel, adaptive learning techniques to improve sensorimotor control of the plant monitoring app, and (vii) test the effectiveness of our solution in real-world field conditions.

8.3.2. Inria International Partners

8.3.2.1. Informal International Partners

8.3.2.1.1. Laboratoire International Associé (LIA): Computing Plant Morphogenesis

The focus of this LIA headed by Teva Vernoux (RDP) and Ottoline Leyser (SLCU) is on plant morphogenesis i.e. the mechanisms allowing the generation of plant shapes at different scales. Both the RDP and SLCU Laboratories are leaders of this field. The scenario for morphogenesis that has recently emerged is that chemical signals controlling cell identities lead to changes in mechanical properties of cells, triggering changes in shapes feeding back on the gene regulatory network. This in turn affects the distribution of chemical signals and mechanical forces, thus channeling morphogenesis. However, our understanding of the molecular and physical basis of morphogenesis in plants or in any other eukaryotic system is still in its infancy due to the complexity and non-linearity of processes involved in morphogenesis dynamics (or Morphodynamics). Understanding morphodynamics requires a modeling environment for the explicit representation of forms at multiple scales and for incorporating complex data from different origins and nature (chemical, mechanical, geometrical). In addition to creating a unique scientific environment, this LIA will gather the critical mass and interdisciplinary expertise required to create such a computational platform and to generate the data to produce an integrated vision of how chemical and mechanical signals interaction drive morphogenesis.

Partners:

- Sainsbury Lab. University of Cambridge (SLCU)

8.4. International Research Visitors

8.4.1. Visits of International Scientists

8.4.1.1. Internships

Farah Ben Naoum, associate professor at the University of Sidi Bel Abbes, paid a one-month visit (July 2018) in the Mosaic group to work with Romain Azaïs and Christophe Godin on algorithms to compute incrementally tree-edit distances based on their directed acyclic graph representation. This visit was funded by Inria and will be followed by another one month visit in March 2019 to complete the writing a related paper.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

- Romain Azaïs is a member of the organizing committee of the *Journées de Statistique 2019* in Nancy.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- C. Godin is:
  - Associate Editor of the journal Frontiers in Plant Sciences, section Plant Biophysics and Modeling.
  - Review Editor of the journal Frontiers in Plant Sciences, section Plant Systems and Synthetic Biology.
  - Member of the Editorial Advisory Board of the new journal *in silico* Plants.
  - Associate Editor of a special issue of the Bulletin of Mathematical Biology on Multi-scale modelling of Tissue Growth and Shape.
- Olivier Ali: Review Editor in *Plant Biophysics & Modeling for Frontiers* since January 2018

9.1.2.2. Reviewer - Reviewing Activities

All team members are regularly involved in reviewer activities for academic journals from different scientific fields: Current Biology, Annals of Applied Probability, Journal of Theoretical Biology, Physical Review Letters, Journal of BioPhysics, Physical Review A, etc.

9.1.3. Invited Talks

- Romain Azaïs:
  - *Arbres de Galton-Watson conditionnés par la taille ou la hauteur : estimation via les modèles limites*. Journée de rencontres scientifiques autour de la statistique pour la biologie et la médecine, Poitiers, February 2018.
  - *Quelques pistes pour l'analyse de données arborescentes*. Rencontres Statistiques Lyonnaises, March 2018 (I/II) and May 2018 (II/II).
  - *Un nouvel éclairage sur le subtree kernel pour données arborescentes*. Séminaire de Probabilités et Statistiques de l'Université de Montpellier, September 2018.
  - *Average number of crossings and rupture detection for piecewise-deterministic processes*. International conference STODEP, Rouen, October 2018.
- Christophe Godin:
  - *Can we manipulate tree forms like numbers?*, co-authored with Romain Azaïs. Workshop on Mathematics for Developmental Biology, BIRS Center for Mathematics in Banff, Canada, Dec. 2017.
  - *Functional-Structural modeling of plants*, co-authored with Arezki Boudaoud. Phytobiom days in Lyon. France, April 2018.
  - *Phyllotaxis at the era of molecular and computational biology: the revival of an old enigma*, co-authored with Teva Vernoux. Jacques Monod series of international conferences in Roscoff, France, Sept 2018.
  - *Coupling mechanical and hydraulic processes in multicellular models of plant development*, co-authored with Ibrahim Cheddadi. Workshop on the Contribution of cell mechanics to cell fate determination and tissue integrity: from an interdisciplinary point of view. Fondation des Treilles, France, Oct. 2018.
- Bruno Leggio (with J. Laussu, E. Faure, C. Godin and P. Lemaire): *Reproducible epidermal morphogenesis in ascidians: an active-reactive mechanical model*. Cell Cycle Day 2018, Montpellier, September 2018.

9.1.4. Scientific Expertise

- C. Godin
  - is a member of the International Scientific Advisory Committee of the Plant Phenotyping and Imaging Research Centre (P2IRC), Saskatchewan, Canada.

  - is a member of the Scientific Board of the Plant Biology and Breeding Department of INRA (BAP).
  - has reviewed a project for the Strategic Basic Research programme of the Flanders fundation for research (FWO), Belgium.

9.1.5. Research Administration

- C. Godin
  - is a member of the Project Committee at Grenoble Rhone-Alpes Research Center.
  - is a member of the Steering Committee of the RDP Lab., Lyon.
  - is a member of the Scientific board of the modeling axis of Labex NUMEV and a member of the direction board of the institut de biologie computationnelle (IBC) in Montpellier.
  - has been President of the 2018 CRCN Selection Jury at Inria Sophia Antipolis-Méditerranée

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Romain Azaïs:
  - *Colles de Mathématiques*, CPGE PCSI, Lycée Jean Perrin, Lyon.
  - One-day course on *Statistical learning*, Bioinformatics Summer School, Angers, July 2018.
- Guillaume Cerutti: *Practicals in modeling for Biosciences*, M2 ENS de Lyon (20h TP)
  Coordinator: Arezki Boudaoud, RDP, ENS de Lyon.
- Christophe Godin:
  - Co-organized with Patrick Lemaire the International Spring School on *Animal and Plant Morphogenesis*, one week in March, Paris ENS Master, Montpellier, France.
  - *Introduction to Microscopy image analysis*. Master Biology ENS de Lyon (2h).
- Bruno Leggio:
  - Conception and supervision of practicals for masters and PhD students during the 2018 *Interdisciplinary spring school on animal and plant morphogenesis*, Hameau de l'Etoile, Montpellier, France, March 2018.
  - One-day class for biophysics students at the University of Montpellier on the *modeling of morphogenesis*, December 2018.

9.2.2. Supervision

- PhD (2015 – 2018): Sarah Bertrand (LIRIS, Université Lumière Lyon 2). *Analyse d'images pour l'identification multi-organes d'espèces végétales*. Guillaume Cerutti. Supervisors: Laure Tougne (LIRIS, Université Lumière Lyon 2) and Guillaume Cerutti.
- PhD in progress (2015 – 2019): Florine Greciet (IECL, Université de Lorraine and Safran). *Modèles markoviens déterministes par morceaux cachés pour la caractérisation de contrainte admissible en fatigue des matériaux*. Supervisors: Anne Gégout-Petit (Inria team BIGS, IECL, Université de Lorraine) and Romain Azaïs.
- Master 1 (3 months in 2018): Florian Ingels. *Méthodes à noyau pour données arborescentes*. Supervisor: Romain Azaïs.
- Licence 3 (2 months in 2018): Clément Legrand. *A COMPLETER*. Supervisor: Christophe Godin.
- Niveau à completer (durée à compléter): Tony Vincent Ang. *A COMPLETER*. Supervisors: Christophe Godin and Frédéric Boudon.

- Niveau à completer (durée à compléter): Renan Berruex. *A COMPLETER*. Supervisors: Christophe Godin and Frédéric Boudon.
- Penser à Hadrien Oliveri, Katia Mirande, Anne Schneider

9.3. Popularization

9.3.1. Articles and contents

- Christophe Godin
  - made a presentation at *My team in 180 seconds*, June 2018 (in French).
  - wrote a paper *Pourquoi les Plantes font des maths* co-authored with Fabrice Besnard and Teva Vernoux, that was published in Pour La Science (Aug 2018) and in the hors-série issue on *La révolution végétale* (Nov 2018). (Nov 2018)

9.3.2. Interventions

- O. Ali has been an invited speaker for an art/science conference and debate organized by the modern art museum of Lyon, October 2018.
- Romain Azaïs won a special mention in the comics contest *Maths et Polar* organized by the website *Image des Mathématiques* of the CNRS, April 2018.
- Guillaume Cerutti, Christophe Godin and Jonathan Legrand have been involved in Déclics initiatives (presentation and discussion with high school students in order to promote science and research careers), October and December 2018.
- Christophe Godin gave an invited 2 hours seminar to 4 classes of high-school students for the Math week (March) at the Lyce ´e International de Valbonne, France.
- H. Oliveri took part in the Class'Code Initiative, a Mooc produced by Inria and OpenClassRoom providing supports and guidelines for people wanting to teach computer science to young childrens.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] A. ARMEZZANI, U. ABAD, O. ALI, A. ANDRES ROBIN, L. VACHEZ, A. LARRIEU, E. MELLEROWICZ, L. TACONNAT, V. BATTU, T. STANISLAS, M. LIU, T. VERNOUX, J. TRAAS, M. SASSI.*Transcriptional induction of cell wall remodelling genes is coupled to microtubule-driven growth isotropy at the shoot apex in Arabidopsis*, in "Development (Cambridge, England)", June 2018, vol. 145, n^{o} 11 [*DOI :* 10.1242/DEV.162255], https://hal.archives-ouvertes.fr/hal-01894645

[2] R. AZAÏS, A. GENADOT.*Estimation of the average number of continuous crossings for non-stationary non-diffusion processes*, in "Journal of Statistical Planning and Inference", 2018, https://arxiv.org/abs/1703.10192 , https://hal.archives-ouvertes.fr/hal-01499914

[3] S. BERTRAND, R. BEN AMEUR, G. CERUTTI, D. COQUIN, L. VALET, L. TOUGNE.*Bark and Leaf Fusion Systems to Improve Automatic Tree Species Recognition*, in "Ecological Informatics", June 2018 [*DOI :* 10.1016/J.ECOINF.2018.05.007], https://hal.archives-ouvertes.fr/hal-01811039

[4] K. BIASUZ, B. LEGGIO, E. FAURE, P. LEMAIRE.*The "computable egg": Myth or useful concept?*, in "Current Opinion in Systems Biology", October 2018, vol. 11, p. 91-97 [*DOI :* 10.1016/J.COISB.2018.09.003], https://hal.archives-ouvertes.fr/hal-01919148

[5] H. OLIVERI, J. J. TRAAS, C. GODIN, O. ALI.*Regulation of plant cell wall stiffness by mechanical stress: a mesoscale physical model*, in "Journal of Mathematical Biology", September 2018, p. 1-29 [*DOI :* 10.1007/S00285-018-1286-Y], https://hal.inria.fr/hal-01691110

[6] S. VERGER, G. CERUTTI, O. HAMANT.*An Image Analysis Pipeline to Quantify Emerging Cracks in Materials or Adhesion Defects in Living Tissues*, in "Bio-protocol ", 2018, vol. 8, n^{o} 19, https://hal.inria.fr/hal-01940178

International Conferences with Proceedings

[7] R. AZAÏS, J.-B. DURAND, C. GODIN.*Approximation of trees by self-nested trees*, in "ALENEX 2019 - Algorithm Engineering and Experiments", San Diego, United States, January 2019, p. 1-24, https://arxiv.org/abs/1810.10860 [*DOI :* 10.10860], https://hal.archives-ouvertes.fr/hal-01294013

Scientific Books (or Scientific Book chapters)

[8] R. AZAÏS, A. GENADOT.*Level Crossings and Absorption of an Insurance Model*, in "Statistical Inference for Piecewise-deterministic Markov Processes", R. AZAÏS, F. BOUGUET (editors), Wiley, August 2018, p. 65-105 [*DOI :* 10.1002/9781119507338.CH3], https://hal.archives-ouvertes.fr/hal-01862266

[9] R. AZAÏS, A. GÉGOUT-PETIT, F. GRECIET.*Rupture Detection in Fatigue Crack Propagation*, in "Statistical Inference for Piecewise-deterministic Markov Processes", Wiley, August 2018, p. 173-207 [*DOI :* 10.1002/9781119507338.CH6], https://hal.archives-ouvertes.fr/hal-01862267

Books or Proceedings Editing

[10] R. AZAÏS, F. BOUGUET (editors). *Statistical Inference for Piecewise-deterministic Markov Processes*, Wiley, August 2018, p. i-xxxiv 1-260 [*DOI :* 10.1002/9781119507338], https://hal.archives-ouvertes.fr/hal-01862248

Scientific Popularization

[11] F. GRECIET, R. AZAÏS, A. GÉGOUT-PETIT.*Detection and modeling of the propagation regimes in fatigue crack propagation*, in "ENBIS 2018 - 18th Annual Conference of the European Network for Business and Industrial Statistics", Nancy, France, September 2018, https://hal.inria.fr/hal-01942243

Other Publications

[12] R. AZAÏS, A. GENADOT, B. HENRY.*Inference for conditioned Galton-Watson trees from their Harris path*, November 2018, https://arxiv.org/abs/1609.01057 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01360650

[13] S. BERTRAND, G. CERUTTI, L. TOUGNE.*Visualization of Leaf Botanical Features Extracted from AlexNet Convolutional Layers*, January 2018, IAMPS - International Workshop on Image Analysis Methods for the Plant Sciences 2018, Poster, https://hal.archives-ouvertes.fr/hal-01691924

[14] G. CERUTTI, C. GALVAN-AMPUDIA, J. LEGRAND, T. VERNOUX, C. GODIN.*A parametric active 4D model for the automatic construction of developmental atlases from confocal images of shoot apical meristems*, January 2018, IAMPS - International Workshop on Image Analysis Methods for the Plant Sciences 2018, Poster, https://hal.inria.fr/hal-01692584

[15] C. GALVAN-AMPUDIA, G. CERUTTI, J. LEGRAND, R. AZAÏS, G. BRUNOUD, S. MOUSSU, C. WENZL, J. U. LOHMANN, C. GODIN, T. VERNOUX.*From spatio-temporal morphogenetic gradients to rhythmic patterning at the shoot apex*, November 2018, working paper or preprint [*DOI :* 10.1101/469718], https://hal.inria.fr/hal-01937928

[16] L. GUIGNARD, U.-M. FIUZA, B. LEGGIO, E. FAURE, J. LAUSSU, L. HUFNAGEL, G. MALANDAIN, C. GODIN, P. LEMAIRE.*Contact-dependent cell-cell communications drive morphological invariance during ascidian embryogenesis*, November 2018, working paper or preprint [*DOI :* 10.1101/238741], https://hal.archives-ouvertes.fr/hal-01938126

[17] B. LEGGIO, J. LAUSSU, A. CARLIER, C. GODIN, P. LEMAIRE, E. FAURE.*MorphoNet: An interactive online morphological browser to explore complex multi-scale data*, November 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01938153

[18] H. OLIVERI, F. ZHAO, O. ALI, J. TRAAS, C. GODIN.*Stress-based regulation of multicellular plant growth: a finite element modeling approach applied to planar leaf morphogenesis*, October 2018, ICSB 2018 - 19th International Conference on Systems Biology, Poster, https://hal.inria.fr/hal-01897027

References in notes

[19] K. BHARATH, P. KAMBADUR, D. DEY, R. ARVIN, V. BALADANDAYUTHAPANI.*Statistical tests for large tree-structured data*, in "Journal of the American Statistical Association", 2016

[20] F. BOUDON, J. CHOPARD, O. ALI, B. GILLES, O. HAMANT, A. BOUDAOUD, J. TRAAS, C. GODIN.*A Computational Framework for 3D Mechanical Modeling of Plant Morphogenesis with Cellular Resolution*, in "PLoS Computational Biology", January 2015, vol. 11, n^{o} 1, p. 1-16 [*DOI :* 10.1371/JOURNAL.PCBI.1003950], https://hal.archives-ouvertes.fr/hal-01142486

[21] Y. LONG, I. CHEDDADI, V. MIRABET, M. DUMOND, C. GODIN, A. BOUDAOUD.*Cellular heterogeneity in pressure and growth emerges from tissue topology and geometry*, 2018 [*DOI :* 10.1101/334664], https://www.biorxiv.org/content/early/2018/05/30/334664

[22] Y. REFAHI, G. BRUNOUD, E. FARCOT, A. JEAN-MARIE, M. PULKKINEN, T. VERNOUX, C. GODIN.*A stochastic multicellular model identifies biological watermarks from disorders in self-organized patterns of phyllotaxis*, in "eLife", jul 2016, vol. 5, e14093, https://doi.org/10.7554/eLife.14093

[23] R. S. SMITH, S. GUYOMARC'H, T. MANDEL, D. REINHARDT, C. KUHLEMEIER, P. PRUSINKIEWICZ.*A plausible model of phyllotaxis*, in "Proceedings of the National Academy of Sciences", 2006, vol. 103, n^{o} 5, p. 1301–1306 [*DOI :* 10.1073/PNAS.0510457103], http://www.pnas.org/content/103/5/1301

[24] A. J. SMOLA, S. VISHWANATHAN.*Fast kernels for string and tree matching*, in "Advances in neural information processing systems", 2003, p. 585–592

[25] S. STOMA, M. LUCAS, J. CHOPARD, M. SCHAEDEL, J. TRAAS, C. GODIN.*Flux-Based Transport Enhancement as a Plausible Unifying Mechanism for Auxin Transport in Meristem Development*, in "PLOS Computational Biology", 10 2008, vol. 4, n^{o} 10, p. 1-15, https://doi.org/10.1371/journal.pcbi.1000207

[26] S. VERGER, Y. LONG, A. BOUDAOUD, O. HAMANT.*A tension-adhesion feedback loop in plant epidermis*, in "eLife", apr 2018, vol. 7, e34460, https://doi.org/10.7554/eLife.34460

Project-Team NANO-D

Algorithms for Modeling and Simulation of Nanosystems

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

IN PARTNERSHIP WITH:
CNRS

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Numerical schemes and simulations

Table of contents

Project-Team NANO-D

Creation of the Team: 2008 January 01, updated into Project-Team: 2014 July 01, end of the Project-Team: 2018 September 30

Keywords:

Computer Science and Digital Science:

A3.4.1. - Supervised learning
A3.4.6. - Neural networks
A5.5.1. - Geometrical modeling
A6.1.4. - Multiscale modeling
A6.1.5. - Multiphysics modeling
A9. - Artificial intelligence

Other Research Topics and Application Domains:

B1.1.1. - Structural biology
B1.1.7. - Bioinformatics
B2.6.3. - Biological Imaging

1. Team, Visitors, External Collaborators

Research Scientists

Stephane Redon [Team leader before 2018, Inria, Senior Researcher, HDR]
Sergei Grudinin [Team leader since 2018, CNRS, Researcher]
Leonard Jaillet [Inria, Starting Research Position]

Technical Staff

Jocelyn Gate [Inria, until Sep 2018, granted by FP7 PoC SAMSON project]
Yassine Naimi [Inria, until Sep 2018]

PhD Students

Guillaume Pages [Inria]
Maria Kadukova [Univ Grenoble Alpes]
Francois Rousse [Inria, until May 2018]
Semeho Edorh [Inria, until Jan 2018]
Alexandre Hoffmann [Inria, until Jan 2018]
Minh Khoa Nguyen [Inria, until Jan 2018]

Post-Doctoral Fellows

Didier Devaurs [Univ Grenoble Alpes, from Dec 2018]
Dmitriy Marin [Inria, until Sep 2018]

Administrative Assistant

Maria Immaculada Presseguer [Inria]

2. Overall Objectives

2.1. Overview

During the twentieth century, the development of macroscopic engineering has been largely stimulated by progress in numerical design and prototyping: cars, planes, boats, and many other manufactured objects are nowadays designed and tested on computers. Digital prototypes have progressively replaced actual ones, and effective computer-aided engineering tools have helped cut costs and reduce production cycles of these macroscopic systems.

The twenty-first century is most likely to see a similar development at the atomic scale. Indeed, the recent years have seen tremendous progress in nanotechnology - in particular in the ability to control matter at the atomic scale. The nanoscience revolution is already impacting numerous fields, including electronics and semiconductors, textiles, energy, food, drug delivery, chemicals, materials, the automotive industry, aerospace and defense, medical devices and therapeutics, medical diagnostics, etc. According to some estimates, the world market for nanotechnology-related products and services will reach one trillion dollars by 2015. Nano-engineering groups are multiplying throughout the world, both in academia and in the industry: in the USA, the MIT has a “NanoEngineering” research group, Sandia National Laboratories created a “National Institute for Nano Engineering”, to name a few; China founded a “National Center for Nano Engineering” in 2003, etc. Europe is also a significant force in public funding of nanoscience and nanotechnology.

Similar to what has happened with macroscopic engineering, powerful and generic computational tools will be employed to engineer complex nanosystems, through modeling and simulation.

Modeling and simulation of natural or artificial nanosystems is still a challenging problem, however, for at least three reasons: (a) the number of involved atoms may be extremely large (liposomes, proteins, viruses, DNA, cell membrane, etc.); (b) some chemical, physical or biological phenomena have large durations (e.g., the folding of some proteins); and (c) the underlying physico-chemistry of some phenomena can only be described by quantum chemistry (local chemical reactions, isomerizations, metallic atoms, etc.). The large cost of modeling and simulation constitutes a major impediment to the development of nanotechnology.

The NANO-D team aims at developing efficient computational methods for modeling and simulation of complex nanosystems, both natural (e.g., the ATPase engine and other complex molecular mechanisms found in biology) and artificial (e.g., NEMS - Nano Electro-Mechanical Systems).

In particular, the group develops novel multiscale, adaptive modeling and simulation methods, which automatically focus computational resources on the most relevant parts of the nanosystems under study.

2.2. Research axes

The goal of the NANO-D group is to help current and future designers of *nanosystems*, i.e. systems studied or designed at the atomic scale (whether natural or artificial, independently of the application domain, including structural biology, material science, chemistry, etc.) by developing the **foundations of a software application which will run on a desktop computer, and will allow for efficient analysis, design, modeling and simulation of nanosystems**.

To achieve this, we will be developing a series of **adaptive methods and algorithms** that allow users to focus computational resources on the parts of the models that they want to simulate, and that allow to finely trade between speed and precision.

In parallel, we will develop the architecture of a new desktop application for virtual prototyping of nanosystems, and will integrate all our algorithms into this application. Furthermore, the architecture of this platform will be open, so that independent developers may add modules, for **multiple application domains** (physics, biology, chemistry, materials, electronics, etc.). With this open platform, we will attempt to federate the research performed in computational nanoscience throughout the world.

This application is called **SAMSON: “Software for Adaptive Modeling and Simulation Of Nanosystems”**.

Our two research axes are:

1. **Developing adaptive algorithms for simulating nanosystems**
   - **Defining adaptive Hamiltonians**: In order to be able to perform simulations with good mathematical properties, we are expanding on our recent work on *adaptively restrained Hamiltonians* [37], *i.e.* modified Hamiltonian representations of molecular systems that are able to switch degrees of freedom on and off during a simulation. These will allow us to finely trade between precision and computational performance, by choosing arbitrarily the number of degrees of freedom. Even though we have already obtained some promising results in this domain, our goal is to develop several different simplification methods.

- **Developing algorithms for incremental potential update**: In order to benefit from performing adaptive particle simulations, we need to develop a series of algorithms that will take advantage of the fact that some (potentially relative) atomic positions are frozen. We have already demonstrated how this is possible for torsion-angle quasi-static simulation of classical bio-molecular force-fields [80], for neighbor search between large rigid molecules [36], and for bond-order reactive force-fields [40]. We are developing new algorithms for incremental neighbor search, energy and force updates corresponding to the adaptive Hamiltonians that we are defining.

2. **Developing algorithms for modeling molecular interactions**
   - **Developing knowledge-driven methods, potentials and algorithms**: Over time, more and more experimental information becomes available. One can use this information to predict and discover new types of molecular interactions and various mechanisms or molecular organization. For example, currently there are more than 50,000 protein structures of a high resolution stored in the Protein Data Bank [38] and over 500,000 structures of small molecules stored in the Cambridge Structural Database [32]. We are developing algorithms for protein-protein interactions and protein-ligand interactions.
   - **Developing parametrization algorithms for interaction potentials**: Molecular models typically require their own potential energy function (or a *forcefield*) to be assigned. However, the development of a new potential function is a very difficult and sometimes challenging task [59]. Therefore, we are developing algorithms for automatic parametrization of new potential functions for some particular representations of a molecular system.
   - **Developing algorithms for exhaustive sampling**: Some application domains, such as computational docking, cryo-EM rigid-body fitting, etc., require sampling in a low-dimensional space. For such applications it is advantageous to perform an exhaustive search rather than accelerated sampling [77]. Therefore, we are developing fast search methods to perform exhaustive search.

3. Research Program

3.1. The need for practical design of nanosystems

Computing has long been an essential tool of engineering. During the twentieth century, the development of macroscopic engineering has been largely stimulated by progress in numerical design and prototyping. Cars, planes, boats, and many other manufactured objects are nowadays, for the most part, designed and tested on computers. Digital prototypes have progressively replaced actual ones, and effective computer-aided engineering tools (e.g., CATIA, SolidWorks, T-FLEX CAD, Alibre Design, TopSolid, etc.) have helped cut costs and reduce production cycles of macroscopic systems [79].

The twenty-first century is most likely to see a similar development at the atomic scale. Indeed, the recent years have seen tremendous progress in nanotechnology. The magazine Science, for example, recently featured a paper demonstrating an example of DNA nanotechnology, where DNA strands are stacked together through programmable self-assembly [49]. In February 2007, the cover of Nature Nanotechnology showed a “nanowheel” composed of a few atoms only. Several nanosystems have already been demonstrated, including a *de-novo* computationally designed protein interface [50], a wheelbarrow molecule [60], a nano-car [83], a Morse molecule [33], etc. Typically, these designs are optimized using semi-empirical quantum mechanics calculations, such as the semi-empirical ASED+ calculation technique [34].

While impressive, these are but two examples of the nanoscience revolution already impacting numerous fields, including electronics and semiconductors [64], textiles [63], [54], energy [69], food [44], drug delivery [52], [85], chemicals [55], materials [45], the automotive industry [31], aerospace and defense [51], medical devices and therapeutics [47], medical diagnostics [86], etc. According to some estimates, the world market for nanotechnology-related products and services will reach one trillion dollars by 2015 [78]. Nano-engineering groups are multiplying throughout the world, both in academia and in the industry: in the USA, the MIT has a "NanoEngineering" research group, Sandia National Laboratories created a "National Institute for Nano Engineering", to name a few; China founded a "National Center for Nano Engineering" in 2003, etc. Europe is also a significant force in public funding of nanoscience and nanotechnology and, in Europe, Grenoble and the Rhone-Alpes area gather numerous institutions and organizations related to nanoscience.

Of course, not all small systems that currently fall under the label "nano" have mechanical, electronic, optical properties similar to the examples given above. Furthermore, current construction capabilities lack behind some of the theoretical designs which have been proposed, such as the planetary gear designed by Eric Drexler at Nanorex. However, the trend is clearly for adding more and more functionality to nanosystems. While designing nanosystems is still very much an art mostly performed by physicists, chemists and biologists in labs throughout the world, there is absolutely no doubt that fundamental engineering practices will progressively emerge, and that these practices will be turned into quantitative rules and methods. Similar to what has happened with macroscopic engineering, powerful and generic software will then be employed to engineer complex nanosystems.

3.2. Challenges of practical nanosystem design

As with macrosystems, designing nanosystems will involve modeling and simulation within software applications: modeling, especially structural modeling, will be concerned with the creation of potentially complex chemical structures such as the examples above, using a graphical user interface, parsers, scripts, builders, etc.; simulation will be employed to predict some properties of the constructed models, including mechanical properties, electronic properties, chemical properties, etc.

In general, design may be considered as an "inverse simulation problem". Indeed, designed systems often need to be optimized so that their properties — predicted by simulation — satisfy specific objectives and constraints (e.g. a car should have a low drag coefficient, a drug should have a high affinity and selectivity to a target protein, a nano-wheel should roll when pushed, etc.). Being the main technique employed to predict properties, simulation is essential to the design process. At the nanoscale, simulation is even more important. Indeed, physics significantly constrains atomic structures (e.g. arbitrary inter-atomic distances cannot exist), so that a tentative atomic shape should be checked for plausibility much earlier in the design process (e.g. remove atomic clashes, prevent unrealistic, high-energy configurations, etc.). For nanosystems, thus, efficient simulation algorithms are required both when modeling structures and when predicting systems properties. Precisely, an effective software tool to design nanosystems should (a) allow for interactive physically-based modeling, where all user actions (e.g. displacing atoms, modifying the system's topology, etc.) are automatically followed by a few steps of energy minimization to help the user build plausible structures, even for large number of atoms, and (b) be able to predict systems properties, through a series of increasingly complex simulations.

3.3. Current simulation approaches

Even though the growing need for effective nanosystem design will still increase the demand for simulation, a lot of research has already gone into the development of efficient simulation algorithms. Typically, two approaches are used: (a) increasing the computational resources (use super-computers, computer clusters, grids, develop parallel computing approaches, etc.), or (b) simulating simplified physics and/or models. Even though the first strategy is sometimes favored, it is expensive and, it could be argued, inefficient: only a few supercomputers exist, not everyone is willing to share idle time from their personal computer, etc. Surely, we would see much less creativity in cars, planes, and manufactured objects all around if they had to be designed on one of these scarce super-resources.

The second strategy has received a lot of attention. Typical approaches to speed up molecular mechanics simulation include lattice simulations [89], removing some degrees of freedom (e.g. keeping torsion angles only [62], [84]), coarse-graining [87], [81], [35], [82], multiple time step methods [75], [76], fast multipole methods [48], parallelization [61], averaging [43], multi-scale modeling [42], [39], reactive force fields [41], [92], interactive multiplayer games for predicting protein structures [46], etc. Until recently, quantum mechanics methods, as well as mixed quantum / molecular mechanics methods were still extremely slow. One breakthrough has consisted in the discovery of linear-scaling, divide-and-conquer quantum mechanics methods [90], [91].

Overall, the computational community has already produced a variety of sophisticated simulation packages, for both classical and quantum simulation: ABINIT, AMBER, CHARMM, Desmond, GROMOS and GROMACS, LAMMPS, NAMD, ROSETTA, SIESTA, TINKER, VASP, YASARA, etc. Some of these tools are open source, while some others are available commercially, sometimes via integrating applications: Ascalaph Designer, BOSS, Discovery Studio, Materials Studio, Maestro, MedeA, MOE, NanoEngineer-1, Spartan, etc. Other tools are mostly concerned with visualization, but may sometimes be connected to simulation packages: Avogadro, PyMol, VMD, Zodiac, etc. The nanoHUB network also includes a rich set of tools related to computational nanoscience.

To the best of our knowledge, however, all methods which attempt to speed up dynamics simulations perform a priori simplification assumptions, which might bias the study of the simulated phenomenon. A few recent, interesting approaches have managed to combine several levels of description (e.g. atomistic and coarse-grained) into a single simulation, and have molecules switch between levels during simulation, including the adaptive resolution method [71], [72], [73], [74], the adaptive multiscale method [68], and the adaptive partitioning of the Lagrangian method [56]. Although these approaches have demonstrated some convincing applications, they all suffer from a number of limitations stemming from the fact that they are either ad hoc methods tuned to fix specific problems (e.g. fix density problems in regions where the level of description changes), or mathematically founded methods that necessitate to "calibrate" potentials so that they can be mixed (i.e. all potentials have to agree on a reference point). In general, multi-scale methods, even when they do not allow molecules to switch between levels of detail during simulation, have to solve the problem of rigorously combining multiple levels of description (i.e. preserve statistics, etc.), of assigning appropriate levels to different parts of the simulated system ("simplify as much as possible, but not too much"), and of determining computable mappings between levels of description (especially, adding back detail when going from coarse-grained descriptions to fine-grained descriptions).

3.4. As-Rigid-As-Possible methods for molecular paths

Last year, in the scope of Minh Khoa Nguyen's PhD, we have adapted the As-Rigid-As-Possible (ARAP) paradigm used in Computer Graphics to generate paths of molecular systems. This year, we continued this line of research with new extensions of the ARAP methodology. One extension led to generate conformational transition paths with low potential-energy barriers for proteins. It was published to the Journal of Computer-Aided Molecular Design, 2018 [66]. Another extension concerned the ART-RRT method which incorporates the ARAP methodology inside tree-based exploration methods to approximate ligand unbinding pathways. This contribution was published to the Journal of Computational Chemistry, 2018 [65]. Finally, the PhD thesis of Minh Khoa Nguyen titled *Efficient exploration of molecular paths from As-Rigid-As-Possible approaches and motion planning methods* was defended in March 2018 [67]. A brief summary of the above-mentioned contributions is presented below.

3.5. Modelling and simulation for the characterization of advanced materials

We have continued our informal collaboration with the *service de Caractérisation des Matériaux et Composants* of CEA, LETI, Minatec initiated in 2017. The collaboration with LETI will offer numerous possibilities of very precise (sub-nanometric) experimental comparisons based on the High-resolution scanning transmission electron microscopy (HRSTEM) using one of the best microscopes on the market (the FEI Titan Ultimate microscope). In this context, we have developed a set of tools tools to manipulate, simulate and measure nanomaterials, with a special focus on crystals that appears in many new materials such as semiconductors. The description of the sublines of research explored so far are detailed below.

4. Highlights of the Year

4.1. Highlights of the Year

- This year we have very successfully participated in the blind assessment of protein structure prediction methods exercise CASP13. We have evaluated the performance of several knowledge-based potentials for protein model quality and protein docking, small-angle scattering approaches Pepsi-SAXS and Pepsi-SANS, cross-linking developments, methods based on normal mode analysis and more. Our team was ranked 1st in three data-assisted CASP13 sub-challenges (SAXS, SANS, and crosslinks), and got into the top-10 predictors in the main category of the prediction of regular targets. We were also interviewed on this subject by the Le Figaro newspaper [88].
- The OneAngström startup was created this year around the development of the SAMSON software platform. Four team members have joined the startup : Stephane Redon, Jocelyn Gate, Dmitriy Marin, and Yassine Naimi.
- Our Ananas analytical symmetry detection method [70] was used in the official assessment of protein assemblies in CASP13 and was also transferred to the PDBe European resource for the collection, organisation and dissemination of data on biological macromolecular structures [30].

4.1.1. Awards

- Our paper "Analytical symmetry detection in protein assemblies. II. Dihedral and cubic symmetries" covered the September 2018 issue of the Journal of Structural Biology [20].
- Our paper "A novel fast Fourier transform accelerated off-grid exhaustive search method for cryo-electron microscopy fitting" covered the the August 2017 issue of Journal of Applied Crystallography [58].
- Our paper "NOLB: Nonlinear Rigid Block Normal Mode Analysis Method" covered May 2017 issue of Journal of Chemical Theory and Computation [57].
- Our predictions were ranked 1st in the SAXS-assisted category of the CASP13 protein structure prediction challenge (cumulative SAXS-assisted z-scores).
- Our predictions were ranked 1st in the SANS-assisted category of the CASP13 protein structure prediction challenge (cumulative SANS-assisted z-scores).
- Our predictions were ranked 1st in the X-link-assisted category of the CASP13 protein structure prediction challenge (cumulative X-link-assisted z-scores).

5. New Software and Platforms

5.1. SAMSON

Software for Adaptive Modeling and Simulation Of Nanosystems
KEYWORDS: Bioinformatics - Simulation - Nanosystems - Structural Biology - Chemistry
SCIENTIFIC DESCRIPTION: Please refer to https://www.samson-connect.net
FUNCTIONAL DESCRIPTION: SAMSON is a software platform for real-time modelling and simulation of natural or artificial nanosystems. The objective is to make SAMSON a generic application for computer-aided design of nanosystems, similar to existing applications for macrosystem prototyping (CATIA, SolidWorks, etc.).

- Contact: Stéphane Redon
- URL: http://nano-d.inrialpes.fr/software/

5.2. DockTrina

A novel protein docking method for modeling the 3D structures of nonsymmetrical triangular trimers
FUNCTIONAL DESCRIPTION: DockTrina is a novel protein docking method for modeling the 3D structures of nonsymmetrical triangular trimers. The method takes as input pair-wise contact predictions from a rigid body docking program. It then scans and scores all possible combinations of pairs of monomers using a very fast root mean square deviation (RMSD) test (see below). Finally, it ranks the predictions using a scoring function which combines triples of pair-wise contact terms and a geometric clash penalty term. The overall approach takes less than 2 min per complex on a modern desktop computer.

- Contact: Sergey Grudinin
- URL: https://team.inria.fr/nano-d/software/docktrina/

5.3. HermiteFit

A new docking algorithm for rapid fitting atomic structures into cryo-EM density maps
FUNCTIONAL DESCRIPTION: HermiteFit is a new docking algorithm for rapid fitting atomic structures into cryo-EM density maps using 3D orthogonal Hermite functions. HermiteFit uses the cross-correlation or the Laplacian-filtered cross-correlation as the fitting criterion. HermiteFit exhaustively rotates the protein density in the Hermite space and then converts the expansion coefficients into the Fourier space for the subsequent fast FFT-based correlation computations.

- Partners: IBS - FZJ Juelich
- Contact: Sergey Grudinin
- URL: https://team.inria.fr/nano-d/software/hermitefit/

5.4. Knodle

KNOwledge-Driven Ligand Extractor
KEYWORDS: Bioinformatics - Machine learning
FUNCTIONAL DESCRIPTION: KNOwledge-Driven Ligand Extractor is a software library for the recognition of atomic types, their hybridization states and bond orders in the structures of small molecules. Its prediction model is based on nonlinear Support Vector Machines. The process of bond and atom properties perception is divided into several steps. At the beginning, only information about the coordinates and elements for each atom is available :

Connectivity is recognized. A search of rings is performed to find the Smallest Set of Smallest Rings (SSSR). Atomic hybridizations are predicted by the corresponding SVM model. Bond orders are predicted by the corresponding SVM model. Aromatic cycles are found. Atomic types are set in obedience to the functional groups. Some bonds are reassigned during this stage.

- Participants: Maria Kadukova and Sergey Grudinin
- Partner: MIPT Moscow
- Contact: Sergey Grudinin
- Publication: Knodle: A Support Vector Machines-Based Automatic Perception of Organic Molecules from 3D Coordinates
- URL: https://team.inria.fr/nano-d/software/Knodle/

5.5. RigidRMSD

A library for rapid computations of the root mean square deviations (RMSDs) corresponding to a set of rigid body transformations of a coordinate vector
KEYWORD: Bioinformatics

FUNCTIONAL DESCRIPTION: RigidRMSD is a library for rapid computations of the root mean square deviations (RMSDs) corresponding to a set of rigid body transformations of a coordinate vector (which can be a molecule in PDB format, for example). Calculation of the RMSD splits into two steps:

Initialization, which is linear in the number of vector entities (or particles in a rigid body). RMSD computation, which is computed in constant time for a single rigid-body spatial transformation (rotation + translation). This step uses the inertia tensor and the the center of mass computed on the first step. Initialization step is performed only once. It makes RigidRMSD particularly useful when computing multiple RMSDs, since each new RMSD calculation takes only constant time.

- Participants: Petr Popov and Sergey Grudinin
- Contact: Sergey Grudinin
- Publication: Rapid determination of RMSDs corresponding to macromolecular rigid body motions
- URL: https://team.inria.fr/nano-d/software/rigidrmsd/

5.6. SAMSON-Drug-design

KEYWORDS: Algorithm - Nanosystems - Structural Biology - Bioinformatics - Chemistry - 3D modeling - Molecular simulation
FUNCTIONAL DESCRIPTION: Arap Interpolation Path : Generate interpolation path between two protein structures by the As-Rigid-As-Possible principle from computer graphics

Ligand unbinding search : Find ligand unbinding pathway with the ART-RRT method. The method uses the T-RRT method from robotics for efficiently searching low-energy paths and the ARAP modeling method from computer graphics for handling flexible motions of the ligand and reducing the number of thedimensions of the search space.

Protein Path search : Find protein conformational transition paths between two given conformations with the ART-RRT method. The method uses the T-RRT method from robotics for searching low-energy paths and the As-Rigid-As-Possible (ARAP) methods from computer graphics for handling the flexibility of the protein and reducing the number of the dimensions of the search space.

- Authors: Leonard Jaillet, Minh Khoa Nguyen and Jocelyn Gaté
- Partner: Inria
- Contact: Stéphane Redon
- URL: http://samson-connect.net

5.7. DeepSymmetry

KEYWORDS: Bioinformatics - 3D modeling - Machine learning - Neural networks
FUNCTIONAL DESCRIPTION: DeepSymmetry is a method based on three-dimensional (3D) convolutional networks that detects structural repetitions in proteins and their density maps. It identifies tandem repeat proteins, proteins with internal symmetries, their symmetry order, and also the corresponding symmetry axes.

- Participants: Guillaume Pages and Sergey Grudinin
- Contact: Sergey Grudinin
- Publication: DeepSymmetry : Using 3D convolutional networks for identification of tandem repeats and internal symmetries in protein structures
- URL: https://team.inria.fr/nano-d/software/deepsymmetry/

5.8. SAMSON-ARAP-Planner

KEYWORDS: 3D - Algorithm - Nanosystems - Bioinformatics - Structural Biology - Chemistry

FUNCTIONAL DESCRIPTION: ARAP planner combines the ARAP method from computer graphics with T-RRT exploration method from robotics for efficiently finding low-energy paths in high-dimensional energy landscapes.

- Authors: Leonard Jaillet and Minh Khoa Nguyen
- Partner: Inria
- Contact: Stéphane Redon
- URL: http://samson-connect.net

5.9. SAMSON-Hydrocarbons

KEYWORDS: Algorithm - Quantum chemistry - Chemistry - Nanosystems - 3D - 3D modeling
FUNCTIONAL DESCRIPTION: Interactive quantum chemistry : This SAMSON Element demonstrates interactive quantum chemistry for small molecules at the ASED-MO level of theory. Choose the ASED-MO (atom superposition and electron delocalization) interaction model when adding a simulator through the 'Simulation' menu. The SAMSON Element also includes an App that makes it possible to visualize how the electron density evolves during interactive simulation.

Brenner interction model : This SAMSON Element contains an adaptive implementation of the Brenner interaction model. Interaction models are one of the five model categories that are used to model nanosystems in SAMSON, along with structural models (for geometry and topology), dynamical models (to represent degrees of freedom), visual models (for visual representations) and property models (to represent properties). The Brenner interaction model is a reactive bond-order potential for hydrocarbon systems. This adaptive implementation makes it possible to interactively simulate large systems. Choose this interaction model when adding a simulator through the 'Simulation' menu.

- Author: Maël Bosson
- Partner: Inria
- Contact: Stéphane Redon
- URL: http://samson-connect.net

5.10. SAMSON-RDKit

KEYWORDS: 2D - 3D - Chemistry - Algorithm - 3D modeling - Structural Biology - Bioinformatics
FUNCTIONAL DESCRIPTION: Based on the RDKit open-source libraries, convert and manage SMILES codes in the SAMSON platform. The RDKit-SMILES Manager element allows you to easily import files (.smi or .txt) containing several SMILES codes or add each code separately. 2D conformation of each code will then be generated and you can save them into svg or png files. Using a checkbox you will be able to select the codes that you want to convert into 3D structures and add them directly into the SAMSON data graph node as structural model. For more information about using this SAMSON Element, please visit https://documentation.samson-connect.net/using-the-rdkit-smiles-manager/.

- Author: Yassine Naimi
- Partner: Inria
- Contact: Stéphane Redon
- URL: https://samson-connect.net/app/main?key=element&uuid=ce09650a-c071-4e84-1f6a-b87706937d5c1

5.11. SAMSON-GROMACS

KEYWORDS: Algorithm - Materials - Chemistry - Bioinformatics - Structural Biology - Nanosystems - 3D modeling - 3-order

FUNCTIONAL DESCRIPTION: This SAMSON Element wraps GROMACS 5.1 force fields and setup tools. Use the "GROMACS setup" app (in the App menu), which wraps the pdb2gmx tool, to generate a structural model suitable for simulation (i.e. add hydrogens, etc.). Then, apply a simulator from the Simulation menu and choose "GROMACS force field" to add a GROMACS interaction model suitable for interactive minimization and simulation (no periodic boundary conditions). Note that, at the moment, at most one structural model should be selected (or in the document, when the selection is empty), and that bond lengths are not yet constrained in this version. This may be combined with the Twister editor to perform large-scale modifications of the structure, and the secondary structure visual model for interactively updated secondary structure prediction. Future updates of this SAMSON Element will wrap more GROMACS tools. Source code for this SAMSON Element will be made available at https://gforge.inria.fr/projects/elements/.

- Authors: Stéphane Redon, Minh Khoa Nguyen and Yassine Naimi
- Partner: Inria
- Contact: Stéphane Redon
- URL: http://samson-connect.net

5.12. SAMSON-Essentials

KEYWORDS: 3D - C++ - OpenGL - Molecular surface - Molecular simulation - Structural Biology - Chemistry - 3D modeling - Bioinformatics - Nanosystems
FUNCTIONAL DESCRIPTION: A set of SAMSON Elements that adds essential features to SAMSON such as import / export of models, import / export of documents, generators, simulators, editors, scripting, app as well as software integrations (autodock vina).

- Authors: Stéphane Redon, Jocelyn Gaté, Guillaume Pages, Dmitriy Marin, Svetlana Artemova, Himani Singhal, Marc Aubert, Marc Piuzzi and Clement Beitone
- Partner: Inria
- Contact: Stéphane Redon
- URL: http://samson-connect.net

5.13. Samson-base

KEYWORDS: Bioinformatics - Simulation - Nanosystems - Structural Biology - Chemistry

- Participants: Evelyne Altariba, Jocelyn Gaté, Noëlle Le Delliou and Stéphane Redon
- Contact: Stéphane Redon

5.14. SAMSON-Connect

KEYWORDS: Web Application - Software platform - Web
FUNCTIONAL DESCRIPTION: SAMSON, SAMSON Elements and the SAMSON Software Development Kit are distributed via the SAMSON Connect website.[2] The site acts as a repository for the SAMSON Elements being uploaded by developers, and users of SAMSON choose and add Elements from SAMSON Connect.

- Authors: Stéphane Redon, Mohamed Yengui and Jocelyn Gaté
- Partner: Inria
- Contact: Stéphane Redon
- URL: http://samson-connect.net

5.15. SAMSON-Updater

KEYWORDS: Webservices - Web Application

FUNCTIONAL DESCRIPTION: Web service to ensure communication between SAMSON and SAMSON-Connect. Features: Add / remove new items to SAMSON if they have been added / deleted on SAMON-Connect Update SAMSON or SAMSON-Elements Authenticate users ...

- Authors: Mohamed Yengui and Jocelyn Gaté
- Partner: Inria
- Contact: Stéphane Redon

5.16. SAMSON-Various-tools

SAMSON-Various-toolsrere

KEYWORDS: 3D - 3D modeling - Chemistry - Algorithm - Bioinformatics - Nanosystems - Structural Biology

FUNCTIONAL DESCRIPTION: Cluster Game : This element has been made in order to help students discover the Lennard Jones interactions. It is a game in which the goal is to optimize the atoms' placement. Contains eight levels and five tutorial levels.

Atoms Selector : This SAMSON element allows selection of atoms in the active document according to a user-provided expression with the usage of NSL-like variables (NSL - Node Specification Language), standard mathematical and logical operations. Parsing and evaluation of expressions is done with the usage of C++ Mathematical Expression Parsing And Evaluation Library 'exprtk' by Arash Partow (https://github.com/ArashPartow/exprtk)

Simple Script : This SAMSON element allows modification of some parameters of atoms using scripting language, standard mathematical and logical operations, and NSL-like variables (NSL - Node Specification Language). The script is applied to each atom independently. Parsing and evaluation of the script is done with the usage of C++ Mathematical Expression Parsing And Evaluation Library 'exprtk' by Arash Partow (https://github.com/ArashPartow/exprtk)

Bond Angle Distribution : This App will compute the bond angle distribution of the selected atoms. The result and its image can then be exported.

Bond Distortion Visualisation : This App permits to visualize with colors the distortion of a molecular or crystalic structure. The angle distortion, the bond distortion and the projected bond distortion can be represented with colors on bonds and atoms.

Frame axis : Basic visual model to show cartesian axis. Open a new visual model and select "Frame axis" to see the frame axis as arrows, lines or both.

Radial Distribution Function : This app computes and draws the radial distribution function of a selection. - Compute the crossed-RDF by selecting 2 different sets of atoms. - Follow the evolution of the RDF by selecting a simulator.

Adaptive Lennard-Jones : An interaction model to compute the forces with an adaptive version of Lennard-Jones potential. The update of forces is done by storing all the position, and at each position update, subtracting previous pair forces, then adding new ones. If both pair particles were frozen by restraining dynamical model, the update is useless and so not done.

STL File Importer : Reads Stereolithography (.stl) binary and ASCII files. Spawns carbon atoms at the intersections of vertices to create quickly new original atomic configurations. Many STL files are avalaible online to generate thousands of new configurations.

StyleSheet Viewer : An app for internal developers to test their skin/styles ...

Animation Player : Create animation from a list of conformations in SAMSON Document view There are 3 play modes: loop, only 1 time, or continuous back-forth. The user can change the frame order by drag and drop in the frame list.

Internal Coordinate Editor : This editor rotate the molecule by defining rotation axis and rotation angle. The user define the rotation axis (represented by an arrow in the view) by clicking on one atom and drag it to the second atom. The rotation angle is defined by the small GUI windows. The editor will try to rotate all the atoms after the head of the arrow and in between the tail and head of the arrow

RMSD : This app calculates RMSD between 2 structures. If it is an amino acid chains, a sequence alignment of the structures is needed (in fasta format). The fasta format has to be obtained from external sources. If it is not a protein, the app will try to match atoms one by one in both structures.

Trajectory Importer : This app import selected pdb files (hold Ctrl + Mouse click for multiple selection) as a trajectory. The result is one single structural model and a list of conformations in SAMSON Document View

Open Babel connector : This app allows users to use Open Babel from inside SAMSON

ARPS demo : This SAMSON Element features a demonstration of ARPS: Adaptively Restrained Particle Simulations, an adaptive simulation technique able to focus computations on the most mobile degrees of freedom. In this demo, a collision cascade may be simulated with various degrees of precision by changing the restrained dynamics threshold and the full dynamics threshold. For example, 0 0 produces a classical, non-adaptive simulation, while 0.625 and 0.7 result in a 10 times speedup in this example (without the graphics overhead). Click on 2D Shock to generate the example, enter the simplification parameters, and press start to simulate the collision cascade. Undo and redo make it possible to zoom on and compare simulations. Please refer to "S. Artemova and S. Redon. Adaptively restrained particle simulations. Physical review letters, 2012" for more details.

Hydrogen bond finder : Find hydrogen bonds in a given structure. The bonds is detected by specifying a threshold distance. Bonds are displayed by yellow lines.

Lennard-Jones model : This SAMSON Element contains a Lennard-Jones interaction model that may be used for several purposes: teaching, learning about van der Waals interactions, developing optimization algorithms, looking for minimum energy Lennard-Jones clusters through interactive simulation, etc. Add this force field to a group of atoms (the atom types do not matter) via the 'Simulation' menu.

Catalogue : BETA version. Bunch of structures easy to load, with images.

Catalogue Generator : BETA. Generator for the Catalogue module.

SAMSON Basic Tutorial : Tutorial for SAMSON : Basic functionnalities. Learn how to create, move and delete an atom, and then create a basic molecule thanks to a step-by-step guide. Experiment in sandbox mode. Test your mastery of SAMSON's tools with 4 speed challenges. The following elements are required : SAMSON Editors, Periodic table, Basic importers. Currently only available in french.

SAMSON Courses : BETA. Module with a custom display, depending on the entry files

SAMSON Courses Creator : BETA. Generator for SAMSON Course module

Charts : A SAMSON Element to allow user plotting something from SAMSON datas interactively

Leap : A driver to control SAMSON with the Leap motion controller

- Authors: Mohamed Nadhir Ben Hadj Abdellatif, Svetlana Artemova, Clement Beitone, Jocelyn Gaté, Dmitriy Marin, Pierre Mehaye, Minh Khoa Nguyen, Guillaume Pages, Stéphane Redon, François Rousse and Joachim Woerly-Moussier
- Partner: Inria
- Contact: Stéphane Redon

5.17. SAMSON-Crystal-Study-Pack

KEYWORDS: Algorithm - Nanosystems - 3D - 3D modeling - Physical simulation
FUNCTIONAL DESCRIPTION: CrystalConstrainer : Constrains borders of a crystal. The crystal should be alignes on xyz axes. The margen defines the maximum distance to the bonding box border for which atoms are constrained. Can use the current positions as constrained pos or some fixed one. Can Constrain each pair of plans separately or by couples (e.g. XY)

CrystalProber : Get some statistics about a crystal. Compute for now the X, Y and Z lattice parameters. Can remove a margen to discard atoms at the border

CrystalRigidityProber : Analyze some crystal properties related to the rigidity and based on forces, such as the Young modulus, Poisson's ration, elastic constants, stiffness, etc.

Keating : Implement the Keating force field. Compute energy and forces according to positions. Follows anikin2011keating

nonharmonic Keating : Develop a nonharmonic Keating model as proposed in Rucker1995anharmonic. As parameter file it uses a .nhk extension. As parameters, it require a nu and theta0 value in addition to the equilibrium distances for each atom type. Also, it requires a0 distances for pairwise atoms, in addition to the alpha and beta parameters.

generalized harmonic Keating : Develop a generalized harmonic Keating force field as proposed in mojica2010modelisation. This model generalizes to atoms of columns III and V of the periodic table. The systems are anisotropics, with specific bond lengths and angles when involving the z direction.

CrystalCharacterizer : Provides functionalities to characterise a crystal.

CrystalVisualizer : Provides functionalities to visualize a crystal. Choose the pointing direction of the eye, according to some representative directions of the crystal mesh. Choose the orientation in the camera planer.

- Author: Leonard Jaillet
- Partner: Inria
- Contact: Stéphane Redon
- URL: http://samson-connect.net

5.18. SAMSON-Materials

KEYWORDS: Algorithm - Materials - 3D - 3D modeling - Nanosystems
FUNCTIONAL DESCRIPTION: Crystal Creator : This SAMSON Elements enables to generate crystals. It contains a SAMSON App to write a unit cell and a SAMSON importer to read CIF format files. Once a unit cell is written or imported, it can be repeated in the directions of the lattice vectors to create a whole crystal. Each repetition is not a mere copy but is generated again so that the defects and the impurities are modeled. A functionality of the associated property model permits to cut the crystal with the Miller indices and expose the important crystalic planes.

Orbital Free DFT : This App computes the electron density of an atomic system. It comes with an interaction model to minimize the atomic structure and a visual model to appreciate the result of computations. The scheme used is the orbital-free DFT, and the pseudo-potential available restrains its use to only 9 elements : Li, Mg, Al, Si, P, Ga, In and Sb.

- Author: François Rousse
- Partner: Inria
- Contact: Stéphane Redon
- URL: http://samson-connect.net

5.19. SAMSON-Planner-Tools

KEYWORDS: 3D - Algorithm - 3D modeling - Molecular simulation - Chemistry - Planning
FUNCTIONAL DESCRIPTION: A set of SAMSON Elements that adds planning features to SAMSON.

- Author: Leonard Jaillet
- Partner: Inria
- Contact: Stéphane Redon
- URL: http://samson-connect.net

5.20. SAMSON-HEX

KEYWORD: Algorithm

FUNCTIONAL DESCRIPTION: Sampling and Docking using the Hex algorithm developed by Dave Ritchie in SAMSON. Docking solutions can be easily displayed and clustered.

- Authors: Sergey Grudinin, Emilie Neveu and Yassine Naimi
- Partner: Inria
- Contact: Stéphane Redon
- URL: https://www.samson-connect.net/

5.21. SAMSON-UFF

KEYWORDS: Algorithm - Bioinformatics - 3D - 3D modeling - Nanosystems - Molecular simulation

FUNCTIONAL DESCRIPTION: Universal Force Field : This SAMSON Element contains a new implementation of the Universal Force Field, with automatic structure perception. In order to use this interaction model, add a simulator to the document from the Simulation menu, and choose "Universal Force Field". The property window of the interaction model makes it possible to customize the perception and setup the interaction model (e.g. choose the cutoff), and displays the various energy types and the total energy.

Interactive Modeling Universal Force Field : It is an extension of UFF that combines the possibility to significantly modify molecular structures (as with reactive force fields) with a broad diversity of supported systems thanks to the universality of UFF. Such an extension lets the user easily build and edit molecular systems interactively while being guided by physics based inter-atomic forces.

- Authors: Leonard Jaillet and Svetlana Artemova
- Partner: Inria
- Contact: Stéphane Redon
- URL: https://samson-connect.net/app/main?key=element&uuid=8cbdc8b1-59e1-6459-d68f-b840275dd5e9

5.22. SAMSON-Normal-modes

KEYWORDS: Algorithm - Bioinformatics - Structural Biology - 3-order - 3D modeling - Nanosystems

FUNCTIONAL DESCRIPTION: Normal mode analysis advanced : This SAMSON Element computes the nonlinear normal modes of a molecular system (protein, RNA, DNA) very quickly using the NOLB algorithm developed by Alexandre Hoffmann and Sergei Grudinin (J. Chem. Theory Comput., 2017, 13 (5), pp 2123-2134, DOI: 10.1021 / acs.jctc.7b00197.). The user indicates the desired number of modes, the interactions cutoff distance and the potential function. For now, the elastic network model potential is the one that is available. In the output, each mode is represented by a slider. The user can visualize the motion of each mode independently by moving its corresponding slider manually or by checking its checkbox and then pressing on the play button. Also, the user can visualize the motion of a combination of modes selecting them before playing the motion. The transformations used in this motion can be set to linear or nonlinear and the amplitude of the motion can be increased/decreased by changing the scaling factor. During this motion, the user can activate a real time minimization using one of the provided algorithms (steepest descent, conjugated gradient or LBGF) and defined values of minimization steps and minimization tolerance. Finally, the user can either save/export a given conformation of the structure or the entire displayed trajectory by going into the "Save Frames" tabulation of the SAMSON element. Please visit https://blog.samson-connect.net/computing-non-linear-normal-modes-of-biomolecules/ for a tutorial.

Normal mode analysis : A light version of the previous element

- Authors: Sergey Grudinin, Alexandre Hoffmann and Yassine Naimi
- Partner: Inria
- Contact: Sergey Grudinin
- URL: http://samson-connect.net

5.23. SAMSON-SAXS

KEYWORDS: Algorithm - Bioinformatics - Structural Biology - Nanosystems - 3D - 3D modeling
FUNCTIONAL DESCRIPTION: The Pepsi-SAXS module rapidly computes small-angle X-ray scattering profiles. It is based on the spherical harmonics expansion method and interactively updates the fits if the molecular structure is modified. Multi-threading is currently only supported for Linux & macOS.

- Authors: Sergey Grudinin, Mariya Garkavenko and Mohamed Nadhir Ben Hadj Abdellatif
- Partner: Inria
- Contact: Sergey Grudinin
- URL: https://samson-connect.net/app/main?key=element&uuid=844be03b-cab2-4420-464b-6f0f9384bc4a

5.24. Ananas

Analytical Analyzer of Symmetries
KEYWORDS: Bioinformatics - Structural Biology
FUNCTIONAL DESCRIPTION: Analytical Analyzer of Symmetries is a software for detection and assessment of the quality of symmetry in a protein assembly.

This software can : Detect the best axes of symmetry for any symmetry group in an assembly containing the right amount of chains, Provide the symmetry-aware RMSD for these axes, Detect the best axis of symmetry for cyclic assemblies with missing subunits, Compute the axes of symmetry with user-provided correspondences.

- Participants: Guillaume Pages and Sergey Grudinin
- Contact: Sergey Grudinin
- Publications: Analytical symmetry detection in protein assemblies. I. Cyclic symmetries - Analytical symmetry detection in protein assemblies. II. Dihedral and Cubic symmetries
- URL: https://team.inria.fr/nano-d/software/ananas/

5.25. Pepsi-SAXS

KEYWORDS: Bioinformatics - Structural Biology - Data modeling
FUNCTIONAL DESCRIPTION: Pepsi-SAXS (PEPSI stands for Polynomial Expansions of Protein Structures and Interactions) is new implementation of the multipole-based scheme initially proposed by Stuhrmann (Stuhrmann, 1970). Overall, our method is significantly faster with a similar accuracy compared to Crysol, FoXS, and the 3D-Zernike implementation from the SAStbx package.

- Participant: Sergey Grudinin
- Partner: MIPT Moscow
- Contact: Sergey Grudinin
- Publication: Pepsi-SAXS : an adaptive method for rapid and accurate computation of small-angle X-ray scattering profiles
- URL: https://team.inria.fr/nano-d/software/pepsi-saxs/

5.26. NOLB

NOn-Linear rigid Block NMA method
KEYWORDS: Structural Biology - Bioinformatics - Elasticity - Proteins - Motion analysis

FUNCTIONAL DESCRIPTION: It's a new conceptually simple and computationally efficient method for non-linear normal mode analysis of macromolecules.

- Participants: Sergey Grudinin and Alexandre Hoffmann
- Contact: Sergey Grudinin
- Publications: NOLB: Nonlinear Rigid Block Normal Mode Analysis Method - RapidRMSD: Rapid determination of RMSDs corresponding to motions of flexible molecules
- URL: https://team.inria.fr/nano-d/software/nolb-normal-modes/

5.27. SBROD

KEYWORDS: Bioinformatics - Machine learning
FUNCTIONAL DESCRIPTION: Smooth orientation-dependent scoring function (SBROD) for coarse-grained protein quality assessment uses only the conformation of the protein backbone, and hence it can be applied to scoring the coarse-grained protein models.

The workflow of SBROD consists in two stages. First, the method extracts features from each protein model in the dataset. Then, the scoring function assigns a score to each processed protein model depending on its features extracted at the first stage. Figure above schematically shows the workflow of SBROD. Here, four types of inter-atomic interactions, described in details below, are taken into account when extracting the features. After these features have been extracted and preprocessed, a Ridge Regression model is trained on them to predict the GDT-TS of protein models.

- Participants: Mikhail Karasikov, Guillaume Pages and Sergey Grudinin
- Contact: Sergey Grudinin
- Publication: Smooth orientation-dependent scoring function for coarse-grained protein quality assessment
- URL: https://team.inria.fr/nano-d/software/sbrod/

5.28. Ornate

KEYWORDS: Bioinformatics - Machine learning - Neural networks
FUNCTIONAL DESCRIPTION: Oriented Routed Neural network with Automatic Typing is a method for protein quality assessment. Ornate is a residue-wise scoring method. It first constructs a three dimensional map representing the structure of the residue, and its neighborhood.

- Participants: Guillaume Pages, BENOIT CHARMETTANT and Sergey Grudinin
- Contact: Sergey Grudinin
- Publication: Protein model quality assessment using 3D oriented convolutional neural networks
- URL: https://team.inria.fr/nano-d/software/ornate/

5.29. SAMSON-AR-LAMMPS

KEYWORDS: Algorithm - Molecular simulation - 3D modeling

- Authors: Sémého Edorh, Krishna Kant Singh and Dmitriy Marin
- Partner: Inria
- Contact: Stéphane Redon
- URL: http://samson-connect.net

6. New Results

6.1. Generating conformational transition paths with low potential-energy barriers for proteins

Participants: Minh Khoa Nguyen, Léonard Jaillet and Stéphane Redon.
Publication: Journal of Computer-Aided Molecular Design, 2018 [66].

The knowledge of conformational transition paths in proteins can be useful for understanding protein mechanisms. Recently, we have introduced the As-Rigid-As-Possible (ARAP) interpolation method, for generating interpolation paths between two protein conformations. The method was shown to preserve well the rigidity of the initial conformation along the path. However, because the method is totally geometry-based, the generated paths may be inconsistent because the atom interactions are ignored. Therefore, we introduce a new method to generate conformational transition paths with low potential-energy barriers for proteins. The method is composed of three processing stages. First, ARAP interpolation is used for generating an initial path. Then, the path conformations are enhanced by a clash remover. Finally, Nudged Elastic Band, a path-optimization method, is used to produce a low-energy path. Large energy reductions are found in the paths obtained from the method than in those obtained from the ARAP interpolation method alone. The results also show that ARAP interpolation is a good candidate for generating an initial path because it leads to lower potential-energy paths than two other common methods for path interpolation (see Figure 1 for an example of optimized transition path).

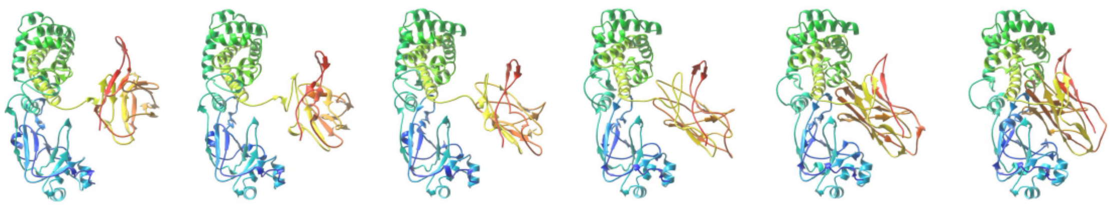


Figure 1. The path for diphtheria toxin after ARAP interpolation and NEB optimization.

6.2. ART-RRT: As-Rigid-As-Possible Exploration of Ligand Unbinding Pathways

Participants: Minh Khoa Nguyen, Leonard Jaillet, Stephane Redon.
Publication: Journal of Computational Chemistry, 2018 [65].

We have proposes a method to efficiently generate approximate ligand unbinding pathways. It combines an efficient tree-based exploration method with a morphing technique from Computer Graphics for dimensionality reduction. This method is computationally cheap and, unlike many existing approaches, does not require a reaction coordinate to guide the search. It can be used for finding pathways with known or unknown directions beforehand. The approach is evaluated on several benchmarks and the obtained solutions are compared with the results from other state-of-the-art approaches. We show that the method is time-efficient and produces pathways in good agreement with other state-of-the-art solutions. These paths can serve as first approximations that can be used, analyzed, or improved with more specialized methods (see Figure 2).

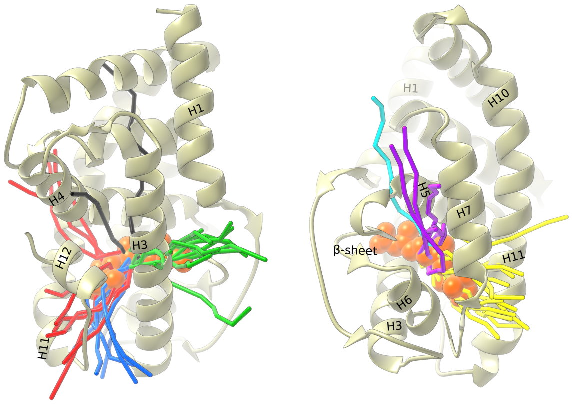


Figure 2. Families of paths (in colored sticks) obtained with ART-RRT for the unbinding of retinoic acid hormone from its receptor. The protein is represented by ribbons and the ligand by orange balls.

6.3. Atomistic modelling and simulation of transmission electron microscopy images: application to intrinsic defects of graphene

Participants: Cyril Guedj, Léonard Jaillet, François Rousse and Stéphane Redon.
Publication: Proceedings of 8th International Conference on Simulation and Modeling Methodologies, Technologies and Applications - Volume 1: SIMULTECH [53].
The characterization of advanced materials and devices in the nanometer range requires complex tools, and the data analysis at the atomic level is required to understand the precise links between structure and properties. We have demonstrated that the atomic-scale modelling of graphene-based defects may be performed efficiently for various structural arrangements using the Brenner module of the SAMSON software platform (cf Figure 3). The signatures of all kinds of defects are computed in terms of energy and scanning transmission electron microscopy simulated images. The results are in good agreement with all theoretical and experimental data available. This original methodology is an excellent compromise between the speed and the precision required by the semiconductor industry and opens the possibility of realistic in-silico research conjugated to experimental nanocharacterisation of these promising materials.

6.4. Impact of hydrogen on graphene-based materials: atomistic modeling and simulation of HRSTEM images.

Participants: Cyril Guedj, Léonard Jaillet, François Rousse and Stéphane Redon.
Oral presentation: AVS 65th International Symposium and Exhibition.
Summary: The hydrogen energy transition is highly probable, because hydrogen is the most abundant element in the universe and represents an ideal "green" source of energy. Meanwhile, the safe hydrogen production and storage remains a major challenge still in progress. To understand and optimize the device efficiency and the interface engineering, it is advantageous to perform advanced nanocharacterizations, linked to numerical modelling and simulations. This task is particularly difficult, because hydrogen is labile and prone to rapid reorganization. This structural evolution may be monitored with transmission electron microscopy (TEM) techniques, but in spite of significant progresses, the direct detection of hydrogen with High Resolution Scanning Transmission Electron Microscopy (HRSTEM) or energy-loss spectroscopy still remains a serious challenge.

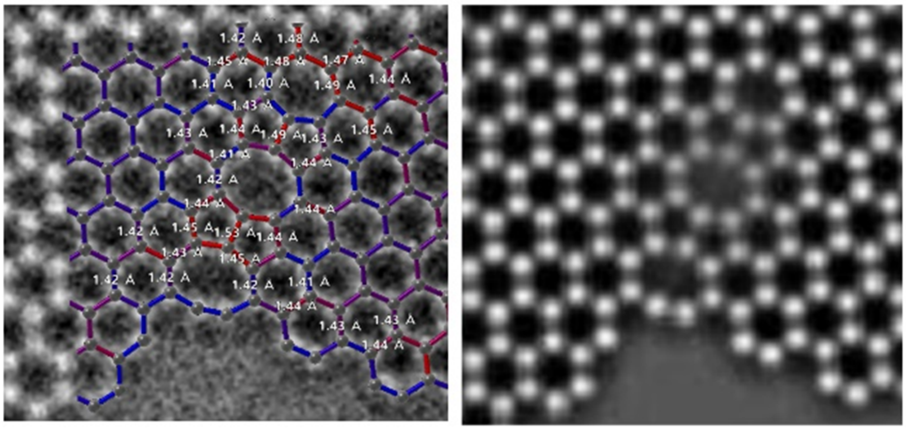


Figure 3. Left: atomistic model of the extended defect 88-7-5555 defect superimposed to the experimental HRTEM image entitled "SALVE-III-project-HRTEM-graphene-vacancy-characteristic-defects.png" (Salve, 2018). Right: corresponding simulated HRTEM image.

We investigate here the interaction of hydrogen with graphene using the Brenner module of the SAMSON software platform and we propose an original methodology to characterize its structural arrangement at the atomic scale by simulating HRSTEM images to interpret experimental results. In particular, we compare the effect of hydrogen on dark field (DF), bright field (BF), high-angle annular dark filed (HAADF) and annular bright field (ABF) images, to estimate the best technique suited to hydrogen detection. In addition, we present the effect of carbon vacancies and adatoms on the stability of hydrogen coverage, associated to the HRSTEM signatures of the most stable configurations. These results provide the necessary building blocks to analyze the structure and energetics of hydrogenated graphene-based materials at the atomic scale.

6.5. Atomistic modelling of diamond-type $Si_xGe_yC_zSn_{1-x-y-z}$ crystals for realistic transmission electron microscopy image simulations

Participants: Leonard Jaillet and Cyril Guedj.

The realistic simulations of transmission electron microscopy (TEM) images requires an accurate definition of the positions of all atoms, which are linked to the mechanical properties of the material. We are working on an approach to build optimized models to represent the lattice parameters and elastic properties of Si, Ge, diamond, alpha-tin and related diamond alloys.

In order to compute precisely the complex $Si_xGe_yC_zSn_{1-x-y-z}$ diamond crystals, a dedicated parametrization of the Keating force field has been proposed. An original periodic boundary strategy has also been provided. Our tool can be used to interpret experimental TEM with a speed several orders of magnitude higher than for ab-initio methods. The method predicts the correct lattice parameters and elastic constants for published experimental results with low deviation. Finally, we have shown that subsequent Monte Carlo simulations predict original self-ordering effects in C in good agreement with the theory. A publication is in preparation on this topic.

6.6. Analytical symmetry detection method AnAnaS

Participants: Guillaume Pagès, Sergei Grudinin, Elvira Kinzina.

Publications: Journal of Structural Biology, 2018 [21], Journal of Structural Biology, 2018 [20].

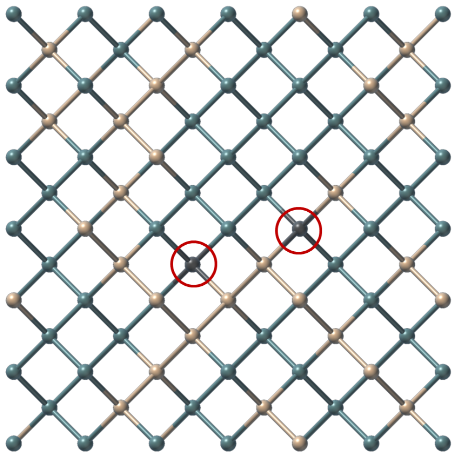


Figure 4. Crystal of $Si_{40}Ge_{60}$ where two carbon atoms (circled in red) have been inserted. The properties of the crystal such as its lattice parameter can be characterized in function of the position of the carbon atoms.

Macromolecules are generally not rigid bodies at physiological temperature and they adopt different conformational states. Thus, if one considers a macromolecular assembly made of N subunits, do we expect that all the units will be structurally identical to each other? Most probably not, since at any given moment of time, each unit may be sampling a different conformational state. For example, there are plenty of X-ray structures of homo-dimers, where the individual monomers are not structurally identical.

In order to quantitatively assess these differences, we developed a method for Analytical Analysis of Symmetries (AnAnaS) in protein complexes. The method is extremely fast, robust and accurate. Two papers describing the method were published [21], [20]. This method is available on the website of the team (https://team.inria.fr/nano-d/software/ananas/).

6.7. Deep Learning for Symmetry detection

Participants: Guillaume Pagès, Sergei Grudinin.
Publication: arXiv preprint, 2018 [29].

We worked on a fully-structural method for detecting symmetries in molecular structures. This allowed us to detect tandem repeats, or even symmetry in density maps. We created a method based on neural network and deep learning, inspired by the advances in computer vision in the past decade. According to our tests on simulated examples, our method is able to detect the order of a cyclic symmetry (which can be 1 for asymmetric structure) with a 92% accuracy, and guesses the direction of the axis of symmetry with an average error of 3°. A manuscript describing this method has been submitted for publication and is available on arXiv [29].

6.8. New method for protein model quality assessment Ornate

Participants: Benoit Charmettant, Guillaume Pagès, Sergei Grudinin.
Publication: bioRxiv preprint, 2018 [28].

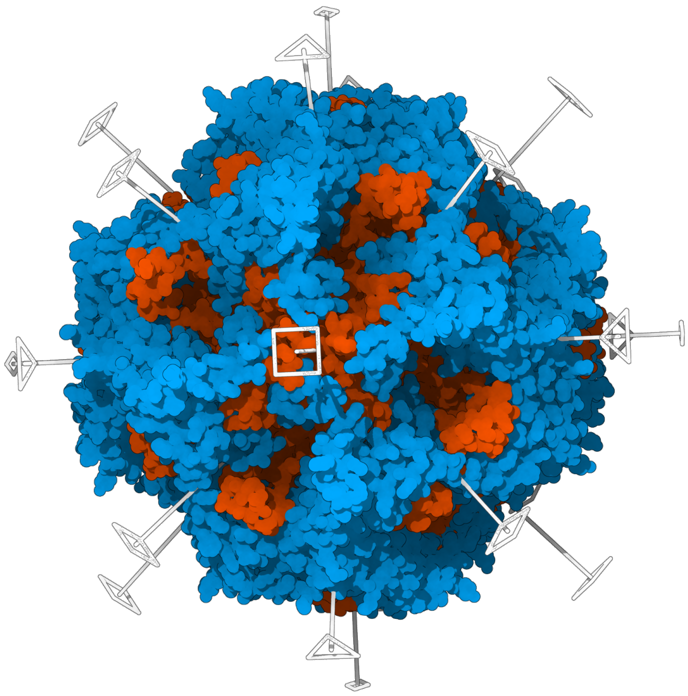


Figure 5. System with an octahedral symmetry, with the symmetry axes displayed in SAMSON.

Protein model quality assessment (QA) is a crucial and yet open problem in structural bioinformatics. It consists of estimating a score to assess whether a given three-dimensional structure is correctly folded or not. The current best methods for single-model QA typically combine results from different approaches, each based on different input features constructed by experts in the field. Then, the prediction model is trained using a machine-learning algorithm. Recently, with the development of convolutional neural networks (CNN), the training paradigm has changed. In computer vision, the expert-developed features have been significantly overpassed by automatically trained convolutional filters. This motivated us to apply a three-dimensional (3D) CNN to the problem of protein model QA.

We developed a novel method for single-model QA called Ornate. Ornate (Oriented Routed Neural network with Automatic Typing) is a residue-wise scoring function that takes as input 3D density maps. It predicts the local (residue-wise) and the global model quality through a deep 3D CNN. Specifically, Ornate aligns the input density map, corresponding to each residue and its neighborhood, with the backbone topology of this residue. This circumvents the problem of ambiguous orientations of the initial models. Also, Ornate includes automatic identification of atom types and dynamic routing of the data in the network. Established benchmarks (CASP 11 and CASP 12) demonstrate the state-of-the-art performance of our approach among single-model QA methods. A manuscript describing this method has been submitted for publication and is available on bioRxiv [28].

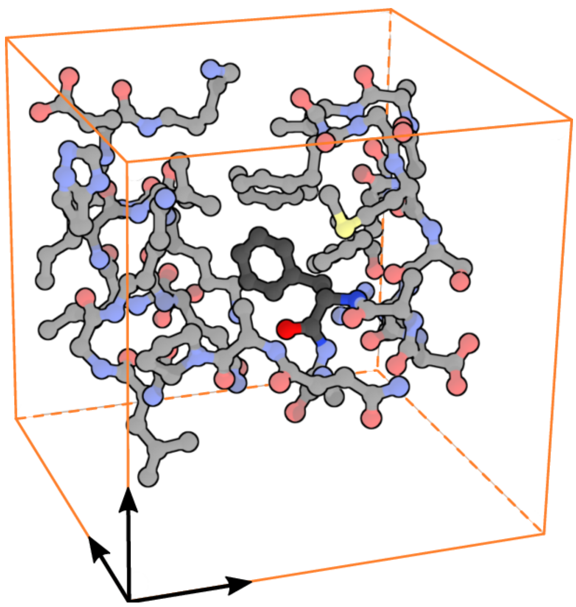


Figure 6. Example of input given to the 3D CNN Ornate.

7. Partnerships and Cooperations

7.1. Regional Initiatives

- Sergei Grudinin has obtained an IDEX UGA grant. It covers 2 years of post-doc of Didier Devaurs, starting from December 2018.
- Doctoral UGA grant covers PhD thesis of Maria Kadukova (supervised by Sergei Grudinin).

7.2. National Initiatives

7.2.1. ANR

In 2018, NANO-D had funding from one ANR program:

- **ANR PRCI**: covers PhD thesis of Guillaume Pages.

7.3. European Initiatives

7.3.1. Collaborations with Major European Organizations

Partner 1: The European Bioinformatics Institute (EMBL-EBI), Protein Data Bank in Europe (PDBe) team, Hinxton (UK)

We are collaborating on the integration of methods developed in the team into the PDBe web resource.

Partner 2: The Institute Laue-Langevin (ILL), the bioSANS team, Grenoble (France)

We are collaborating on the development of neutron small-angle scattering software.

Partner 3: The Ecole polytechnique fédérale de Lausanne (EPFL), Laboratory for Biomolecular Modeling, Lausanne (Switzerland)

We are collaborating on the integrative structural biology approaches.

7.4. International Initiatives

7.4.1. Inria International Partners

7.4.1.1. Declared Inria International Partners : BIOTOOLS

Title: Novel Computational Tools for Structural Bioinformatics

International Partner (Institution - Laboratory - Researcher):

MIPT (Russia (Russian Federation)) - Department of Control and Applied Mathematics - Vadim Strijov

Duration: 2016 - 2020

Start year: 2016

Abstract : The general scientific objectives of the forthcoming collaboration are the new developments of computational tools for structural bioinformatics. In particular, we plan to collaborate on several subjects: 1. Development of tractable approximations for intractable combinatorial problems in structural biology. 2. Development of new computational tools for scattering experiments. 3. Machine learning for structural bioinformatics.

7.4.1.2. Informal International Partners

- University of Stony Brook, lab of Dima Kozakov (USA). We have been collaborating on the development of novel protein docking methods.
- University of Vilnius, department of Bioinformatics (Lithuania). We have been collaborating on the development of novel protein docking methods.
- KU Copenhagen (Denmark), department of Chemistry. We collaborate on the integrative structural biology approaches.
- Autonomous University of Madrid (Spain), Bioinformatics Unit. We collaborate on the development of computational methods for protein flexibility.
- Francis Crick Institute, London (UK), Biomolecular Modelling Laboratory. We collaborate on the development of flexible protein docking methods.

7.4.2. Participation in Other International Programs

Our team has obtained the PHC Gilibert grant for a 2-year collaboration with the Vilnius University (Lithuania). Our partner is the Department of Bioinformatics, http://www.bti.vu.lt/en/departments/department-of-bioinformatics.

7.5. International Research Visitors

7.5.1. Visits of International Scientists

- Karina Dos Santos Machado, lecturer at the Federal University of Rio Grande (FURG, Brazil), Oct 2018 - Oct 2019.

7.5.1.1. Internships

- Amal Akkari (Mohammed V University, Rabat, Morocco), Jun 2018 - Nov 2018.
- Khalid Mustafin (MIPT Moscow, Russia), Sep 2018 - Feb 2019.

7.5.2. Visits to International Teams

Sergei Grudinin has visited and gave talks in the following research labs :

- Department of Bioinformatics, University of Vilnius, Lithuania, May 10-11, 2018.
- The team of Dima Kozakov,Stony Brook University, USA, November 12th-14th, 2018.
- The team of Simon Billinge, Columbia University, USA, November 15th, 2018.
- Department of Biology, University of Copenhagen, Denmark, November 19th, 2018.
- Department of Chemistry, University of Copenhagen, Denmark, November 22nd, 2018.
- Department of Bioinformatics, State Belorussian University, Minsk, Belarus, Dec 28, 2018.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Reviewer

- Léonard Jaillet has been reviewer for the IROS 2018 conference (International Conference on Intelligent Robots and Systems).
- Sergei Grudinin was a reviewer for the BIBM'18 conference (The IEEE International Conference on Bioinformatics and Biomedicine), and 6th International Work-Conference on Bioinformatics and Biomedical Engineering (IWBBIO 2018).

8.1.2. Journal

8.1.2.1. Reviewer - Reviewing Activities

Sergei Grudinin has reviewed submissions for the following journals : PLOS Computational Biology, Journal of Computer–Aided Molecular Design, Bioinformatics, Computational Biology and Chemistry, Journal of Computational Chemistry, Proteins, Nature, BMC Bioinformatics, The Journal of Physical Chemistry, IEEE Access, Accounts of Chemical Research, Computational and Structural Biotechnology Journal.

8.1.3. Invited Talks

- Sergei Grudinin gave an invited talk 'Algorithms for Protein–Protein Docking' at the Meet–U 2018 course on structural bioinformatics, Paris, 2018.
- Sergei Grudinin gave an invited talk 'Using Machine Learning for Structure–Based Predictions of Protein–Ligand Interactions' at the 7th French–Japanese Workshop on Computational Methods in Chemistry, Strasburg, 2018.
- Sergei Grudinin gave an invited talk 'Novel Methods for Structural Bioinformatics' at the GDR BIM / GT Méthodes Algorithmiques pour les Structures et Interactions, Paris 2018.
- Sergei Grudinin gave an invited talk 'Using machine learning to predict protein structure and interactions' at the VIth International Conference "Chemistry, Structure and Function of Biomolecules", Minsk, 2018.

- Sergei Grudinin gave an invited talk 'Predicting protein interactions with protein flexibility and small-angle scattering profiles' at the Modeling of Protein Interactions (MPI) 2018 conference, November 8-10, Lawrence, KS, USA.
- Sergei Grudinin gave an invited talk on data-assisted modeling of protein structures at the CASP13 conference, Iberostar Paraiso, Riviera Maya, Mexico December 1-4, 2018.
- Sergei Grudinin gave an invited talk 'Artificial Intelligence for Learning Protein Interactions' at the 4th International Conference on Mathematical and Computational Medicine, December 3-7, 2018, Cancun, Mexico.

8.1.4. Leadership within the Scientific Community

Sergei Grudinin with his colleagues from ILL has organized a new data-assisted (SANS) sub-challenge for the CASP13 community-wise protein structure prediction exercise.

8.1.5. Scientific Expertise

Sergei Grudinin reviewed an application for the OPUS funding scheme at the National Science Center, Poland.

8.2. Teaching - Supervision - Juries

8.2.1. Supervision

PhD : Phd thesis defence of Minh Khoa Nguyen, Université Grenoble Alpes, 2018
Title: Efficient exploration of molecular paths from As-Rigid-As-Possible approaches and motion planning methods [67].
Thesis committee: Emmanuel Mazer, Léonard Jaillet, Stéphane Redon, Juan Cortes, Charles Robert, Dirk Stratmann.
Summary: In this dissertation, we are particularly interested in developing new methods to find for a system made of a single protein or a protein and a ligand, the pathways that allow a transition from one state to another. During a few past decades, a vast amount of computational methods has been proposed to address this problem. However, these methods still have to face two challenges: the high dimensionality of the representation space, associated to the large number of atoms in these systems, and the complexity of the interactions between these atoms. This dissertation proposes two novel methods to efficiently find relevant pathways for such biomolecular systems. The methods are fast and their solutions can be used, analyzed or improved with more specialized methods. The first proposed method generates interpolation pathways for biomolecular systems using the As-Rigid-As-Possible (ARAP) principle from computer graphics. The method is robust and the generated solutions best preserve the local rigidity of the original system. An energy-based extension of the method is also proposed, which significantly improves the solution paths. However, in the scenarios requiring complex deformations, this approach may still generate unnatural paths. Therefore, we propose a second method called ART-RRT, which combines the ARAP principle for reducing the dimensionality, with the Rapidly-exploring Random Trees from robotics for efficiently exploring possible pathways. This method not only gives a variety of pathways in reasonable time but the pathways are also low-energy and clash-free, with the local rigidity preserved as much as possible. The mono-directional and bi-directional versions of the ART-RRT method were applied for finding ligand-unbinding and protein conformational transition pathways, respectively. The results were found to be in good agreement with experimental data and other state-of-the-art solutions.

PhD : Phd thesis defence of Alexandre Hoffmann, Université Grenoble Alpes, 2018
Title: Docking Flexible Proteins using Polynomial Expansions.
Thesis committee: Valérie Perrier, Slavica Jonic, Florence Tama, Sergei Grudinin, Marc Delarue, Roland Hildebrand.
Summary: This thesis focuses on two main axes. The first axis is the development of a new method that exhaustively samples both rigid-body and collective motions computed via normal mode analysis (NMA). We first present a method that combines the advantages of the fast Fourier

transform (FFT)-based exhaustive search, which samples all the conformations of a system under study on a grid, with a local optimization technique that guarantees to find the nearest optimal off-grid and flexible conformation. The algorithm first samples a quadratic approximation of a scoring function on a 6D grid. Then, the method performs the flexible search by maximizing the quadratic approximation of the cost function within a certain search space. We then present a multi-step version of our algorithm, which finds the collective motions that maximize the docking score with respect to the rigid-body degrees of freedom (DOFs). The method exhaustively samples both rigid- body and collective motions by maximizing the soft maximum over the rigid body DOFs of the docking/fitting cost function. Both methods were applied to docking problems on both real and artificial example and we were able to design a benchmark in which the “fit then refine” approach fails at finding the correct conformation while our method succeeds.

The second axis is the development of a new extrapolation of motions computed by NMA. We show that it is possible, with minimal computations, to extrapolate the instantaneous motions computed by NMA in the rotations-translations of blocks (RTB) subspace as an almost pure rotation around a certain axis. We applied this non-linear block (NOLB) method on various biological systems and were able to, firstly, retrieve biologically relevant motions and secondly, to demonstrate that the NOLB method generates structures with a better topology than a linear NMA method.

PhD : Phd thesis defence of Semeho Prince A. Edorh, Université Grenoble Alpes, 2018
Title: Incremental Algorithm for long range interactions [11].
Thesis committee: Stephane Redon, Olivier Coulaud, Matthias Bolten, Jean-Louis Barrat, Stefano Mossa, Jerôme Mathe.
Summary: Particle simulations have become an essential tool in various fields such as physics, astrophysics, biology, chemistry, climatology, and engineering, to name few. Usually, these computer simulations produce a temporal evolution of the system of interest by describing the motion of particles. In order to perform reliable simulations, we must provide an accurate description of interaction forces undergone by each particle. In most cases, these forces mirror inter-particle interactions and depend on relative coordinates of the particles. Moreover, pairwise long-range interactions are generally the cornerstone of particle simulations, an example being gravitational forces that are so essential in astrophysics. In molecular simulations, electrostatic forces are the most common illustration of long-range interactions. Furthermore, due to their computational cost, pairwise long-range interactions are the bottleneck of particle simulations. Therefore, sophisticated algorithms must be used for efficient evaluations of these interactions. In this thesis, we thus propose algorithms which may reduce the cost of long-range interactions when the studied system is governed by a particular dynamics. Precisely, these so-called «incremental» algorithms are effective for simulations where a part of the system remains frozen awhile. In particular, our algorithms will be validated on systems whose particles are governed by the so-called Adaptively Restrained Molecular Dynamics (ARMD) which is a promising approach in molecular dynamics simulations. Although several incremental algorithms introduced by this thesis will be devoted to molecular dynamics simulations, we believe that they can be generalized to all kinds of long-range interactions.

PhD in progress : Maria Kadukova, "Novel computational approaches for protein ligand interactions", Sep 2016-, supervisors: Sergei Grudinin (France) and Vladimir Chupin (MIPT, Russia).

PhD in progress : Guillaume Pagès, "Novel computational developments for protein structure prediction", Apr 2016-, supervisors: Sergei Grudinin (Inria), Valentin Gordeliy (IBS).

8.3. Popularization

8.3.1. Articles and contents

Sergei Grudinin and Guillaume Pagès were interviewed by Le Figaro about the progress of deep learning and artificial intelligence in protein structure prediction [88].

9. Bibliography

Major publications by the team in recent years

[1] S. GRUDININ, M. GARKAVENKO, A. KAZENNOV.*Pepsi-SAXS : an adaptive method for rapid and accurate computation of small-angle X-ray scattering profiles*, in "Acta Crystallographica Section D: Biological Crystallography", May 2017, vol. D73, p. 449 - 464 [*DOI :* 10.1107/S2059798317005745], https://hal.inria.fr/hal-01516719

[2] A. HOFFMANN, S. GRUDININ.*NOLB: Nonlinear Rigid Block Normal Mode Analysis Method*, in "Journal of Chemical Theory and Computation", April 2017, vol. 13, n° 5, p. 2123-2134 [*DOI :* 10.1021/ACS.JCTC.7B00197], https://hal.inria.fr/hal-01505843

[3] L. JAILLET, S. ARTEMOVA, S. REDON.*IM-UFF: extending the Universal Force Field for interactive molecular modeling*, in "Journal of Molecular Graphics and Modelling", October 2017, vol. 77, p. 350 - 362 [*DOI :* 10.1016/J.JMGM.2017.08.023], https://hal.inria.fr/hal-01676519

[4] M. KADUKOVA, S. GRUDININ.*Convex-PL: a novel knowledge-based potential for protein-ligand interactions deduced from structural databases using convex optimization*, in "Journal of Computer-Aided Molecular Design", October 2017, vol. 31, n° 10, p. 943–958 [*DOI :* 10.1007/S10822-017-0068-8], https://hal.inria.fr/hal-01591154

[5] E. NEVEU, D. RITCHIE, P. POPOV, S. GRUDININ.*PEPSI-Dock: a detailed data-driven protein–protein interaction potential accelerated by polar Fourier correlation*, in "Bioinformatics", August 2016, vol. 32, n° 7, p. i693-i701 [*DOI :* 10.1093/BIOINFORMATICS/BTW443], https://hal.archives-ouvertes.fr/hal-01358645

[6] M. K. NGUYEN, L. JAILLET, S. REDON.*As-Rigid-As-Possible molecular interpolation paths*, in "Journal of Computer-Aided Molecular Design", April 2017, vol. 31, n° 4, p. 403 - 417 [*DOI :* 10.1007/S10822-017-0012-Y], https://hal.inria.fr/hal-01676132

[7] G. PAGÈS, S. GRUDININ.*Analytical symmetry detection in protein assemblies. II. Dihedral and Cubic symmetries*, in "Journal of Structural Biology", September 2018, vol. 203, n° 3, p. 185-194 [*DOI :* 10.1016/J.JSB.2018.05.005], https://hal.inria.fr/hal-01816449

[8] G. PAGÈS, E. KINZINA, S. GRUDININ.*Analytical symmetry detection in protein assemblies. I. Cyclic symmetries*, in "Journal of Structural Biology", August 2018, vol. 203, n° 2, p. 142-148 [*DOI :* 10.1016/J.JSB.2018.04.004], https://hal.inria.fr/hal-01779893

[9] P. POPOV, S. GRUDININ.*Knowledge of Native Protein–Protein Interfaces Is Sufficient To Construct Predictive Models for the Selection of Binding Candidates*, in "Journal of Chemical Information and Modeling", September 2015, vol. 55, n° 10, p. 2242–2255 [*DOI :* 10.1021/ACS.JCIM.5B00372], https://hal.inria.fr/hal-01229886

[10] D. W. RITCHIE, S. GRUDININ.*Spherical polar Fourier assembly of protein complexes with arbitrary point group symmetry*, in "Journal of Applied Crystallography", February 2016, vol. 49, n° 1, p. 158-167 [*DOI :* 10.1107/S1600576715022931], https://hal.inria.fr/hal-01261402

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] S. P. A. EDORH.*Incremental Algorithm for long range interactions*, Université Grenoble - Alpes, October 2018, https://hal.inria.fr/tel-01980307

[12] M. K. NGUYEN.*Efficient exploration of molecular paths from As-Rigid-As-Possible approaches and motion planning methods*, Université Grenoble Alpes, March 2018, https://tel.archives-ouvertes.fr/tel-01978418

Articles in International Peer-Reviewed Journal

[13] G. DEREVYANKO, S. GRUDININ, Y. BENGIO, G. LAMOUREUX.*Deep convolutional networks for quality assessment of protein folds*, in "Bioinformatics", 2018, https://arxiv.org/abs/1801.06252 - 8 pages [*DOI :* 10.1093/BIOINFORMATICS/BTY494], https://hal.inria.fr/hal-01702857

[14] S. P. A. EDORH, S. REDON.*Incremental update of electrostatic interactions in adaptively restrained particle simulations*, in "Journal of Computational Chemistry", July 2018, vol. 39, n^{o} 20, p. 1455-1469 [*DOI :* 10.1002/JCC.25215], https://hal.inria.fr/hal-01761906

[15] M. KADUKOVA, S. GRUDININ.*Docking of small molecules to farnesoid X receptors using AutoDock Vina with the Convex-PL potential: lessons learned from D3R Grand Challenge 2*, in "Journal of Computer-Aided Molecular Design", January 2018, vol. 32, n^{o} 1, p. 151–162 [*DOI :* 10.1007/S10822-017-0062-1], https://hal.inria.fr/hal-01591157

[16] M. KARASIKOV, G. PAGÈS, S. GRUDININ.*Smooth orientation-dependent scoring function for coarse-grained protein quality assessment*, in "Bioinformatics", December 2018, p. 1-8 [*DOI :* 10.1093/BIOINFORMATICS/BTY1037], https://hal.inria.fr/hal-01971128

[17] E. NEVEU, P. POPOV, A. HOFFMANN, A. MIGLIOSI, X. BESSERON, G. DANOY, P. BOUVRY, S. GRUDININ.*RapidRMSD: Rapid determination of RMSDs corresponding to motions of flexible molecules*, in "Bioinformatics", August 2018, vol. 34, n^{o} 16, p. 2757–2765 [*DOI :* 10.1093/BIOINFORMATICS/BTY160], https://hal.inria.fr/hal-01735214

[18] M. K. NGUYEN, L. JAILLET, S. REDON.*ART-RRT: As-Rigid-As-Possible exploration of ligand unbinding pathways*, in "Journal of Computational Chemistry", April 2018, vol. 39, n^{o} 11, p. 665-678 [*DOI :* 10.1002/JCC.25132], https://hal.inria.fr/hal-01973778

[19] M. K. NGUYEN, L. JAILLET, S. REDON.*Generating conformational transition paths with low potential-energy barriers for proteins*, in "Journal of Computer-Aided Molecular Design", August 2018, vol. 32, n^{o} 8, p. 853-867 [*DOI :* 10.1007/S10822-018-0137-7], https://hal.inria.fr/hal-01973757

[20] G. PAGÈS, S. GRUDININ.*Analytical symmetry detection in protein assemblies. II. Dihedral and Cubic symmetries*, in "Journal of Structural Biology", September 2018, vol. 203, n^{o} 3, p. 185-194 [*DOI :* 10.1016/J.JSB.2018.05.005], https://hal.inria.fr/hal-01816449

[21] G. PAGÈS, E. KINZINA, S. GRUDININ.*Analytical symmetry detection in protein assemblies. I. Cyclic symmetries*, in "Journal of Structural Biology", August 2018, vol. 203, n^{o} 2, p. 142-148 [*DOI :* 10.1016/J.JSB.2018.04.004], https://hal.inria.fr/hal-01779893

[22] P. POPOV, S. GRUDININ.*Eurecon: Equidistant Uniform Rigid-body Ensemble Constructor*, in "Journal of Molecular Graphics and Modelling", March 2018, vol. 80, p. 313-319 [*DOI :* 10.1016/J.JMGM.2018.01.015], https://hal.inria.fr/hal-01702810

[23] K. K. SINGH, S. REDON.*Single-pass Incremental Force Updates for Adaptively Restrained Molecular Dynamics*, in "Journal of Computational Chemistry", March 2018, vol. 39, n^{o} 8, p. 412-423 [*DOI :* 10.1002/JCC.25126], https://hal.inria.fr/hal-01635863

[24] G. STOLTZ, Z. TRSTANOVA.*Langevin dynamics with general kinetic energies*, in "Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal", May 2018, vol. 16, n^{o} 2, p. 777-806, https://arxiv.org/abs/1609.02891 [*DOI :* 10.1137/16M110575X], https://hal.archives-ouvertes.fr/hal-01364821

International Conferences with Proceedings

[25] C. GUEDJ, L. JAILLET, F. ROUSSE, S. REDON.*Atomistic Modelling and Simulation of Transmission Electron Microscopy Images: Application to Intrinsic Defects of Graphene*, in "8th International Conference on Simulation and Modeling Methodologies, Technologies and Applications", Porto, Portugal, Proceedings of 8th International Conference on Simulation and Modeling Methodologies, Technologies and Applications, July 2018 [*DOI :* 10.5220/0006829200150024], https://hal.inria.fr/hal-01973626

Conferences without Proceedings

[26] C. GUEDJ, L. JAILLET, F. ROUSSE, S. REDON.*Impact of Hydrogen on Graphene-based Materials: Atomistic Modeling and Simulation of HRSTEM Images*, in "AVS 65th International Symposium & Exhibition", Long Beach, United States, October 2018, https://hal.inria.fr/hal-01973651

Other Publications

[27] S. GRUDININ.*User Guide for NOLB : Non-Linear Normal Mode Analysis*, February 2018, working paper or preprint, https://hal.inria.fr/hal-01565129

[28] G. PAGÈS, B. CHARMETTANT, S. GRUDININ.*Protein model quality assessment using 3D oriented convolutional neural networks*, October 2018, working paper or preprint [*DOI :* 10.1101/432146], https://hal.inria.fr/hal-01899468

[29] G. PAGÈS, S. GRUDININ.*DeepSymmetry : Using 3D convolutional networks for identification of tandem repeats and internal symmetries in protein structures*, October 2018, https://arxiv.org/abs/1810.12026 - working paper or preprint, https://hal.inria.fr/hal-01903624

References in notes

[30] *Protein Data Bank in Europe*, 2018, https://www.ebi.ac.uk/pdbe/

[31] B. AHMADI, M. KASSIRIHA, K. KHODABAKHSHI, E. R. MAFI.*Effect of nano layered silicates on automotive polyurethane refinish clear coat*, in "Progress in Organic Coatings", 2007, vol. 60, n^{o} 2, p. 99 - 104 [*DOI :* 10.1016/J.PORGCOAT.2007.07.008], http://www.sciencedirect.com/science/article/pii/S0300944007001464

[32] F. H. ALLEN.*The Cambridge Structural Database: a quarter of a million crystal structures and rising*, in "Acta Crystallographica Section B", Jun 2002, vol. 58, n^{o} 3 Part 1, p. 380–388, http://dx.doi.org/10.1107/S0108768102003890

[33] F. AMPLE, S. AMI, C. JOACHIM, F. THIEMANN, G. RAPENNE.*A Morse manipulator molecule for the modulation of metallic shockley surface states*, in "Chemical Physics Letters", 2007, vol. 434, p. 280-285 [*DOI :* 10.1016/J.CPLETT.2006.12.021], http://www.sciencedirect.com/science/article/pii/S0009261406018148

[34] F. AMPLE, C. JOACHIM.*A semi-empirical study of polyacene molecules adsorbed on a Cu(1 1 0) surface*, in "Surface Science", 2006, vol. 600, n^{o} 16, p. 3243 - 3251 [*DOI :* 10.1016/J.SUSC.2006.06.015], http://www.sciencedirect.com/science/article/pii/S003960280600700X

[35] A. ARKHIPOV, P. FREDDOLINO, K. IMADA, K. NAMBA, K. SCHULTEN.*Coarse-grained molecular dynamics simulations of a rotating bacterial flagellum*, in "Biophysical Journal", 2006, vol. 91, p. 4589-4597

[36] S. ARTEMOVA, S. GRUDININ, S. REDON.*A comparison of neighbor search algorithms for large rigid molecules*, in "Journal of Computational Chemistry", 2011, vol. 32, n^{o} 13, p. 2865–2877, http://dx.doi.org/10.1002/jcc.21868

[37] S. ARTEMOVA, S. REDON.*Adaptively Restrained Particle Simulations*, in "Phys. Rev. Lett.", Nov 2012, vol. 109, http://link.aps.org/doi/10.1103/PhysRevLett.109.190201

[38] H. M. BERMAN, J. WESTBROOK, Z. FENG, G. GILLILAND, T. N. BHAT, H. WEISSIG, I. N. SHINDYALOV, P. E. BOURNE.*The Protein Data Bank*, in "Nucleic Acids Research", 2000, vol. 28, n^{o} 1, p. 235-242 [*DOI :* 10.1093/NAR/28.1.235], http://nar.oxfordjournals.org/content/28/1/235.abstract

[39] X. BLANC, C. LE BRIS, F. LEGOLL.*Analysis of a prototypical multiscale method coupling atomistic and continuum mechanics*, in "ESAIM: Mathematical Modelling and Numerical Analysis", 2005, vol. 39, n^{o} 04, p. 797-826, http://dx.doi.org/10.1051/m2an:2005035

[40] M. BOSSON, S. GRUDININ, X. BOUJU, S. REDON.*Interactive physically-based structural modeling of hydrocarbon systems*, in "Journal of Computational Physics", 2012, vol. 231, n^{o} 6, p. 2581 - 2598 [*DOI :* 10.1016/J.JCP.2011.12.006], http://www.sciencedirect.com/science/article/pii/S0021999111007042

[41] D. W. BRENNER.*Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films*, in "Phys. Rev. B", Nov 1990, vol. 42, p. 9458–9471, http://link.aps.org/doi/10.1103/PhysRevB.42.9458

[42] T. CAGIN, G. WANG, R. MARTIN, G. ZAMANAKOS, N. VAIDEHI, D. T. MAINZ, W. A. GODDARD III.*Multiscale modeling and simulation methods with applications to dendritic polymers*, in "Computational and Theoretical Polymer Science", 2001, vol. 11, n^{o} 5, p. 345 - 356 [*DOI :* 10.1016/S1089-3156(01)00026-5], http://www.sciencedirect.com/science/article/pii/S1089315601000265

[43] E. CANCES, F. CASTELLA, P. CHARTIER, E. FAOU, C. LE BRIS, F. LEGOLL, G. TURINICI.*Long-time averaging for integrable Hamiltonian dynamics*, in "Numerische Mathematik", 2005, vol. 100, p. 211-232, 10.1007/s00211-005-0599-0, http://dx.doi.org/10.1007/s00211-005-0599-0

[44] Q. CHAUDHRY, M. SCOTTER, J. BLACKBURN, B. ROSS, A. BOXALL, L. CASTLE, R. AITKEN, R. WATKINS.*Applications and implications of nanotechnologies for the food sector*, in "Food Additives & Contaminants: Part A", 2008, vol. 25, n^{o} 3, p. 241-258, http://www.tandfonline.com/doi/abs/10.1080/02652030701744538

[45] X. CHEN, S. S. MAO.*Titanium Dioxide Nanomaterials: Synthesis, Properties, Modifications, and Applications*, in "ChemInform", 2007, vol. 38, n^{o} 41, http://dx.doi.org/10.1002/chin.200741216

[46] S. COOPER, F. KHATIB, A. TREUILLE, J. BARBERO, J. LEE, M. BEENEN, A. LEAVER-FAY, D. BAKER, Z. POPOVIC, F. PLAYERS.*Predicting protein structures with a multiplayer online game*, in "Nature", 2010, vol. 466, p. 756-760

[47] M. CURRELI, A. H. NADERSHAHI, G. SHAHI.*Emergence of nanomedical devices for the diagnosis and treatment of cancer: the journey from basic science to commercialization*, in "International Journal of Technology Transfer and Commercialisation", 2008, vol. 7, n^{o} 4, p. 290-307

[48] E. DARVE.*The Fast Multipole Method: Numerical Implementation*, in "Journal of Computational Physics", 2000

[49] H. DIETZ, S. M. DOUGLAS, W. M. SHIH.*Folding DNA into Twisted and Curved Nanoscale Shapes*, in "Science", 2009, vol. 325, n^{o} 5941, p. 725-730 [*DOI :* 10.1126/SCIENCE.1174251], http://www.sciencemag.org/content/325/5941/725.abstract

[50] S. J. FLEISHMAN, T. A. WHITEHEAD, D. C. EKIERT, C. DREYFUS, J. E. CORN, E.-M. STRAUCH, I. A. WILSON, D. BAKER.*Computational Design of Proteins Targeting the Conserved Stem Region of Influenza Hemagglutinin*, in "Science", 2011, vol. 332, n^{o} 6031, p. 816-821 [*DOI :* 10.1126/SCIENCE.1202617], http://www.sciencemag.org/content/332/6031/816.abstract

[51] G. FOX-RABINOVICH, B. BEAKE, K. YAMAMOTO, M. AGUIRRE, S. VELDHUIS, G. DOSBAEVA, A. ELFIZY, A. BIKSA, L. SHUSTER.*Structure, properties and wear performance of nano-multilayered TiAlCrSiYN/TiAlCrN coatings during machining of Ni-based aerospace superalloys*, in "Surface and Coatings Technology", 2010, vol. 204, p. 3698 - 3706 [*DOI :* 10.1016/J.SURFCOAT.2010.04.050], http://www.sciencedirect.com/science/article/pii/S0257897210003178

[52] M. GOLDBERG, R. LANGER, X. JIA.*Nanostructured materials for applications in drug delivery and tissue engineering*, in "Journal of Biomaterials Science, Polymer Edition", 2007, vol. 18, n^{o} 3, p. 241-268 [*DOI :* DOI:10.1163/156856207779996931], https://doi.org/10.1163/156856207779996931

[53] C. GUEDJ, L. JAILLET, F. ROUSSE, S. REDON.*Atomistic Modelling and Simulation of Transmission Electron Microscopy Images: Application to Intrinsic Defects of Graphene*, in "Proceedings of 8th International Conference on Simulation and Modeling Methodologies, Technologies and Applications - Volume 1: SIMULTECH,", SciTePress, 2018, p. 15-24, http://dx.doi.org/10.5220/0006829200150024

[54] J.-H. HE.*An elementary introduction to recently developed asymptotic methods and nanomechanics in textile engineering*, in "International Journal of Modern Physics B", 2008, vol. 22, n^{o} 21, p. 3487-3578

[55] S. HELVEG.*Structure and Dynamics of Nanocatalysts*, in "Microscopy and Microanalysis", 2010, vol. 16, n^{o} Supplement S2, p. 1712-1713, http://dx.doi.org/10.1017/S1431927610055005

[56] A. HEYDEN, D. G. TRUHLAR.*Conservative Algorithm for an Adaptive Change of Resolution in Mixed Atomistic/Coarse-Grained Multiscale Simulations*, in "Journal of Chemical Theory and Computation", 2008, vol. 4, n^o^ 2, p. 217-221, http://pubs.acs.org/doi/abs/10.1021/ct700269m

[57] A. HOFFMANN, S. GRUDININ.*NOLB: Nonlinear Rigid Block Normal Mode Analysis Method*, in "Journal of Chemical Theory and Computation", April 2017, vol. 13, n^o^ 5, p. 2123-2134 [*DOI :* 10.1021/ACS.JCTC.7B00197], https://hal.inria.fr/hal-01505843

[58] A. HOFFMANN, V. PERRIER, S. GRUDININ.*A novel fast Fourier transform accelerated off-grid exhaustive search method for cryo-electron microscopy fitting*, in "Journal of Applied Crystallography", August 2017, vol. 50, n^o^ 4, p. 1036-1047 [*DOI :* 10.1107/S1600576717008172], https://hal.inria.fr/hal-01553293

[59] V. HORNAK, R. ABEL, A. OKUR, B. STROCKBINE, A. ROITBERG, C. SIMMERLING.*Comparison of multiple Amber force fields and development of improved protein backbone parameters*, in "Proteins: Structure, Function, and Bioinformatics", 2006, vol. 65, n^o^ 3, p. 712–725, http://dx.doi.org/10.1002/prot.21123

[60] C. JOACHIM, H. TANG, F. MORESCO, G. RAPENNE, G. MEYER.*The design of a nanoscale molecular barrow*, in "Nanotechnology", 2002, vol. 13, n^o^ 3, 330, http://stacks.iop.org/0957-4484/13/i=3/a=318

[61] L. KALÉ, R. SKEEL, M. BHANDARKAR, R. BRUNNER, A. GURSOY, N. KRAWETZ, J. PHILLIPS, A. SHINOZAKI, K. VARADARAJAN, K. SCHULTEN.*NAMD2: Greater Scalability for Parallel Molecular Dynamics*, in "Journal of Computational Physics", 1999, vol. 151, n^o^ 1, p. 283 - 312 [*DOI :* 10.1006/JCPH.1999.6201], http://www.sciencedirect.com/science/article/pii/S0021999199962010

[62] Z. LI, H. A. SCHERAGA.*Monte Carlo-minimization approach to the multiple-minima problem in protein folding*, in "Proceedings of the National Academy of Sciences", 1987, vol. 84, n^o^ 19, p. 6611-6615, http://www.pnas.org/content/84/19/6611.abstract

[63] L. LO, Y. LI, K. YEUNG, C. YUEN.*Indicating the development stage of nanotechnology in the textile and clothing industry*, in "International Journal of Nanotechnology", 2007, vol. 4, n^o^ 6, p. 667-679

[64] W. LU, C. M. LIEBER.*Nanoelectronics from the bottom up*, in "Nature materials", 2007, vol. 6, n^o^ 11, p. 841-850

[65] M. K. NGUYEN, L. JAILLET, S. REDON.*ART-RRT: As-Rigid-As-Possible exploration of ligand unbinding pathways*, in "Journal of computational chemistry", 2018, vol. 39, n^o^ 11, p. 665–678

[66] M. K. NGUYEN, L. JAILLET, S. REDON.*Generating conformational transition paths with low potential-energy barriers for proteins*, in "Journal of Computer-Aided Molecular Design", 2018, vol. 32, n^o^ 8, p. 853–867

[67] M. K. NGUYEN.*Efficient exploration of molecular paths from As-Rigid-As-Possible approaches and motion planning methods*, Université Grenoble Alpes, 2018

[68] S. O. NIELSEN, P. B. MOORE, B. ENSING.*Adaptive Multiscale Molecular Dynamics of Macromolecular Fluids*, in "Phys. Rev. Lett.", Dec 2010, vol. 105, 237802, http://link.aps.org/doi/10.1103/PhysRevLett.105.237802

[69] A. NIKITIN, X. LI, Z. ZHANG, H. OGASAWARA, H. DAI, A. NILSSON.*Hydrogen Storage in Carbon Nanotubes through the Formation of Stable C-H Bonds*, in "Nano Letters", 2008, vol. 8, n^{o} 1, p. 162-167, PMID: 18088150, http://pubs.acs.org/doi/abs/10.1021/nl072325k

[70] G. PAGÈS, S. GRUDININ.*Ananas : Analytical symmetry detection in protein assemblies*, Inria, 2018, https://team.inria.fr/nano-d/software/ananas/

[71] M. PRAPROTNIK, L. DELLE SITE, K. KREMER.*Adaptive resolution scheme for efficient hybrid atomistic-mesoscale molecular dynamics simulations of dense liquids*, in "Phys. Rev. E", Jun 2006, vol. 73, 066701, http://link.aps.org/doi/10.1103/PhysRevE.73.066701

[72] M. PRAPROTNIK, S. MATYSIAK, L. D. SITE, K. KREMER, C. CLEMENTI.*Adaptive resolution simulation of liquid water*, in "Journal of Physics: Condensed Matter", 2007, vol. 19, n^{o} 29, 292201, http://stacks.iop.org/0953-8984/19/i=29/a=292201

[73] M. PRAPROTNIK, L. D. SITE, K. KREMER.*A macromolecule in a solvent: Adaptive resolution molecular dynamics simulation*, in "The Journal of Chemical Physics", 2007, vol. 126, n^{o} 13, 134902, http://aip.scitation.org/doi/abs/10.1063/1.2714540?journalCode=jcp

[74] M. PRAPROTNIK, L. D. SITE, K. KREMER.*Multiscale Simulation of Soft Matter: From Scale Bridging to Adaptive Resolution*, in "Annual Review of Physical Chemistry", 2008, vol. 59, n^{o} 1, p. 545-571, http://www.annualreviews.org/doi/abs/10.1146/annurev.physchem.59.032607.093707

[75] P. PROCACCI, T. DARDEN, M. MARCHI.*A Very Fast Molecular Dynamics Method To Simulate Biomolecular Systems with Realistic Electrostatic Interactions*, in "The Journal of Physical Chemistry", 1996, vol. 100, n^{o} 24, p. 10464-10468, http://pubs.acs.org/doi/abs/10.1021/jp960295w

[76] X. QIAN, T. SCHLICK.*Efficient multiple-time-step integrators with distance-based force splitting for particle-mesh-Ewald molecular dynamics simulations*, in "Journal of Chemical Physics", 2002, vol. 116, p. 5971-5983

[77] D. W. RITCHIE, G. J. KEMP.*Protein docking using spherical polar Fourier correlations*, in "Proteins: Structure, Function, and Bioinformatics", 2000, vol. 39, n^{o} 2, p. 178–194

[78] M. C. ROCO.*The long view of nanotechnology development: the National Nanotechnology Initiative at 10 years*, in "Journal of Nanoparticle Research", 2010

[79] B. ROOKS.*A shorter product development time with digital mock-up*, in "Assembly Automation", 1998, vol. 18, n^{o} 1, p. 34-38 [*DOI :* DOI:10.1108/01445159810201405], http://www.ingentaconnect.com/content/mcb/033/1998/00000018/00000001/art00004

[80] R. ROSSI, M. ISORCE, S. MORIN, J. FLOCARD, K. ARUMUGAM, S. CROUZY, M. VIVAUDOU, S. REDON.*Adaptive torsion-angle quasi-statics: a general simulation method with applications to protein structure analysis and design*, in "Bioinformatics", 2007, vol. 23, n^{o} 13 [*DOI :* 10.1093/BIOINFORMATICS/BTM191], http://bioinformatics.oxfordjournals.org/content/23/13/i408.abstract

[81] R. E. RUDD.*Coarse-Grained Molecular Dynamics for Computer Modeling of Nanomechanical Systems*, in "International Journal for Numerical Methods in Engineering", 2004

[82] A. SHIH, P. FREDDOLINO, A. ARKHIPOV, K. SCHULTEN.*Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations*, in "Journal of Structural Biology", 2007, vol. 157, p. 579-592

[83] Y. SHIRAI, A. J. OSGOOD, Y. ZHAO, Y. YAO, L. SAUDAN, H. YANG, C. YU-HUNG, L. B. ALEMANY, T. SASAKI, J.-F. MORIN, J. M. GUERRERO, K. F. KELLY, J. M. TOUR.*Surface-Rolling Molecules*, in "Journal of the American Chemical Society", 2006, vol. 128, n° 14, p. 4854-4864, PMID: 16594722, http://pubs.acs.org/doi/abs/10.1021/ja058514r

[84] E. G. STEIN, L. M. RICE, A. T. BRÜNGER.*Torsion-Angle Molecular Dynamics as a New Efficient Tool for NMR Structure Calculation*, in "Journal of Magnetic Resonance", 1997, vol. 124, n° 1, p. 154 - 164 [*DOI :* 10.1006/JMRE.1996.1027], http://www.sciencedirect.com/science/article/pii/S1090780796910277

[85] X. SUN, Z. LIU, K. WELSHER, J. ROBINSON, A. GOODWIN, S. ZARIC, H. DAI.*Nano-graphene oxide for cellular imaging and drug delivery*, in "Nano Research", 2008, vol. 1, p. 203-212, 10.1007/s12274-008-8021-8, http://dx.doi.org/10.1007/s12274-008-8021-8

[86] D. TOMALIA, L. REYNA, S. SVENSON.*Dendrimers as multi-purpose nanodevices for oncology drug delivery and diagnostic imaging*, in "Biochemical Society Transactions", 2007, vol. 35, p. 61-67

[87] N. VAIDEHI, W. A. GODDARD III.*Domain Motions in Phosphoglycerate Kinase using Hierarchical NEIMO Molecular Dynamics Simulations*, in "The Journal of Physical Chemistry A", 2000, vol. 104, n° 11, p. 2375-2383, http://pubs.acs.org/doi/abs/10.1021/jp991985d

[88] C. VANLERBERGHE.*DeepMind, l'IA de Google, est championne en biologie moléculaire*, Dec 2018, http://sante.lefigaro.fr/article/deepmind-la-filiale-de-google-est-championne-en-biologie-moleculaire/

[89] T. VETTOREL, A. Y. GROSBERG, K. KREMER.*Statistics of polymer rings in the melt: a numerical simulation study*, in "Physical Biology", 2009, vol. 6, n° 2, 025013, http://stacks.iop.org/1478-3975/6/i=2/a=025013

[90] W. YANG.*Direct calculation of electron density in density-functional theory*, in "Phys. Rev. Lett.", Mar 1991, vol. 66, p. 1438–1441, http://link.aps.org/doi/10.1103/PhysRevLett.66.1438

[91] W. YANG.*Electron density as the basic variable: a divide-and-conquer approach to the ab initio computation of large molecules*, in "Journal of Molecular Structure: THEOCHEM", 1992, vol. 255, n° 0, p. 461 - 479 [*DOI :* 10.1016/0166-1280(92)85024-F], http://www.sciencedirect.com/science/article/pii/016612809285024F

[92] A. C. T. VAN DUIN, S. DASGUPTA, F. LORANT, W. A. GODDARD III.*ReaxFF: A Reactive Force Field for Hydrocarbons*, in "The Journal of Physical Chemistry A", 2001, vol. 105, n° 41, p. 9396-9409, http://pubs.acs.org/doi/abs/10.1021/jp004368u

Project-Team NECS

Networked Controlled Systems

IN COLLABORATION WITH: Grenoble Image Parole Signal Automatique (GIPSA)

IN PARTNERSHIP WITH:
CNRS

Institut polytechnique de Grenoble

Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team NECS

Creation of the Project-Team: 2007 January 01, end of the Project-Team: 2018 December 31

Keywords:

Computer Science and Digital Science:

A1. - Architectures, systems and networks
A1.2. - Networks
A1.2.6. - Sensor networks
A1.2.7. - Cyber-physical systems
A1.2.9. - Social Networks
A1.5. - Complex systems
A3. - Data and knowledge
A3.1. - Data
A6. - Modeling, simulation and control
A6.1. - Methods in mathematical modeling
A6.2. - Scientific computing, Numerical Analysis & Optimization
A6.4. - Automatic control

Other Research Topics and Application Domains:

B7. - Transport and logistics
B7.1. - Traffic management
B7.2. - Smart travel

1. Team, Visitors, External Collaborators

Research Scientists

Carlos Canudas de Wit [Team leader, CNRS, Senior Researcher, HDR]
Maria Laura Delle Monache [Inria, Researcher]
Paolo Frasca [CNRS, Researcher]
Federica Garin [Inria, Researcher]

Faculty Members

Hassen Fourati [Univ. Grenoble Alpes, Associate Professor]
Alain Kibangou [Univ. Grenoble Alpes, Associate Professor, HDR]
Aida Makni [Univ. Grenoble Alpes, Assistant Professor, until Aug. 2018]

Technical Staff

Vadim Bertrand [CNRS]

PhD Students

Stéphane Durand [Univ. Grenoble Alpes, until Sep. 2018]
Sebin Gracy [Univ. Grenoble Alpes, until Sep. 2018]
Nicolas Martin [CNRS]
Stéphane Mollier [CNRS]
Thembani Moyo [Univ. of Johannesburg (South Africa), from May 2018]
Muhammad Umar B. Niazi [CNRS]
Denis Nikitin [CNRS, from Sep. 2018]
Bassel Othman [IFPEN, from Oct. 2018]

Martin Rodriguez [CNRS]
Nigina Toktassynova [Kazakh National Research Technical University (Kazakhstan)]
Liudmila Tumash [CNRS, from Sep. 2018]
Makia Zmitri [CNRS, from Oct. 2018]

Post-Doctoral Fellows

Giacomo Casadei [CNRS, until Sep. 2018]
Thibault Liard [Inria]

Visiting Scientists

Diego Deplano [Univ. di Cagliari (Italy), from Sep. 2018]
George Gunter [Vanderbilt Univ. Nashville (USA), Nov. 2018]
Maolong Lyu [TU Delft (The Netherlands), from Oct. 2018]
Olga Lucia Quintero Montoya [Universidad EAFIT Medellin (Colombia), Sep. 2018]
Raphael Stern [Vanderbilt Univ. Nashville (USA), Nov. 2018]

Administrative Assistants

Myriam Etienne [Inria]
Lydie Leon [CNRS, until Jul. 2018]
Julie Perrin [CNRS, from Jul. 2018]

2. Overall Objectives

2.1. Context and overall goal of the project

NECS is a joint INRIA/GIPSA-LAB team, bi-located at the INRIA-Rhône-Alpes Center in Montbonnot and at GIPSA-LAB (http://www.gipsa-lab.grenoble-inp.fr) in the Saint-Martin-d'Hères campus, both locations being in the Grenoble area. NECS team's research is focused on Networked Controlled Systems.
The research field of Networked Controlled Systems deals with feedback systems controlled over networks, but also concerns systems that naturally exhibit a network structure (e.g., traffic, electrical networks, etc.).

The first system category results from the arrival of new control problems posed by the consideration of several factors, such as: new technological components (e.g., wireless, RF, communications, local networks, etc.), increase of systems complexity (e.g., increase in vehicle components), the distributed location of sensor and actuator, and computation constraints imposed by their embedded nature. In this class of systems, the way that the information is transferred and processed (information constraints), and the manner in which the computation resources are used (resources management), have a substantial impact in the resulting stability and performance properties of the feedback controlled systems. One main challenge here is the co-design of control together with one or more other components of different nature. The NECS team has tackled co-design problems concerning:

- Control under communications and network constraints;
- Control under resources constraints.

The second category of systems is motivated by the natural network structure in which the original systems are built. Examples are biologic networks, traffic networks, and electrical networks. The complex nature of such systems makes the classical centralized view of the control design obsolete. New distributed and/or collaborative control and estimation algorithms need to be devised as a response to this complexity. Even if the dynamic behavior of each individual system is still important, the aggregated behavior (at some macroscopic level), and its interconnection graph properties become of dominant importance. To build up this research domain, the team has put a strong focus on traffic (vehicular) networks, and in some associated research topics capturing problems that are specific to these complex network systems (distributed estimation, graph-discovering, etc).

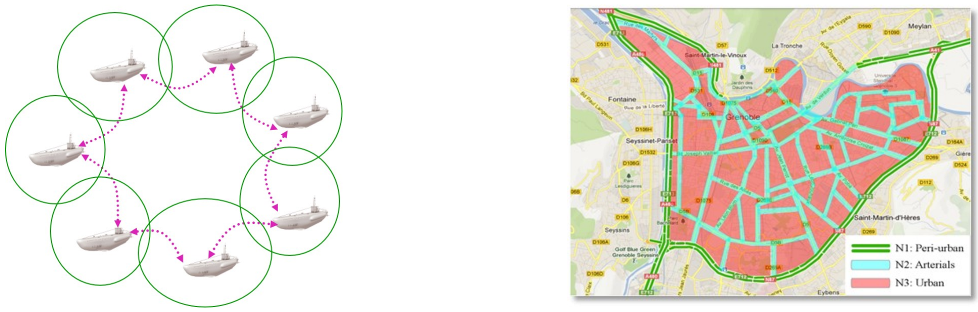


Figure 1. Left: a system of autonomous agents, where the network structure is created by the feedback, used to coordinate agents towards a common goal. Right: a system naturally having a network structure.

3. Research Program

3.1. Introduction

NECS team deals with Networked Control Systems. Since its foundation in 2007, the team has been addressing issues of control under imperfections and constraints deriving from the network (limited computation resources of the embedded systems, delays and errors due to communication, limited energy resources), proposing co-design strategies. The team has recently moved its focus towards general problems on *control of network systems*, which involve the analysis and control of dynamical systems with a network structure or whose operation is supported by networks. This is a research domain with substantial growth and is now recognized as a priority sector by the IEEE Control Systems Society: IEEE has started a new journal, IEEE Transactions on Control of Network Systems, whose first issue appeared in 2014.

More in detail, the research program of NECS team is along lines described in the following sections.

3.2. Distributed estimation and data fusion in network systems

This research topic concerns distributed data combination from multiple sources (sensors) and related information fusion, to achieve more specific inference than could be achieved by using a single source (sensor). It plays an essential role in many networked applications, such as communication, networked control, monitoring, and surveillance. Distributed estimation has already been considered in the team. We wish to capitalize and strengthen these activities by focusing on integration of heterogeneous, multidimensional, and large data sets:

- Heterogeneity and large data sets. This issue constitutes a clearly identified challenge for the future. Indeed, heterogeneity comes from the fact that data are given in many forms, refer to different scales, and carry different information. Therefore, data fusion and integration will be achieved by developing new multi-perception mathematical models that can allow tracking continuous (macroscopic) and discrete (microscopic) dynamics under a unified framework while making different scales interact with each other. More precisely, many scales are considered at the same time, and they evolve following a unique fully-integrated dynamics generated by the interactions of the scales. The new multi-perception models will be integrated to forecast, estimate and broadcast useful system states in a distributed way. Targeted applications include traffic networks and navigation.
- Multidimensionality. This issue concerns the analysis and the processing of multidimensional data, organized in multiway array, in a distributed way. Robustness of previously-developed algorithms will be studied. In particular, the issue of missing data will be taken into account. In addition,

since the considered multidimensional data are generated by dynamic systems, dynamic analysis of multiway array (or tensors) will be considered. The targeted applications concern distributed detection in complex networks and distributed signal processing for collaborative networks. This topic is developed in strong collaboration with UFC (Brazil).

3.3. Network systems and graph analysis

This is a research topic at the boundaries between graph theory and dynamical systems theory.

A first main line of research will be to study complex systems whose interactions are modeled with graphs, and to unveil the effect of the graph topology on system-theoretic properties such as observability or controllability. In particular, on-going work concerns observability of graph-based systems: after preliminary results concerning consensus systems over distance-regular graphs, the aim is to extend results to more general networks. A special focus will be on the notion of ‘generic properties’, namely properties which depend only on the underlying graph describing the sparsity pattern, and hold true almost surely with a random choice of the non-zero coefficients. Further work will be to explore situations in which there is the need for new notions different from the classical observability or controllability. For example, in opinion-forming in social networks or in formation of birds flocks, the potential leader might have a goal different from classical controllability. On the one hand, his goal might be much less ambitious than the classical one of driving the system to any possible state (e.g., he might want to drive everybody near its own opinion, only, and not to any combination of different individual opinions), and on the other hand he might have much weaker tools to construct his control input (e.g., he might not know the whole system’s dynamics, but only some local partial information). Another example is the question of detectability of an unknown input under the assumption that such an input has a sparsity constraint, a question arising from the fact that a cyber-physical attack might be modeled as an input aiming at controlling the system’s state, and that limitations in the capabilities of the attacker might be modeled as a sparsity constraint on the input.

A second line of research will concern graph discovery, namely algorithms aiming at reconstructing some properties of the graph (such as the number of vertices, the diameter, the degree distribution, or spectral properties such as the eigenvalues of the graph Laplacian), using some measurements of quantities related to a dynamical system associated with the graph. It will be particularly challenging to consider directed graphs, and to impose that the algorithm is anonymous, i.e., that it does not makes use of labels identifying the different agents associated with vertices.

3.4. Collaborative and distributed network control

This research line deals with the problem of designing controllers with a limited use of the network information (i.e. with restricted feedback), and with the aim to reach a pre-specified global behavior. This is in contrast to centralized controllers that use the whole system information and compute the control law at some central node. Collaborative control has already been explored in the team in connection with the underwater robot fleet, and to some extent with the source seeking problem. It remains however a certain number of challenging problems that the team wishes to address:

- Design of control with limited information, able to lead to desired global behaviors. Here the graph structure is imposed by the problem, and we aim to design the “best” possible control under such a graph constraint [0]. The team would like to explore further this research line, targeting a better understanding of possible metrics to be used as a target for optimal control design. In particular, and in connection with the traffic application, the long-standing open problem of ramp metering control under minimum information will be addressed.

[0] Such a problem has been previously addressed in some specific applications, particularly robot fleets, and only few recent theoretical works have initiated a more systematic system-theoretic study of sparsity-constrained system realization theory and of sparsity-constrained feedback control.

- Clustering control for large networks. For large and complex systems composed of several sub-networks, feedback design is usually treated at the sub-network level, and most of the times without taking into account natural interconnections between sub-networks. The team is exploring new control strategies, exploiting the emergent behaviors resulting from new interconnections between the network components. This requires first to build network models operating in aggregated clusters, and then to re-formulate problems where the control can be designed using the cluster boundaries rather than individual control loops inside of each network. Examples can be found in the transportation application domain, where a significant challenge will be to obtain dynamic partitioning and clustering of heterogeneous networks in homogeneous sub-networks, and then to control the perimeter flows of the clusters to optimize the network operation. This topic is at the core of the Advanced ERC project Scale-FreeBack.

3.5. Transportation networks

This is currently the main application domain of the NECS team. Several interesting problems in this area capture many of the generic networks problems identified before (e.g., decentralized/collaborative traffic optimal control, density balancing using consensus concepts, data fusion, distributed estimation, etc.). Several specific actions have been continued/launched to this purpose: improvement and finalization of the Grenoble Traffic Lab(GTL), EU projects (SPEEDD, ERC-AdG Scale-FreeBack). Further research goals are envisioned, such as:

- Modeling of large scale traffic systems. We aim at reducing the complexity of traffic systems modeling by engaging novel modeling techniques that make use of clustering for traffic networks while relying on its specific characteristics. Traffic networks will be aggregate into clusters and the main traffic quantities will be extrapolated by making use of this aggregation. Moreover, we are developing an extension of the Grenoble Traffic Lab (GTL) for downtown Grenoble which will make use of GPS and probe data to collect traffic data in the city center.
- Modeling and control of intelligent transportation systems. We aim at developing a complete micro-macro modeling approach to describe and model the new traffic dynamics that is developing thanks to mixed (simple, connected and automated) vehicles in the roads. This will require cutting edge mathematical theory and field experiments.

4. Application Domains

4.1. A large variety of application domains

Sensor and actuator networks are ubiquitous in modern world, thanks to the advent of cheap small devices endowed with communication and computation capabilities. Potential application domains for research in networked control and in distributed estimation are extremely various, and include the following examples.

- Intelligent buildings, where sensor information on CO_2 concentration, temperature, room occupancy, etc. can be used to control the heating, ventilation and air conditioning (HVAC) system under multi-objective considerations of comfort, air quality, and energy consumption.
- Smart grids: the operation of electrical networks is changing from a centralized optimization framework towards more distributed and adaptive protocols, due to the high number of small local energy producers (e.g., solar panels on house roofs) that now interact with the classic large power-plants.
- Disaster relief operations, where data collected by sensor networks can be used to guide the actions of human operators and/or to operate automated rescue equipment.
- Surveillance using swarms of Unmanned Aerial Vehicles (UAVs), where sensor information (from sensors on the ground and/or on-board) can be used to guide the UAVs to accomplish their mission.

- Environmental monitoring and exploration using self-organized fleets of Autonomous Underwater Vehicles (AUVs), collaborating in order to reach a goal such as finding a pollutant source or tracing a seabed map.
- Infrastructure security and protection using smart camera networks, where the images collected are shared among the cameras and used to control the cameras themselves (pan-tilt-zoom) and ensure tracking of potential threats.

In particular, NECS team is currently focusing in the areas described in detail below.

4.2. Intelligent transportation systems

Throughout the world, roadways are notorious for their congestion, from dense urban network to large freeway systems. This situation tends to get worse over time due to the continuous increase of transportation demand whereas public investments are decreasing and space is lacking to build new infrastructures. The most obvious impact of traffic congestion for citizens is the increase of travel times and fuel consumption. Another critical effect is that infrastructures are not operated at their capacity during congestion, implying that fewer vehicles are served than the amount they were designed for. Using macroscopic fluid-like models, the NECS team has initiated new researches to develop innovative traffic management policies able to improve the infrastructure operations. The research activity is on two main challenges: (1) modeling and forecasting, so as to provide accurate information to users, e.g., travel times; and (2) control, via ramp-metering and/or variable speed limits. The Grenoble Traffic Lab (see http://necs.inrialpes.fr/pages/grenoble-traffic-lab.php) is an experimental platform, collecting traffic infrastructure information in real time from Grenoble South Ring, together with innovative software e.g. for travel-time prediciton, and a show-case where to graphically illustrate results to the end-user. This activity is done in close collaboration with local traffic authorities (DIR-CE, CG38, La Metro), and with the start-up company Karrus (http://www.karrus-its.com/)

4.3. Inertial navigation

The team is exploring techniques and approaches from estimation, filtering and machine learning, in order to use inertial sensor units in pedestrian navigation, attitude estimation, augmented reality and human activities recognition. These units are composed of accelerometers, magnetometers and gyroscopes, sensors that we find usually in smartphones. This area of research in the team will evolve towards multimodal navigation, cooperative and collaborative navigation in indoor and outdoor environments.

5. Highlights of the Year

5.1. Highlights of the Year

- The team organized the international ERC Scale-FreeBack workshop on “Analysis and Control of Large-Scale Complex Networks”, Grenoble, September 10-11th, 2018 (http://scale-freeback.eu/workshop-on-analysis-and-control-of-large-scale-complex-networks-10-11-sept-2018-grenoble/)
- P. Frasca is Senior Member of the IEEE

6. New Software and Platforms

6.1. GTL

Grenoble Traffic Lab

FUNCTIONAL DESCRIPTION: The Grenoble Traffic Lab (GTL) initiative, led by the NeCS team, is a real-time traffic data Center (platform) that collects traffic road infrastructure information in real-time with minimum latency and fast sampling periods. The main elements of the GTL are: a real-time data-base, a show room, and a calibrated micro-simulator of the Grenoble South Ring. Sensed information comes from a dense wireless sensor network deployed on Grenoble South Ring, providing macroscopic traffic signals such as flows, velocities, densities, and magnetic signatures. This sensor network was set in place in collaboration with Inria spin-off Karrus-ITS, local traffic authorities (DIR-CE, CG38, La Metro), and specialized traffic research centers. In addition to real data, the project also uses simulated data, in order to validate models and to test the ramp-metering, the micro-simulator is a commercial software (developed by TSS AIMSUN ©). More details at http://necs.inrialpes.fr/pages/grenoble-traffic-lab.php

- Participants: Alain Kibangou, Andres Alberto Ladino Lopez, Anton Andreev, Carlos Canudas-De-Wit, Dominik Pisarski, Enrico Lovisari, Fabio Morbidi, Federica Garin, Hassen Fourati, Iker Bellicot, Maria laura Delle monache, Paolo Frasca, Pascal Bellemain, Pietro Grandinetti, Rémi Piotaix, Rohit Singhal and Vadim Bertrand
- Contact: Carlos Canudas-De-Wit
- URL: http://necs.inrialpes.fr/pages/grenoble-traffic-lab.php

6.2. Benchmarks Attitude Smartphones

KEYWORDS: Experimentation - Motion analysis - Sensors - Performance analysis - Smartphone
SCIENTIFIC DESCRIPTION: We investigate the precision of attitude estimation algorithms in the particular context of pedestrian navigation with commodity smartphones and their inertial/magnetic sensors. We report on an extensive comparison and experimental analysis of existing algorithms. We focus on typical motions of smartphones when carried by pedestrians. We use a precise ground truth obtained from a motion capture system. We test state-of-the-art attitude estimation techniques with several smartphones, in the presence of magnetic perturbations typically found in buildings. We discuss the obtained results, analyze advantages and limits of current technologies for attitude estimation in this context. Furthermore, we propose a new technique for limiting the impact of magnetic perturbations with any attitude estimation algorithm used in this context. We show how our technique compares and improves over previous works.

- Participants: Hassen Fourati, Nabil Layaïda, Pierre Genevès and Thibaud Michel
- Partner: GIPSA-Lab
- Contact: Pierre Genevès
- URL: http://tyrex.inria.fr/mobile/benchmarks-attitude/

7. New Results

7.1. Network systems: modeling, analysis, and estimation

7.1.1. Network reduction towards a scale-free structure preserving physical properties

Participants: N. Martin, P. Frasca, C. Canudas de Wit [Contact person].

In the context of the ERC project, we are addressing a problem of graph reduction, where a given arbitrary weighted graph is reduced to a (smaller) scale-free graph while preserving a consistency with the initial graph and some physical properties. This problem can be formulated as a minimization problem. We give specifications to this general problem to treat a particular case: to this end we define a metric to measure the scale-freeness of a graph and another metric to measure the similarity between two graphs with different dimensions, based on a notion of spectral centrality. Moreover, through the reduction we also preserve a property of mass conservation (essentially, Kirchoff's first law). We study the optimization problem and, based on the gained insights, we derive an algorithm allowing to find an approximate solution. Finally, we have simulated the algorithm both on synthetic networks and on real-world examples of traffic networks that represent the city of Grenoble. These results are presented in [57] and in [31]. We also developed an application to the control of epidemics [58].

7.1.2. Cyber-Physical Systems: a control-theoretic approach to privacy and security

Participants: F. Garin [Contact person], A. Kibangou, S. Gracy.

Cyber-physical systems are composed of many simple components (agents) with interconnections giving rise to a global complex behaviour. Interesting recent research has been exploring how the graph describing interactions affects control-theoretic properties such as controllability or observability, namely answering the question whether a small group of agents would be able to drive the whole system to a desired state, or to retrieve the state of all agents from the observed local states only.

A related problem is observability in the presence of an unknown input, where the input can represent a failure or a malicious attack, aiming at disrupting the normal system functioning while staying undetected. We study linear network systems, and we aim at characterizing input and state observability (ISO), namely the conditions under which both the whole network state and the unknown input can be reconstructed from some measured local states. We complement the classical algebraic characterizations with novel structural results, which depend only on the graph of interactions (equivalently, on the zero pattern of the system matrices). More precisely, we obtain two kinds of results (see [24], [25] and the PhD thesis of S. Gracy): structural results, true for almost all interaction weights, and strongly structural results, true for all non-zero interaction weights. We consider both the case where the system graph is time-invariant, and the case where it varies in time.

When the conditions for ISO are satisfied, one can run algorithms in the same vein as a Kalman filter, in order to reconstruct the state and the unknown input from noisy measurements. These algorithms are known for the case where the input can be reconstructed with only one time-step of delay with respect to the measurements; in [54] we propose a (suboptimal) filter for the case when this is not possible, i.e., more measurements are needed for the input reconstruction.

7.1.3. Heterogeneity and uncertainty in distributed estimation from relative measurements

Participants: C. Ravazzi, N. K. Chan, P. Frasca [Contact person].

This work, presented in [34], has studied the problem of estimation from relative measurements in a graph, in which a vector indexed over the nodes has to be reconstructed from pairwise measurements of differences between its components associated to nodes connected by an edge. In order to model heterogeneity and uncertainty of the measurements, we assume them to be affected by additive noise distributed according to a Gaussian mixture. In this original setup, we formulate the problem of computing the Maximum-Likelihood (ML) estimates and we design two novel algorithms, based on Least Squares regression and Expectation-Maximization (EM). The first algorithm (LSEM) is centralized and performs the estimation from relative measurements, the soft classification of the measurements, and the estimation of the noise parameters. The second algorithm (Distributed LS-EM) is distributed and performs estimation and soft classification of the measurements, but requires the knowledge of the noise parameters. We provide rigorous proofs of convergence for both algorithms and we present numerical experiments to evaluate their performance and compare it with solutions from the literature. The experiments show the robustness of the proposed methods against different kinds of noise and, for the Distributed LS-EM, against errors in the knowledge of noise parameters.

7.1.4. Average state estimation in large-scale multi-cluster networks

Participants: U. Niazi, A. Kibangou, C. Canudas de Wit [Contact person].

In the context of the ERC project, we are addressing the problem of estimation of a functional of non-observed states. Indeed, large-scale network systems can be unobservable from the dedicated state measurements at few nodes. By resorting to an aggregation of multiple clusters of unmeasured nodes, we are investigating the observability and detectability of average states of the clusters. The approach is to obtain a reduced network system whose state vector contains the average states of the clusters. The notion of average observability is defined with respect to the observability of this reduced network system. For average observability, we have stated a necessary condition and a sufficient condition depending solely on the structure of the network. Average detectability, which is a milder notion than average observability, is also studied and a sufficient condition, under which an open-loop average state observer converges, is provided. This condition requires clusters of unmeasured nodes to have negatively balanced local outflow centrality.

7.2. Control of multi-agent systems and opinion dynamics

7.2.1. Open multi-agent systems: Dynamic consensus

Participants: W. S. Rossi, P. Frasca [Contact person].

In [53] we investigate a dynamic consensus problem for an open multi-agent system. Open multi-agent systems are characterized by a time-varying set of agents connected by a network: agents may leave and new agents may join the network at any time, thus the term "open". The dynamic consensus problem consists in achieving agreement about the time-varying average of a set of reference signals that are assumed to be the agents' inputs. Dynamic consensus has recently found application in the context of distributed estimation for electric demand-side management, where a large population of connected domestic appliances needs to estimate its future average power consumption. Since the considered network of devices changes as new appliances log in and out, there is a need to develop and characterize dynamic consensus algorithms for these open scenarios. In this paper we give several initial contributions both to a general theory of open multi-agent systems and to the specific problem of dynamic consensus within this context. On the theoretical side, we propose a formal definition of open multi-agent system, a suitable notion of stability, and some sufficient conditions to establish it. On the applied side, we design a novel dynamic consensus algorithm, the Open Proportional Dynamic Consensus algorithm. We characterize some of its convergence properties in the proposed open-multi-agent systems framework and we illustrate its evolution by numerical simulations.

7.2.2. Robust average consensus over unreliable networks

Participants: F. Acciani, P. Frasca [Contact person], G. Heijenk, A. Stoorvogel.

Packet loss is a serious issue in wireless consensus networks, as even few failures might prevent a network to converge to the desired consensus value. In the last four years, we have devised some possible ways to compensate for the errors caused by packet collisions, by modifying the updating weights. Since these modifications may result in a reduced convergence speed, a gain parameter is used to increase the convergence speed, and an analysis of the stability of the network is performed, leading to a criterion to choose such gain to guarantee network stability. For the implementation of the compensation method, we propose a new communication algorithm, which uses both synchronous and asynchronous mechanisms to achieve average consensus and to deal with uncertainty in packet delivery. The paper [14] provides a complete account of our results.

7.2.3. Asynchronous opinion dynamics on the k-nearest-neighbors graph

Participants: W. S. Rossi, P. Frasca [Contact person].

This work is about a new model of opinion dynamics with opinion-dependent connectivity. We assume that agents update their opinions asynchronously and that each agent's new opinion depends on the opinions of the k agents that are closest to it. In the paper [63], we show that the resulting dynamics is substantially different from comparable models in the literature, such as bounded-confidence models. We study the equilibria of the dynamics, observing that they are robust to perturbations caused by the introduction of new agents. We also prove that if the number of agents n is smaller than 2k, the dynamics converge to consensus. This condition is only sufficient.

7.2.4. Quantization effects in opinion dynamics

Participants: F. Ceragioli, P. Frasca [Contact person].

This work deals with continuous-time opinion dynamics that feature the interplay of continuous opinions and discrete behaviors. In our model, the opinion of one individual is only influenced by the behaviors of fellow individuals. The key technical difficulty in the study of these dynamics is that the right-hand sides of the equations are discontinuous and thus their solutions must be intended in some generalized sense: in our analysis, we consider both Carathéodory and Krasovskii solutions. We first prove the existence and completeness of Carathéodory solutions from every initial condition and we highlight a pathological behavior of Carathéodory solutions, which can converge to points that are not (Carathéodory) equilibria. Notably, such points can be arbitrarily far from consensus and indeed simulations show that convergence to nonconsensus configurations is common. In order to cope with these pathological attractors, we study Krasovskii solutions. We give an estimate of the asymptotic distance of all Krasovskii solutions from consensus and we prove its tightness by an example of equilibrium such that this distance is quadratic in the number of agents. This fact implies that quantization can drastically destroy consensus. However, consensus is guaranteed in some special cases, for instance, when the communication among the individuals is described by either a complete or a complete bipartite graph. These results are reported in details in [19], whereas the book chapter [66] puts them in the broader context of consensus-seeking dynamics with discontinuous right-hand side.

7.2.5. Message-passing computation of harmonic influence in social networks

Participants: W. S. Rossi, P. Frasca [Contact person].

In the study of networks, identifying the most important nodes is of capital importance. The concept of Harmonic Influence has been recently proposed as a metric for the importance of nodes in a social network. This metric evaluates the ability for one node to sway the opinions of the other nodes in the network, under the assumption of a linear diffusion of opinions in the network. A distributed message passing algorithm for its computation has been proposed by Vassio et al., 2014, but its convergence guarantees were limited to trees and regular graphs. In [36], we prove that the algorithm converges on general graphs. In [64], we offer two additional contributions to its study. We evaluate how the presence of communities in the network impacts the algorithm performance, and how the algorithm performs on networks which change topology during the execution of the algorithm.

7.2.6. Distributed control and game theory: self-optimizing systems

Participants: F. Garin [Contact person], B. Gaujal [POLARIS], S. Durand.

The design of distributed algorithms for a networked control system composed of multiple interacting agents, in order to drive the global system towards a desired optimal functioning, can benefit from tools and algorithms from game theory. This is the motivation of the Ph.D. thesis of Stéphane Durand, a collaboration between POLARIS and NECS teams.

The focus of this thesis is on the complexity of the best response algorithm to find a Nash equilibrium for potential games. Best response is a simple greedy algorithm, known to converge to a Nash equilibrium if players play one after the other in a round-robin way, but with a worst-case complexity which is exponential in the number of players. We consider instead its average complexity over the ensemble of random potential games, showing that such average complexity is surprisingly low, only linear in the number of players. Then we focus on removing the need of a centralised scheduler enforcing the round robin order of play. In [52], [21] we consider agents activated according to independent local Poisson clocks, and we show that (despite the possible overlaps of the computations of some players), we can still obtain convergence, with an average complexity of order $n \log n / \log \log n$, where n is the number of players. In [51] we show how to take advantage of the structure of the interactions between players in a network game: noninteracting players can play simultaneously. This improves best response algorithm, both in the centralized and in the distributed case.

7.2.7. Control of switched interconnected large-scale systems

Participants: H. Fourati [Contact person], D. Belkhiat, D. Jabri.

We proposed in [27] a new design of a decentralized output-feedback tracking control for a class of switched large-scale systems with external bounded disturbances. The controller proposed herein is synthesized to satisfy the robust H_∞ tracking performance with local disturbance attenuation levels. Based on multiple switched Lyapunov functions, sufficient conditions proving the existence of the proposed controller are formulated in terms of Linear Matrix Inequalities (LMI).

7.3. Transportation networks and vehicular systems

7.3.1. Density and flow reconstruction in urban traffic networks

Participants: C. Canudas de Wit [Contact person], H. Fourati, A. Kibangou, A. Ladino, M. Rodriguez-Vega.

In [56], we consider the problem of joint reconstruction of flow and density in a urban traffic network using heterogeneous sources of information. The traffic network is modeled within the framework of macroscopic traffic models, where we adopt Lighthill-Whitham-Richards model (LWR) conservation equation characterized by a piecewise linear fundamental diagram. The estimation problem considers two key principles. First, the error minimization between the measured and reconstructed flows and densities, and second the equilibrium state of the network which establishes flow propagation within the network. Both principles are integrated together with the traffic model constraints established by the supply/demand paradigm. Finally the problem is cast as a constrained quadratic optimization with equality constraints in order to shrink the feasible region of estimated variables. Some simulation scenarios based on synthetic data for a manhattan grid network are provided in order to validate the performance of the proposed algorithm.

In [62], we addressed the conditions imposed on the number and location of fixed sensors such that all flows in the network can be uniquely reconstructed. We determine the minimum number of sensors needed to solve the problem given partial information of turning ratios, and then we propose a linear time algorithm for their allocation in a network. Using these results in addition to floating car data, we propose a method to reconstruct all traffic density and flow.

7.3.2. Discrete-time system optimal dynamic traffic assignment (SO-DTA) with partial control for horizontal queuing networks

Participants: S. Samaranayake, J. Reilly, W. Krichene, M. L. Delle Monache [Contact person], P. Goatin [Acumes, Inria], A. Bayen.

Dynamic traffic assignment (DTA) is the process of allocating time-varying origin-destination (OD) based traffic demand to a set of paths on a road network. There are two types of traffic assignment that are generally considered, the user equilibrium or Wardrop equilibrium allocation (UE-DTA), in which users minimize individual travel-time in a selfish manner, and the system optimal allocation (SODTA) where a central authority picks the route for each user and seeks to minimize the aggregate total travel-time over all users. It can be shown that the price of anarchy (PoA), the worst-case ratio of the system delay caused by the selfish behavior over the system optimal solution, may be arbitrarily large even in simple networks. System optimal (SO) traffic assignment on the other hand leads to optimal utilization of the network resources, but is hard to achieve in practice since the overriding objective for individual drivers in a road network is to minimize their own travel-time. It is well known that setting a toll on each road segment corresponding to the marginal delay of the demand moves the user equilibrium towards a SO allocation. In [37], we formulate the system optimal dynamic traffic assignment problem with partial control (SO-DTAPC), using a Godunov discretization of the Lighthill-Williams-Richards (LWR) partial differential equation (PDE) with a triangular flux function. We propose solving the SO-DTA-PC problem with the non-convex traffic dynamics and limited OD data with complete split ratios as a non-linear optimal control problem. This formulation generalizes to multiple sources and multiple destinations. We show that the structure of our dynamical system allows for very efficient computation of the gradient via the discrete adjoint method.

7.3.3. Priority-based Riemann solver for traffic flow on networks

Participants: M. L. Delle Monache [Contact person], P. Goatin [Acumes, Inria], B. Piccoli.

In [20] we introduce a novel solver for traffic intersection which considers priorities among the incoming roads as the first criterion and maximization of flux as the second. The main idea is that the road with the highest priority will use the maximal flow taking into account also outgoing roads constraints. If some room is left for additional flow then the road with the second highest priority will use the left space and so on. A precise definition of the new Riemann solver, called Priority Riemann Solver, is based on a traffic distribution matrix , a priority vector and requires a recursion method. The general existence theorem for Riemann solvers on junctions can not be applied in the present case.Therefore, we achieve existence via a new set of general properties.

7.3.4. Dissipation of stop-and-go waves via control of autonomous vehicles

Participants: R. Stern, S. Cui, M. L. Delle Monache [Contact person], R. Bhadani, M. Bunting, M. Churchill, N. Hamilton, R. Haulcy, H. Pohlmann, F. Wu, B. Piccoli, B. Seibold, J. Sprinkle, D. B. Work.

Traffic waves are phenomena that emerge when the vehicular density exceeds a critical threshold. Considering the presence of increasingly automated vehicles in the traffic stream, a number of research activities have focused on the influence of automated vehicles on the bulk traffic flow. In [38], we demonstrate experimentally that intelligent control of an autonomous vehicle is able to dampen stop-and-go waves that can arise even in the absence of geometric or lane changing triggers. Precisely, our experiments on a circular track with more than 20 vehicles show that traffic waves emerge consistently, and that they can be dampened by controlling the velocity of a single vehicle in the flow. We compare metrics for velocity, braking events, and fuel economy across experiments. These experimental findings suggest a paradigm shift in traffic management: flow control will be possible via a few mobile actuators (less than 5%) long before a majority of vehicles have autonomous capabilities.

7.3.5. Cooperative adaptive cruise control over unreliable networks

Participants: F. Acciani, P. Frasca [Contact person], G. Heijenk, E. Semsar-Kazerooni, A. Stoorvogel.

Cooperative Adaptive Cruise Control (CACC) is a promising technique to increase highway throughput, safety and comfort for vehicles. Enabled by wireless communication, CACC allows a platoon of vehicles to achieve better performance than Adaptive Cruise Control; however, since wireless is employed, problems related to communication unreliability arise. In [45], we design a digital controller to achieve platoon stability, enhanced by an observer to increase robustness against packet losses. Our results confirms the interest of using an observer in combination with a local and cooperative digital controller.

7.3.6. Heterogeneity in synchronization: an adaptive control approach, with applications to vehicle platooning

Participants: S. Baldi, P. Frasca [Contact person].

Heterogeneity is a substantial obstacle to achieve synchronisation of interconnected systems (that is, in control) In order to overcome heterogeneity, advanced control techniques are needed, such as the use of “internal models” or of adaptive techniques. In a series of papers motivated by multi-vehicle platooning and coordinated autonomous driving, we have explored the application of adaptive control techniques. Our results cover both the cases of state-feedback [15] and of output-feedback [16], under the assumption that the topology of the interconnections has no circuits. Further investigation has shown that restrictive assumption can be relaxed (at least for state-feedback on some specific topologies) [47]. This understanding paves the road to use these techniques not only to stabilise heterogeneous platoons, but also to manage their merging or splitting operations [48].

7.3.7. Modeling traffic on roundabout

Participants: M. L. Delle Monache [Contact person], A. Rat, S. Hammond, B. Piccoli.

In [50] we introduce a Riemann solver for traffic flow on a roundabout with two lanes. The roundabout is modeled as a sequence of junctions. The Riemann solver provides a solution at junctions by taking into consideration traffic distribution, priorities, and the maximization of through flux. We prove existence and uniqueness of the solution of the Riemann problem and show some results numerically. This work stems from the fact that there is a general notion among transportation professionals that having a longer additional lane length at a double-lane roundabout entry yields better performances. In [55], we investigate this notion using Lighthill-Whitham-Richards Model. Using Lighthill- Whitham-Richards model, a double-lane roundabout with additional lane design at the entry is analyzed. The additional lane lengths are varied at the entry in order to study the effect of different additional lane lengths on roundabout performance. The results obtained with the PDE model were then compared with similar lane length variations in VISSIM.

7.3.8. Two dimensional models for traffic

Participants: S. Mollier, M. L. Delle Monache, C. Canudas de Wit [Contact person], B. Seibold.

The work deals with the problem of modeling traffic flow in urban area, e. g. a town. More precisely, the goal is to design a two-dimensional macroscopic traffic flow model suitable to model large network as the one of a city. Macroscopic traffic models are inspired from fluid dynamic. They represent vehicles on the road by a density and describe their evolution with partial differential equations. Usually, these models are one dimensional models and, for instance, give a good representation of the evolution of traffic states in highway. The extension of these 1D models to a network is possible thanks to models of junction but can be tedious according to the number of parameters to fit. In the last few years, the idea of models based on a two dimensional conservation laws arose in order to represent traffic flow in large and dense networks. This study starts with a simple model [33] for homogeneous network and where a preferred direction of traffic exists. Our aim is to extend gradually this model by adding complexity. As this approach is uncommon, we investigate a way to compare the results of this model with microsimulation in [73] using Aimsun. Then, in the literature, the network is mainly assumed to be homogeneous. However, in a large-scale scenario, it is unlikely that the traffic network characteristics–such as speed limit, number of lanes, or the network geometry–remain constant throughout the network. Therefore, we introduce a first extension [59] where the fundamental diagram is space-dependent and varies with respect to the area considered. Finally, we have studied more recently a possible way to relax the assumption of a preferred direction of flow by considering several layers of density such that each layer describe a different direction of flow. In this case, the model becomes a system of conservation and is hyperbolic-elliptic which imply special caution in the choice of the numerical method.

7.4. Multisensor data fusion for navigation

7.4.1. Sensors fusion for attitude estimation

Participants: H. Fourati [Contact person], Z. Zhou, J. Wu.

Attitude estimation consists in the determination of rigid body orientation in 3D space (principally in terms of Euler angles, rotation matrix, or quaternion). As a key problem for multisensor attitude determination, Wahba's problem has been studied for almost 50 years. In [42], we present a novel linear approach to solve this problem. We name the proposed method the fast linear attitude estimator (FLAE) because it is faster than known representative algorithms. The original Wahba's problem is extracted to several 1-D equations based on quaternions. They are then investigated with pseudoinverse matrices establishing a linear solution to n-D equations, which are equivalent to the conventional Wahba's problem. To obtain the attitude quaternion in a robust manner, an eigenvalue-based solution is proposed. Symbolic solutions to the corresponding characteristic polynomial are derived, showing higher computation speed. Also, to verify the feasibility in embedded application, an experiment on the accelerometer–magnetometer combination is carried out where the algorithms are compared via C++ programming language. From other side, the integration of the Accelerometer and Magnetometer (AM) provides continuous, stable and accurate attitude information for land-vehicle navigation without magnetic distortion and external acceleration. However, magnetic disturbance and linear acceleration strongly degrade the overall system performance. As an important complement, the Global Navigation Satellite System (GNSS) produces the heading estimates, thus it can potentially benefit the AM system. Such a GNSS/AM system for attitude estimation is mathematically converted to a multi-observation vector pairs matching problem in [44]. The optimal and sub-optimal attitude determination and their time-varying recursive variants are all comprehensively investigated and discussed. The developed methods are named as the Optimal Linear Estimator of Quaternion (OLEQ), Suboptimal-OLEQ (SOLEQ) and Recursive-OLEQ (ROLEQ) for different application scenarios. The theory is established based on our previous contributions, and the multi-vector matrix multiplications are decomposed with the eigenvalue factorization. Some analytical results are proven and given, which provides the reader with a brand new viewpoint of the attitude determination and its evolution. With the derivations of the two-vector case, the n-vector case is then naturally formed. The algorithms are then implemented using the C++ programming language on the designed hardware with a GNSS module, three-axis accelerometer and three-axis magnetometer, giving an effective validation of them in real-world applications. In [39], a super fast attitude solution is obtained for consumer electronics accelerometer-magnetometer combination. The quaternion parameterizing the orientation is analytically derived from a least-square optimization that maintains very simple form. Like previously developed approaches, this algorithm does not require predetermined magnetometer reference vector. In [41], we present a novel sequential multiplicative quaternion attitude estimation method from various vector sensor outputs. The unique linear constitution of the algorithm leads to its specific name of Recursive Linear Quaternion Estimator (RLQE). The algorithm's architecture is designed to use each single pair of vector observation linearly so that the vector observations can be arbitrarily chosen and fused. The closed-form covariance of the RLQE is derived that builds up the existence of a highly reliable RLQE Kalman filter (RLQE-KF). In [65], to generate the virtual-gyro output in the case of gyroscope failures, virtual-gyro Kalman filter is established for angular rate estimation base on attitude estimation results.

7.4.2. *Attitude estimation applied in augmented reality*

Participants: H. Fourati [Contact person], T. Michel, P. Genevès, N. Layaïda.

We investigate the precision of attitude estimation algorithms in the particular context of pedestrian navigation with commodity smartphones and their inertial/magnetic sensors. A particular attention was paid to the study of attitude estimation in the context of augmented reality motions when using smartphones [32]. We report on an extensive comparison and experimental analysis of existing algorithms. We focus on typical motions of smartphones when carried by pedestrians. We use a precise ground truth obtained from a motion capture system. We test state-of-the-art and built-in attitude estimation techniques with several smartphones, in the presence of magnetic perturbations typically found in buildings. We discuss the obtained results, analyze advantages and limits of current technologies for attitude estimation in this context. Furthermore, we propose a new technique for limiting the impact of magnetic perturbations with any attitude estimation algorithm used in this context.

7.4.3. *Attitude determination for satellite*

Participants: H. Fourati [Contact person], S. Pourtakdoust, Csug Team, E. Kerstel.

Recently, we started to work on attitude estimation for satellites. In [29], we are focused on the development and verification of a heat attitude model (HAM) for satellite attitude determination. Within this context, the Sun and the Earth are considered as the main external sources of radiation that could affect the satellite surface temperature changes. Assuming that the satellite orbital position (navigational data) is known, the proposed HAM provides the satellite surface temperature with acceptable accuracy and also relates the net heat flux (NHF) of three orthogonal satellite surfaces to its attitude via the inertial to satellite transformation matrix. The proposed HAM simulation results are verified through comparison with commercial thermal analysis tools. The proposed HAM has been successfully utilized in some researches for attitude estimation, and further studies for practical implementations are still ongoing. Actually, we are establishing a project around quantum communication experiments under Nanobob CubeSat mission [28]. Some attitude estimation algorithms will be deployed to orient the satellite to the ground station.

7.4.4. *Sensors fusion for distance measurement in pedestrian navigation*

Participants: H. Fourati [Contact person], Z. Zhou, J. Wu.

We developed in [43] a foot-mounted pedestrian navigation system prototype with the emphasis on distance measuring with an inertial measurement unit (IMU) which implies the characteristics of pedestrian gait cycle and thus can be used as a crucial step indicator for distance calculation. An adaptive time- and frequency-domains joint distance measuring method is proposed by utilizing the means of behaviors classification. Two key issues are studied: step detection and step length determination. For the step detection part, first behavior classification along with state transition strategy is designed to identify typical pedestrian behaviors including standing still, walking, running and irregular swing. Then a four-stage step detection method is proposed to adaptively determine both step frequency and threshold in a flexible window. Based on the behavior classification results, a two-segment functional based step length model is established to adapt the walking and running behaviors.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. *Control of Cyber-Social Systems (C2S2)*

C2C2 is a two-year project funded by the University Grenoble Alpes, MSTIC department. Evolving from recent research on network systems, this exploratory project has the objective to concentrate on cyber-social systems, that is, complex systems with interacting social and technological components. A strong motivation for this novel research direction comes from the need for innovative tools for the management of vehicular traffic. In this application, state-of-the-art approaches concentrate on hard control actions, like traffic lights: instead, future management methods should exploit soft control actions aimed at controlling the traffic demand, that is, the aggregated behaviors of the drivers.

8.1.2. *Understanding data accidents for traffic safety (DATASAFE)*

DATASAFE is a two years project funded by Grenoble Data Institute, with the aim to understand from real traffic data the behavior of traffic in the moments preceding an accident. The general approach is to use novel statistical techniques in order to learn traffic characteristics that can be used to develop new traffic models. Bayesian approaches are used to (supervised) classification and (unsupervised) clustering in order to respectively predict collision occurrences and discover traffic patterns.

8.1.3. *Modeling autonomous vehicles in traffic flow (MAVIT)*

MAVIT is a two year project funded by the University Grenoble Alpes, MSTIC department. The goal of this project is to develop a unified micro-macro approach for traffic management, involving human and autonomous vehicles drivers by providing analytical and numerical tools for traffic modeling, estimation and control. We will work towards field operational tests, by using instrumented cars to collect data on AVs trajectory and their interaction with the traffic flow with human drivers. The proposed research provides new mathematical models, computational/software tools, and engineering solutions for the control of human controlled vehicles via intelligently controlled AVs in the traffic stream. Moreover, the control of traffic via moving actuators provides a new alternative to contemporary control technologies, such as ramp metering and variable speed limits; even when AVs comprise a tiny fraction of the total fleet, these techniques may be viable, and rapidly configurable. This research considers new types of traffic models, new control algorithms for traffic flow regulation, and new sensing and control paradigms that are enabled by a small number of controllable systems anticipated in a flow. Specifically, the research focuses on new (1) micro-macro models to model few AVs in a flow; (2) estimation techniques for AV interactions with the traffic flow; (3) developing and assessing dynamical controllers to mitigate traffic events.

8.1.4. *NanoSatellite Project: Advanced modelling and Control of attitude dynamics for quantum communication (SPACE)*

SPACE is a two-year project funded by the IDEX University Grenoble Alpes. It aims to launch an exploratory study to find the required minimal data we need to collect and combine for software design of Nanosatellite Attitude Determination and Control System (ADCS).

8.1.5. *caPture de mOuvements humainS par cenTrales inertielles/d'attitUde et smaRtphonEs : Vers l'analyse d'anomalies neurologiques et fonctions motrices (POSTURE)*

POSTURE is a one-year project funded by CDP NeuroCog. The project is focused on 1) the identification and characterization of postures and reference movements in humans using appropriate algorithms of classification (machine learning), and 2) the analysis of the effects of the number, location and orientation of the inertial sensors on the performance of the methods of identification and classification of postures and movements.

8.2. National Initiatives

8.2.1. Models of Bubbles in Online Social Networks (MOB)

MOB is a PEPS S2IH INS2I 2018 interdisciplinary project. This exploratory project focuses on the effects of online recommendation systems on social dynamics, which may entail the formation of «filter bubbles» that distort the experience of the users. The project will develop a mathematical model to demonstrate these effects and propose designs for their mitigation. The research will be conducted by a blend of tools from dynamical systems, network theory, complex systems, and control systems.

8.2.2. AgileWorld-MRSEI

AgileWorld is an ANR-MRSEI project (2018-2020), which aims at building an European network for an innovative training on road transportation systems in a connected world. The funding will help to prepare and then submit a proposal for the MSCA-ITN 2019 call.

8.3. European Initiatives

8.3.1. Collaborations in European Programs, Except FP7 & H2020

8.3.1.1. Scale-FreeBack

Type: ERC Advanced Grant

Duration: Sep. 2016 to Aug. 2021

Coordinator: C. Canudas de Wit

Inria contact: C. Canudas de Wit

Abstract: The overall aim of Scale-FreeBack is to develop holistic scale-free control methods of controlling complex network systems in the widest sense, and to set the foundations for a new control theory dealing with complex physical networks with an arbitrary size. Scale-FreeBack envisions devising a complete, coherent design approach ensuring the scalability of the whole chain (modelling, observation, and control). It is also expected to find specific breakthrough solutions to the problems involved in managing and monitoring large-scale road traffic networks. Field tests and other realistic simulations to validate the theory will be performed using the equipment available at the Grenoble Traffic Lab center (see GTL), and a microscopic traffic simulator replicating the full complexity of the Grenoble urban network.

See also: http://scale-freeback.eu

8.4. International Initiatives

8.4.1. Inria Associate Teams Not Involved in an Inria International Lab

8.4.1.1. MEMENTO

Title: ModEling autonoMous vEhicles iN Traffic flOw

International Partner:

Vanderbilt University, Nashville (United States) - Dan Work

Start year: 2018

See also: http://necs.inrialpes.fr/memento/index.html

In recent years, the strategic priorities of automotive and transportation systems focus on research, development and adoption of automation-related technologies as they emerge. As these technology developments are introduced in the traffic stream, an open question is how the mathematical models that are at the heart of transportation planning and operations will need to be advanced to accommodate these changes. The goal of the NeCS-Vanderbilt, MEMENTO, associate team is to create a multidisciplinary environment to model autonomous vehicles (AV) in human traffic flow. Specifically, our goal is to develop a unified micro-macro approach for traffic management, involving human drivers and autonomous vehicles by providing analytical and numerical tools for traffic modeling, estimation and control. We will work towards field operational tests, by using instrumented cars to collect data on AVs trajectories and their interaction with the traffic flow with human drivers.

8.4.2. Participation in Other International Programs

8.4.2.1. TICO-MED

TicoMed (Traitement du signal, Traitement numérique multidimensionnel de l'Information avec applications aux Télécommunications et au génie Biomédical) is a French-Brazilian project funded by CAPES-COFECUB for the period 2015-2018. It involves University of Nice Sophia Antipolis (I3S Laboratory), CNAM, SUP-ELEC, University of Grenoble Alpes (Gipsa-Lab), Universidade Federal do Ceara, Universidade Federal do Rio de Janeiro, and Universidade Federal do Santa Catarina as partners.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

Prof. Andrea Tosin (Politecnico di Torino, Italy) visited the team in February 2018 in the frame of the Scale-FreeBack ERC project. He gave a talk on "Control strategies for road risk mitigation in kinetic traffic modeling". He exchanged ideas with Carlos Canudas De Wit, Paolo Frasca, Stephane Mollier, Maria Laura Delle Monache and Thibault Liard.

Prof. Sandro Zampieri (Univ. Padova, Italy) visited the team in February 2018 in the frame of the Scale-FreeBack ERC project, to work with Carlos Canudas De Wit and Giacomo Casadei.

Prof. Karl Henrik Johansson (KTH, Stockholm, Sweden) visited the team in March 2018 in the frame of the Scale-FreeBack ERC project, and gave a talk on “Control of vehicle platoons and their influence on traffic”.

Prof. Dan Work (Vanderbilt University (USA)), visited the team in July 2018 to work with Maria Laura Delle Monache and Thibault Liard, in the framework of the associated team MEMENTO.

Mauro Franceschelli (University of Cagliari, Italy) visited the team in October 2018 to collaborate with P. Frasca.

Prof. Olga Lucia Quintero Montoya, Univ. EAFIT, Medellin, Colobmia, visited the team in September 2018. She worked with C. Canudas de Wit on the theoretical development of a normalized macroscopic fundamental diagram for urban traffic.

George Gunter and Raphael Stern (University of Illinois at Urbana-Champaign and Vanderbilt University (USA)) visited the team in November 2018 to work with Maria Laura Delle Monache and Thibault Liard, in the framework of the associated team MEMENTO.

Maolong Lyu is a PhD student from TU Delft (Netherlands) under the supervision of Prof. Simone Baldi. He visited the team for two months to work with M.L. Delle Monache and P. Frasca on string stability for microscopic traffic flow models describing mixed traffic (human drivers and autonomous vehicles).

Diego Deplano is a PhD student from Univ. Cagliari (Italy) under the supervision of Prof. Alessandro Giua. He is visiting the team since Sept. 2018, working with C. Canudas de Wit.

8.5.1.1. Internships

Alexandre Olikier, “Open multi-agent systems with fixed size and possibly not complete topologies”, December 2017–June 2018. Université catholique de Louvain, Belgium. Jointly advised by Paolo Frasca and Julien Hendrickx.

8.5.2. Visits to International Teams

8.5.2.1. Research Stays Abroad

P. Frasca is a Visiting Scientist at the IEIIT-CNR Institute, National Research Council CNR, Torino, Italy. By this collaboration, he performs research on distributed estimation in sensor networks and distributed control of social networks. He visited Torino three times in 2018.

P. Frasca is a Visiting Faculty at the Department of Applied Mathematics, University of Twente, Enschede, The Netherlands. By this collaboration, he performs research on vehicle platooning and on the dynamics of social media. He visited Enschede three times in 2018.

M. L. Delle Monache visited University of Alabama (USA) in April 2018.

M. L. Delle Monache visited Vanderbilt University (USA) in May 2018, and T. Liard visited the same university in May-June and December 2018. These visits are in the frame of the MEMENTO associate team.

F. Garin visited Rutgers University (Philadelphia, USA) in April 2018, to initiate a collaboration with Prof. Benedetto Piccoli and his students on metabolic networks.

A. Kibangou visited the University of Johannesburg (South Africa) in October 2018. During his stay, he gave a lecture to students of Department of Town and Regional Planning of Univ. of Johannesburg on Mobility and traffic management.

N. Martin visited Imura Laboratory at Tokyo University of Technology from June 20th to August 20th, in the frame of the JSPS summer program. The aim of this collaboration is to integrate controllability and/or observability notions in the network reduction problem at the core of this Ph.D. work.

M. U. B. Niazi visited Professor Jacquelien Scherpen at the University of Groningen, Netherlands, in October 2018, to work on model reduction for network systems.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. Member of the Organizing Committees

The team organized the international ERC Scale-FreeBack workshop on “Analysis and Control of Large-Scale Complex Networks”, Grenoble, September 10-11th, 2018 (http://scale-freeback.eu/workshop-on-analysis-and-control-of-large-scale-complex-networks-10-11-sept-2018-grenoble/).

M. L. Delle Monache organized a mini-symposium on “Modélisation et gestion du trafic routier”, 44e Congrès National d’Analyse Numérique, May 2018 (with P. Goatin, Acumes team).

M. L. Delle Monache organized a workshop on “Traffic flow control via PDE techniques”, CDC, December 2018 (with Nikolaos Bekiaris-Liberis, Delphine Bresch-Pietri and Rafael Vazquez).

Team members organized the following invited sessions at the European Control Conference ECC 2018, Cyprus, June 2018:

- “Model reduction and control in large-scale networks” (P. Frasca and C. Canudas de Wit)
- “Multi-agent network games” (F. Garin, with S. Grammatico from Delft Univ. of Technology)

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

C. Canudas de Wit has served as Associate Editor at Large for the American Control Conference ACC 2019.

P. Frasca has served as Associate Editor in the conference editorial boards for the 7th IFAC Workshop on Distributed Estimation and Control in Networked Systems and the 23rd International Symposium on Mathematical Theory of Networks and Systems (MTNS).

F. Garin is Associate Editor in the IEEE Control System Society Conference Editorial Board (this year, she served for CDC 2018, ACC 2019) and Associate Editor in the European Control Association (EUCA) Conference Editorial Board (this year, she served for ECC 2019).

H. Fourati was member of:

- the International Program Committee (IPC) of international conferences STA'18, ICCAD'18, ICITE'18.
- the International Program Committee (IPC)/Associate Editor for contributed papers for the IEEE Conference on Control Technology and Applications (CCTA'18), Copenhagen (Denmark), Aug. 2018;
- the Technical Program Committee (TCP) for the International Conference on Indoor Positioning and Indoor Navigation (IPIN'18), Nantes (France), Sep. 2018;
- the committee of the reviewing phase of the 21st Euro Working Group on Transportation Meeting (EWGT'18), Braunschweig (Germany), Sep. 2018.

9.1.2.2. Reviewer

Team members have been reviewers for several conferences, including the most prestigious ones in their research area: IEEE Conference on Decision and Control CDC, European Control Conference ECC, American Control Conference ACC, European Signal Processing Conference, IEEE International Conference on Robotics and Automation ICRA, IEEE/RSJ International Conference on Intelligent Robots and Systems IROS, IFAC Workshop on Distributed Estimation and Control in Networked Systems (NecSys), IFAC Workshop on Control for Transportartion Systems (CTS), IEEE Intelligent Transportation Systems Society Conference, Transportation Research Board Annual Meeting.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

C. Canudas de Wit is Associate Editor of the IEEE Transactions on Control of Networks Systems IEEE-TCNS (since June 2013) and Editor of the Asian Journal of Control AJC (since 2010).

P. Frasca is Subject Editor of the International Journal of Robust and Nonlinear Control (Wiley) (since February 2014), Associate Editor of the IEEE Control System Letters (from February 2017) and Associate Editor of the Asian Journal of Control (Wiley) (since January 2017).

H. Fourati is Associate Editor of the Asian Journal of Control (Wiley) (since January 2016) and of the Open Transportation Journal https://benthamopen.com/TOTJ/editorial-board. He has also been guest editor of the special issue “Multi-sensor Integrated Navigation and Location based services applications” for International Journal of Distributed Sensor Networks (IJDSN), 2017-2018 (http://journals.sagepub.com/topic/collections-dsn/dsn-1-msinalbsa/dsn) and lead guest editor of the special issue “Recent Advances on Data Fusion, Estimation in Navigation and Control” for Asian Journal of Control (AJC), 2018.

9.1.3.2. Reviewer - Reviewing Activities

Team members have been reviewers for several journals, including the most prestigious ones in their research area: IEEE Trans. on Automatic Control, IEEE Trans. on Control of Network Systems, IEEE Trans. on Signal Processing, Automatica, IEEE Signal Processing Letters, Systems and Control Letters, Int. Journal of Robust and Nonlinear Control, Elsevier Transportation Research Part B, IEEE Trans. on Intelligent Transportation Systems, IEEE/ASME Trans. on Mechatronics, IEEE Trans. on Instrumentations and Measurements, IEEE Sensors journal, IEEE Trans. on Robotics, AIMS Networks and Heterogeneous Network (NHM), Wiley Mathematical Methods in the Applied Sciences (MMAS), Journal of Mathematical Analysis and Applications (JMMA), Journal of Nonlinear Science and Applications (JNSA), Journal of the Franklin Institute, AMS Mathematical Reviews, Asian Journal of Control.

9.1.4. Invited Talks

- C. Canudas de Wit, “Model and Control of Large scale systems", plenary talk at IFAC-NecSys’18, Groningen, The Netherlands, August 2018.
- C. Canudas de Wit, “Control of Large scale Urban networks: a new perspective", plenary talk at IFAC-CTS’18, Savona, Italy, June 2018
- C. Canudas de Wit, “Scale-FreeBack", pitch talk at Transport Research Arena, Vienna, Austria, April 2018 .
- C. Canudas de Wit, “Towards Scale-Free Control of Large-scale Traffic Networks” and M. L. Delle Monache, “Micro-macro traffic modeling for estimation and control”, First SoPhy International Workshop on Societal-Scale Cyber-Physical Transport Systems Workshop, Stockholm, Sweden, September 2018.
- M. L. Delle Monache, “Can big data help traffic flow control?”, Workshop: Traffic flow control via PDE Techniques, IEEE-CDC Conference, Miami Beach, USA, December 2018.
- M. L. Delle Monache, “Control and estimation of traffic flow using autonomous vehicles”, Joint meeting of the Italian Mathematical Union, the Italian Society of Industrial and Applied Mathematics and the Polish Mathematical Society, Wroclaw, Poland, September 2018.
- M. L. Delle Monache, “Riemann solver for a macroscopic double-lane roundabout model”, 15th IFAC Symposium on Control in Transportation Systems, Savona, Italy, June 2018.
- M. L. Delle Monache, “Control of traffic: from ramp metering to autonomous vehicles”, Institute for Software Integrated Systems, Vanderbilt University, USA, May 2018.
- M. L. Delle Monache, “Control of traffic flow: from ramp metering to autonomous vehicles”, Seminar of the department of mathematics, University of Alabama, USA, April 2018
- M. L. Delle Monache, “Two-dimensional macroscopic model for large scale traffic network”, Incontro Scientifico su Modellizzazione ed Analisi di Problemi di Folle e Traffico, Politecnico di Torino, Italy, April 2018.
- M. L. Delle Monache, “Les mathématiques cachées du trafic routier”, ISN Conference, Académie de Grenoble, Inria Grenoble – Rhône-Alpes, France, March 2018.
- P. Frasca, “The harmonic influence in social networks and its distributed computation by message passing”, IXXI, ENS Lyon, July 3, 2018
- P. Frasca, “Randomization and quantization in opinion dynamics”, IRSTEA, Clermont-Ferrand, March 14, 2018.
- F. Garin, “Input-and-state observability of structured systems”, Paths in Mathematical Systems Theory workshop, Torino (Italy), Feb. 2018.

9.1.5. Leadership within the Scientific Community

Team members participate to the following technical committees of IEEE Control Systems Society and of the International Federation of Automatic Control:

- IEEE-CSS Technical Committee “Networks and Communications Systems” (P. Frasca and F. Garin);
- IFAC Technical Committee 1.5 on Networked Systems (P. Frasca and C. Canudas de Wit);
- IFAC Technical Committee 2.5 on Robust Control (P. Frasca);
- IFAC-TC7.1 Automotive Control (C. Canudas de Wit);
- IFAC-TC7.4 Transportation systems (C. Canudas de Wit).

C. Canudas de Wit is member of the advisory board (2017-21) of the project ”Societal-Scale Cyber-Physical Transport Systems” supported by the Swedish Strategic Research Foundation, KTH Sweden.

P. Frasca is member of the “Comité de Direction du GdR MACS ”, term 2019-2023.

P. Frasca reviewed project proposals for the ERC and the Italian Ministry of Scientific Research; A. Kibangou reviewed project proposals for ANR, NRF (South-African research agency), and ERC.

9.1.6. Research Administration

C. Canudas de Wit is a member of the COST-Inria-RA (Conseil d’Orientation Scientifique et Technologique, Inria Rhône-Alpes), since 2017.
F. Garin is member of two local committees at Inria Rhône-Alpes: Comité des Emplois Scientifiques (post-docs, délégations) since 2015 and Comité des Études Doctorales (PhD grants CORDI-S) since 2016.
A. Kibangou is:

- Elected member of the research department MSTIC (mathematics, information and communication sciences) of Univ. Grenoble Alpes
- Co-head of the PCS (Pervasive Computing Systems) action of Persyval-Lab
- Academic director (L2) IUT1 (GEII)
- Co-head for higher studies opportunities (Responsable poursuite d’études) (IUT1-GEII)

H. Fourati is

- Member of the Department of Electrical Engineering Council, IUT1 Grenoble, France (2018-2021)
- Member of CNU61 (Conseil national des universités, Génie informatique, Automatique et Traitement du Signal) since 2016.
- In charge of communication mission and visits to high school within the Department of Electrical Engineering, IUT1 Grenoble, France (2017-present).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master and PhD: M.L. Delle Monache, Traffic flow and crowd dynamics: modeling and computing, 18h, ED MSTII, Univ. Grenoble Alpes, France.

Master: F. Garin, Distributed Algorithms and Network Systems, 13.5h, M2, Univ. Grenoble Alpes, France.

Licence: H. Fourati, Mathématiques, 30h, L2, IUT1 (GEII1), Univ. Grenoble Alpes, France.

Licence: H. Fourati, Informatique Industrielle, 95h, L1, IUT 1 (GEII), Univ. Grenoble Alpes, France.

Licence: A. Kibangou, Automatique, 75h, L2, IUT1(GEII), Univ. Grenoble Alpes, France.

Licence: A. Kibangou, Mathématiques, 44h, L2, IUT1 (GEII), Univ. Grenoble Alpes, France.

Licence: A. Kibangou, Mathématiques, 126h, L1, IUT1 (GEII), Univ. Grenoble Alpes, France.

9.2.2. Supervision

PhD: Andrés Alberto Ladino Lopez, Traffic state estimation and prediction in freeways and urban networks, Univ. Grenoble Alpes, March 2018, co-advised by C. Canudas de Wit, A. Kibangou and H. Fourati.

PhD: Sebin Gracy, Input and state observability of linear network systems with application to security of cyber-physical systems, Univ. Grenoble Alpes, Nov. 2018, co-advised by A. Kibangou and F. Garin.

PhD: Stéphane Durand, Analysis of best response dynamics in potential games, Univ. Grenoble Alpes, Dec. 2018, co-advised by B. Gaujal and F. Garin.

PhD in progress: Stéphane Mollier, Aggregated Scale-Free Models for 2-D Large-scale Traffic Systems, from Oct. 2016, co-advised by C. Canudas de Wit, M. L. Delle Monache and B. Seibold.

PhD in progress: Nigina Toktassynova, Simulation and research of industrial flow control systems of the enterprise based on MES, from Oct. 2016, co-advised by H. Fourati and Batyrbek Suleimenov (Kazakh National Research Technical University).

PhD in progress: Nicolas Martin, On-line partitioning algorithms for evolutionary scale-free networks, from Dec. 2016, co-advised by C. Canudas de Wit and P. Frasca.

PhD in progress: Martin Rodriguez-Vega, Traffic density, traveling time and vehicle emission estimation in large-scale traffic networks, from Oct. 2017, co-advised by C. Canudas de Wit and H. Fourati.

PhD in progress: Muhammad Umar B. Niazi, State-state estimation design and optimal sensor placement algorithms for large-scale evolutionary dynamical networks, from Dec. 2017, co-advised by C. Canudas de Wit and A. Kibangou.

PhD in progress: Thembani Moyo, Origin and destination modeling and estimation for smart mobility, from May 2018, co-advised by A. Kibangou and W. Musakwa (Univ. of Johannesburg).

PhD in progress: Denis Nikitin, Scalable large-scale control of network aggregates, from Sept. 2018, co-advised by C. Canudas de Wit and P. Frasca.

PhD in progress: Liudmila Tumash, Traffic control in large-scale urban networks, from Sept. 2018, co-advised by C. Canudas de Wit and M. L. Delle Monache.

PhD in progress: Bassel Othman, Dynamic optimization of road traffic in a large-scale urban network from Oct. 2018, co-advised by C. Canudas de Wit and G. De Nunzio.

PhD in progress: Makia Zmitri, Estimating the attitude by IMU, magnetic and vision measures: an automatic control approach, from Oct. 2018, co-advised by H. Fourati and C. Prieur.

9.2.3. Juries

- P. Frasca was committee member of the PhD defence of Zhiyang Ju. Thesis: *Persistent Communication Connectivity of Multi-agent Systems*. University of Melbourne, Australia. PhD advisor: Dragan Nesic and Iman Shames. 2018
- P. Frasca was committee member of the PhD defence of Pierre-Yves Chevalier. Thesis: *Inhomogeneous Products of Stochastic Matrices with Application to Consensus Systems*. Université catholique de Louvain, Louvain-la-Neuve, Belgium. Ph.D. advisors: Julien Hendrickx and Raphael Jungers, June 2018
- P. Frasca was committee member of the PhD defence of Domenico Tangredi. Thesis: *Consensus in Heterogeneous Opinion Dynamics Networks*. University of Sannio, Benevento, Italy. Ph.D. advisor: Francesco Vasca, 2018.
- H. Fourati was committee member of the PhD defence of Fadoua Taia-Alaoui, IFSTTAR / Univ. Nantes, Dec. 2018.
- F. Garin was scientific assessor for the promotion of Maben Rabi to 'oavlönad docent', Chalmers University of Technology (Sweden), Jan. 2018.

- A. Kibangou was a reviewer of the PhD thesis of Thomas Brault "Étude des algorithmes de fusion de données pour estimer l'orientation d'un objet", Sorbonne Universités, defended November 8th, 2018.
- A. Kibangou was a reviewer of the thesis of Smart Dumba "Modelling Signalised Intersections' Capacity under the Impact of Minibus Public Transport in Harare, Zimbabwe", University of Zimbabwe. To be defended in January 2019.

9.3. Popularization

9.3.1. Articles and contents

P. Frasca has co-authored a column on the IEEE Control Systems Magazine about the activities of the Technical Committee on Networks and Communications Systems [67].

9.3.2. Education

M. L. Delle Monache gave a talk on "Les mathématiques cachées du trafic routier" to the high school teachers in the framework of the ISN Conference in collaboration with the Académie de Grenoble. The video of the conference is available at https://www.canal-u.tv/video/inria/les_mathematiques_cachees_du_trafic_routier.44275.

9.3.3. Interventions

Vadim Bertrand presented demos of the GTL-Ville:

- at *Club PTV Vision* (https://discover.ptvgroup.com/Club_PTV_Vision_2018), on Oct. 4th, in Paris
- at *Rencontres Inria-industrie* (https://www.inria.fr/innovation/recherche-partenariale-transfert/rencontres-inria-industrie/presentation), on Nov. 20th, in Paris

10. Bibliography

Major publications by the team in recent years

[1] G. BIANCHIN, P. FRASCA, A. GASPARRI, F. PASQUALETTI.*The Observability Radius of Networks*, in "IEEE Transactions on Automatic Control", June 2017, vol. 62, n^{o} 6, p. 3006-3013 [*DOI :* 10.1109/TAC.2016.2608941], https://hal.archives-ouvertes.fr/hal-01528187

[2] G. DE NUNZIO, C. CANUDAS DE WIT, P. MOULIN, D. DI DOMENICO.*Eco-Driving in Urban Traffic Networks Using Traffic Signals Information*, in "International Journal of Robust and Nonlinear Control", 2016, n^{o} 26, p. 1307–1324 [*DOI :* 10.1002/RNC.3469], https://hal.archives-ouvertes.fr/hal-01297629

[3] R. FABBIANO, F. GARIN, C. CANUDAS DE WIT.*Distributed Source Seeking without Global Position Information*, in "IEEE Transactions on Control of Network Systems", March 2018, vol. 5, n^{o} 1, p. 228-238 [*DOI :* 10.1109/TCNS.2016.2594493], https://hal.archives-ouvertes.fr/hal-01354294

[4] F. FAGNANI, P. FRASCA.*Introduction to Averaging Dynamics over Networks*, Springer, 2017 [*DOI :* 10.1007/978-3-319-68022-4], https://hal.archives-ouvertes.fr/hal-01614915

[5] H. FOURATI.*Multisensor Data Fusion: From Algorithms and Architectural Design to Applications (Book)*, Series: Devices, Circuits, and Systems, CRC Press, Taylor & Francis Group LLC, August 2015, 663, https://hal.inria.fr/hal-01169514

[6] S. GRACY, F. GARIN, A. Y. KIBANGOU.*Input and State Observability of Network Systems with Time-Varying Topology*, in "IEEE transactions on control of network systems", 2019 [*DOI :* 10.1109/TCNS.2018.2880304], https://hal.archives-ouvertes.fr/hal-01918497

[7] A. Y. KIBANGOU, C. COMMAULT.*Observability in Connected Strongly Regular Graphs and Distance Regular Graphs*, in "IEEE Transactions on Control of Network Systems", December 2014, vol. 1, n^o^ 4, p. 360-369 [*DOI :* 10.1109/TCNS.2014.2357532], https://hal.archives-ouvertes.fr/hal-01092954

[8] A. MAKNI, H. FOURATI, A. Y. KIBANGOU.*Energy-aware Adaptive Attitude Estimation Under External Acceleration for Pedestrian Navigation*, in "IEEE/ASME Transactions on Mechatronics", June 2016, vol. 21, n^o^ 3, p. 1366-1375 [*DOI :* 10.1109/TMECH.2015.2509783], https://hal.inria.fr/hal-01241403

[9] D. PISARSKI, C. CANUDAS DE WIT.*Nash Game Based Distributed Control Design for Balancing of Traffic Density over Freeway Networks*, in "IEEE Transactions on Control of Network Systems", 2016, vol. 3, n^o^ 2, p. 149-161 [*DOI :* 10.1109/TCNS.2015.2428332], https://hal.archives-ouvertes.fr/hal-01251805

[10] R. E. STERN, S. CUI, M. L. DELLE MONACHE, R. BHADANI, M. BUNTING, M. CHURCHILL, N. HAMILTON, R. HAULCY, H. POHLMANN, F. WU, B. PICCOLI, B. SEIBOLD, J. SPRINKLE, D. B. WORK.*Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments*, in "Transportation research. Part C, Emerging technologies", 2018, vol. 89, p. 205-221 [*DOI :* 10.1016/J.TRC.2018.02.005], https://hal.inria.fr/hal-01614638

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] S. DURAND.*Analysis of best response dynamics in potential games*, Université Grenoble Alpes, December 2018

[12] S. GRACY.*Input and state observability of linear network systems with application to security of cyber-physical systems*, Université Grenoble Alpes, November 2018

[13] A. LADINO LOPEZ.*Traffic state estimation and prediction in freeways and urban networks*, Université Grenoble Alpes, March 2018, https://tel.archives-ouvertes.fr/tel-01867240

Articles in International Peer-Reviewed Journal

[14] F. ACCIANI, P. FRASCA, G. HEIJENK, A. STOORVOGEL.*Achieving robust average consensus over lossy wireless networks*, in "IEEE Transactions on Control of Network Systems", 2018, p. 1 - 11 [*DOI :* 10.1109/TCNS.2018.2800407], https://hal.archives-ouvertes.fr/hal-01699122

[15] S. BALDI, P. FRASCA.*Adaptive synchronization of unknown heterogeneous agents: an adaptive virtual model reference approach*, in "Journal of The Franklin Institute", 2018, p. 1-21 [*DOI :* 10.1016/J.JFRANKLIN.2018.01.022], https://hal.archives-ouvertes.fr/hal-01702085

[16] S. BALDI, S. YUAN, P. FRASCA.*Output synchronization of unknown heterogeneous agents via distributed model reference adaptation*, in "IEEE Transactions on Control of Network Systems", 2018, p. 1 - 10 [*DOI :* 10.1109/TCNS.2018.2844260], https://hal.archives-ouvertes.fr/hal-01809667

[17] G. CASADEI, D. ASTOLFI.*Multi-pattern output consensus in networks of heterogeneous nonlinear agents with uncertain leader: a nonlinear regression approach*, in "IEEE Transactions on Automatic Control", August 2018, vol. 63, n^o 8, p. 2581-2587 [*DOI :* 10.1109/TAC.2017.2771316], https://hal.archives-ouvertes.fr/hal-01626655

[18] G. CASADEI, V. BERTRAND, B. GOUIN, C. CANUDAS DE WIT.*Aggregation and travel time calculation over large scale traffic networks: An empiric study on the Grenoble City*, in "Transportation research. Part C, Emerging technologies", 2018, vol. 95, p. 713-730 [*DOI :* 10.1016/J.TRC.2018.07.033], https://hal.archives-ouvertes.fr/hal-01953560

[19] F. CERAGIOLI, P. FRASCA.*Consensus and Disagreement: The Role of Quantized Behaviors in Opinion Dynamics*, in "SIAM Journal on Control and Optimization", March 2018, vol. 56, n^o 2, p. 1058 - 1080 [*DOI :* 10.1137/16M1083402], https://hal.archives-ouvertes.fr/hal-01741358

[20] M. L. DELLE MONACHE, P. GOATIN, B. PICCOLI.*Priority-based Riemann solver for traffic flow on networks* , in "Communications in Mathematical Sciences", 2018, vol. 16, n^o 1, p. 185-211, https://hal.inria.fr/hal-01336823

[21] S. DURAND, F. GARIN, B. GAUJAL.*Distributed best response dynamics with high playing rates in potential games*, in "Performance Evaluation", October 2018, vol. 129, p. 40-59 [*DOI :* 10.1016/J.PEVA.2018.09.007], https://hal.inria.fr/hal-01940128

[22] R. FABBIANO, F. GARIN, C. CANUDAS DE WIT.*Distributed Source Seeking without Global Position Information*, in "IEEE Transactions on Control of Network Systems", March 2018, vol. 5, n^o 1, p. 228-238 [*DOI :* 10.1109/TCNS.2016.2594493], https://hal.archives-ouvertes.fr/hal-01354294

[23] P. FRASCA, S. TARBOURIECH, L. ZACCARIAN.*Hybrid models of opinion dynamics with opinion-dependent connectivity*, in "Automatica", February 2019, vol. 100, p. 153-161 [*DOI :* 10.1016/J.AUTOMATICA.2018.11.017], https://hal.archives-ouvertes.fr/hal-01940187

[24] S. GRACY, F. GARIN, A. KIBANGOU.*Structural and Strongly Structural Input and State Observability of Linear Network Systems*, in "IEEE Transactions on Control of Network Systems", December 2018, vol. 5, n^o 4, p. 2062-2072 [*DOI :* 10.1109/TCNS.2017.2782489], https://hal.archives-ouvertes.fr/hal-01663175

[25] S. GRACY, F. GARIN, A. Y. KIBANGOU.*Input and State Observability of Network Systems with Time-Varying Topology*, in "IEEE Transactions on Control of Network Systems", 2019 [*DOI :* 10.1109/TCNS.2018.2880304], https://hal.archives-ouvertes.fr/hal-01918497

[26] P. GRANDINETTI, C. CANUDAS DE WIT, F. GARIN.*Distributed Optimal Traffic Lights Design for Large-Scale Urban Networks*, in "IEEE Transactions on Control Systems Technology", March 2018, vol. PP, n^o 99, p. 1-14 [*DOI :* 10.1109/TCST.2018.2807792], https://hal.inria.fr/hal-01727937

[27] D. JABRI, D. E. C. BELKHIAT, H. FOURATI.*Design of a decentralized tracking control for a class of switched large-scale systems*, in "Journal of Control Engineering and Applied Informatics", September 2018, vol. 20, n^o 3, p. 24 - 35, https://hal.inria.fr/hal-01929596

[28] E. KERSTEL, A. GARDELEIN, M. BARTHELEMY, H. FOURATI, Y. GILOT, E. LECOARER, J. RODRIGO, T. SEQUIES, V. BORNE, G. BOURDAROT, J.-Y. BURLET, A. CHRISTIDIS, J. SEGURA, B. BOULANGER, V. BOUTOU, M. BOUZAT, M. CHABANOL, L. FESQUET, M. MOULIN, J.-M. NIOT, R. P. BASTOS, B. ROBU,

E. ROLLAND, S. TORU, M. FINK, R. URSIN, S. K. JOSHI.*Nanobob: a CubeSat mission concept for quantum communication experiments in an uplink configuration*, in "European physical journal quantum technology", December 2018, vol. 5, n^o 6, p. 1-30 [*DOI :* 10.1140/EPJQT/S40507-018-0070-7], https://hal.inria.fr/hal-01929079

[29] A. LABIBIAN, S. H. POURTAKDOUST, A. ALIKHANI, H. FOURATI.*Development of a radiation based heat model for satellite attitude determination*, in "Aerospace Science and Technology", November 2018, vol. 82-83, p. 479 - 486 [*DOI :* 10.1016/J.AST.2018.09.031], https://hal.inria.fr/hal-01929229

[30] A. MAKNI, A. KIBANGOU, H. FOURATI.*Data Fusion-Based Descriptor Approach for Attitude Estimation underaccelerated maneuvers*, in "Asian Journal of Control", January 2019, https://hal.archives-ouvertes.fr/hal-01982463

[31] N. MARTIN, P. FRASCA, C. CANUDAS DE WIT.*Large-scale network reduction towards scale-free structure*, in "IEEE Transactions on Network Science and Engineering", 2018, p. 1-12 [*DOI :* 10.1109/TNSE.2018.2871348], https://hal.archives-ouvertes.fr/hal-01885140

[32] T. MICHEL, P. GENEVÈS, H. FOURATI, N. LAYAÏDA.*Attitude Estimation for Indoor Navigation and Augmented Reality with Smartphones*, in "Pervasive and Mobile Computing", March 2018 [*DOI :* 10.1016/J.PMCJ.2018.03.004], https://hal.inria.fr/hal-01650142

[33] S. MOLLIER, M. L. DELLE MONACHE, C. CANUDAS DE WIT.*A simple example of two dimensional model for traffic: discussion about assumptions and numerical methods*, in "Transportation Research Record (TRR), Journal of the Transportation Research Board", August 2018, p. 1-20, https://hal.archives-ouvertes.fr/hal-01665285

[34] C. RAVAZZI, N. CHAN, P. FRASCA.*Distributed estimation from relative measurements of heterogeneous and uncertain quality*, in "IEEE transactions on Signal and Information Processing over Networks", 2018, p. 1 - 13 [*DOI :* 10.1109/TSIPN.2018.2869117], https://hal.archives-ouvertes.fr/hal-01877623

[35] M. RODRIGUEZ-VEGA, C. CANUDAS DE WIT, H. FOURATI.*Location of turning ratio and flow sensors for flow reconstruction in large traffic networks*, in "Transportation Research", January 2019, https://hal.archives-ouvertes.fr/hal-01958601

[36] W. S. ROSSI, P. FRASCA.*On the convergence of message passing computation of harmonic influence in social networks*, in "IEEE Transactions on Network Science and Engineering", 2018, vol. PP, n^o 99, p. 1-14 [*DOI :* 10.1109/TNSE.2018.2792401], https://hal.archives-ouvertes.fr/hal-01683634

[37] S. SAMARANAYAKE, J. REILLY, W. KRICHENE, M. L. DELLE MONACHE, P. GOATIN, A. BAYEN.*Discrete-time system optimal dynamic traffic assignment (SO-DTA) with partial control for horizontal queuing networks*, in "Transportation Science", 2018, vol. 52, n^o 4, p. 982-1001 [*DOI :* 10.1287/TRSC.2017.0800], https://hal.inria.fr/hal-01095707

[38] R. E. STERN, S. CUI, M. L. DELLE MONACHE, R. BHADANI, M. BUNTING, M. CHURCHILL, N. HAMILTON, R. HAULCY, H. POHLMANN, F. WU, B. PICCOLI, B. SEIBOLD, J. SPRINKLE, D. B. WORK.*Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments*, in "Transportation research. Part C, Emerging technologies", 2018, vol. 89, p. 205-221 [*DOI :* 10.1016/J.TRC.2018.02.005], https://hal.inria.fr/hal-01614638

[39] J. WU, Z. ZHOU, H. FOURATI, Y. CHENG.*A Super Fast Attitude Determination Algorithm for Consumer-Level Accelerometer and Magnetometer*, in "IEEE Transactions on Consumer Electronics", August 2018, vol. 64, n^o 3, p. 375 - 381 [*DOI :* 10.1109/TCE.2018.2859625], https://hal.inria.fr/hal-01922922

[40] J. WU, Z. ZHOU, H. FOURATI, R. LI, M. LIU.*Generalized Linear Quaternion Complementary Filter for Attitude Estimation from Multi-Sensor Observations: An Optimization Approach*, in "IEEE Transactions on Automation Science and Engineering", 2019, p. 1-14, https://hal.archives-ouvertes.fr/hal-01958652

[41] J. WU, Z. ZHOU, H. FOURATI, M. LIU.*Recursive Linear Continuous Quaternion Attitude Estimator From Vector Observations*, in "IET Radar, Sonar & Navigation", November 2018, https://hal.inria.fr/hal-01923022

[42] J. WU, Z. ZHOU, B. GAO, R. LI, Y. CHENG, H. FOURATI.*Fast Linear Quaternion Attitude Estimator Using Vector Observations*, in "IEEE Transactions on Automation Science and Engineering", January 2018, forthcoming [*DOI :* 10.1109/TASE.2017.2699221], https://hal.inria.fr/hal-01513263

[43] Z. ZHOU, S. MO, J. WU, H. FOURATI.*Behaviors classification based distance measuring system for pedestrians via a foot-mounted inertial sensor*, in "Asian Journal of Control", January 2019, p. 1-12 [*DOI :* 10.1002/ASJC.0000], https://hal.inria.fr/hal-01929624

[44] Z. ZHOU, J. WU, J. WANG, H. FOURATI.*Optimal, Recursive and Sub-optimal Linear Solutions to Attitude Determination from Vector Observations for GNSS/Accelerometer/Magnetometer Orientation Measurement*, in "Remote Sensing", March 2018, vol. 10, n^o 3, p. 1-28 [*DOI :* 10.3390/RS10030377], https://hal.inria.fr/hal-01722260

International Conferences with Proceedings

[45] F. ACCIANI, P. FRASCA, A. STOORVOGEL, E. SEMSAR-KAZEROONI, G. HEIJENK.*Cooperative adaptive cruise control over unreliable networks: an observer-based approach to increase robustness to packet loss*, in "ECC 2018 - 16th European Control Conference", Limassol, Cyprus, June 2018, p. 1399-1404, https://hal.archives-ouvertes.fr/hal-01721773

[46] D. ASTOLFI, G. CASADEI, R. POSTOYAN.*Emulation-based semiglobal output regulation of minimum phase nonlinear systems with sampled measurements*, in "16th European Control Conference, ECC'18", Limassol, Cyprus, June 2018, https://hal.archives-ouvertes.fr/hal-01840726

[47] S. BALDI, M. RIDHO ROSA, P. FRASCA.*Adaptive state-feedback synchronization with distributed input: the cyclic case*, in "7th IFAC Workshop on Distributed Estimation and Control in Networked Systems", Groningen, Netherlands, 7th IFAC Workshop on Distributed Estimation and Control in Networked Systems, August 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01945733

[48] S. BALDI, M. RIDHO ROSA, P. FRASCA, E. KOSMATOPOULOS.*Platooning merging maneuvers in the presence of parametric uncertainty*, in "NecSys 2018 - 7th IFAC Workshop on Distributed Estimation and Control in Networked Systems", Groningen, Netherlands, 7th IFAC Workshop on Distributed Estimation and Control in Networked Systems, August 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01945760

[49] G. CASADEI, C. CANUDAS DE WIT, S. ZAMPIERI.*Scale Free Controllability of Large-Scale Networks: an Output Controllability Approach*, in "CDC 2018 - 57th IEEE Conference on Decision and Control", Miami, FL, United States, December 2018, p. 1-8, https://hal.archives-ouvertes.fr/hal-01953539

[50] M. L. DELLE MONACHE, S. HAMMOND, B. PICCOLI.*Riemann solver for a macroscopic double-lane roundabout model*, in "15th IFAC Symposium on Control in Transportation Systems (CTS 2018)", Savona, Italy, June 2018, p. 55-60 [*DOI :* 10.1016/J.IFACOL.2018.07.010], https://hal.inria.fr/hal-01816703

[51] S. DURAND, F. GARIN, B. GAUJAL.*Distributed Best Response Algorithms for Potential Games*, in "ECC 2018 - 16th European Control Conference", Limassol, Cyprus, June 2018, p. 1-6, https://hal.inria.fr/hal-01726836

[52] S. DURAND, F. GARIN, B. GAUJAL.*Efficiency of Best Response Dynamics with High Playing Rates in Potential Games*, in "IFIP WG 7.3 Performance 2018 - 36th International Symposium on Computer Performance, Modeling, Measurements and Evaluation", Toulouse, France, December 2018, p. 1-2, https://hal.inria.fr/hal-01940150

[53] M. FRANCESCHELLI, P. FRASCA.*Proportional Dynamic Consensus in Open Multi-Agent Systems*, in "CDC 2018 - 57th IEEE Conference on Decision and Control", Miami, United States, December 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01945840

[54] F. GARIN, S. GRACY, A. Y. KIBANGOU.*Unbiased Filtering for State and Unknown Input with Delay*, in "NecSys 2018 - 7th IFAC Workshop on Distributed Estimation and Control in Networked Systems", Groningen, Netherlands, IFAC, August 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01850957

[55] S. HAMMOND, A. RAT, M. L. DELLE MONACHE, B. PICCOLI.*Operational Performance of a Double-Lane Roundabout with Additional Lane Length Design: Lighthill-Whitham-Richards Model Analysis*, in "TRB 2018 - Transportation Research Board 97th Annual Meeting", Washington, D.C., United States, January 2018, p. 1-14, https://hal.inria.fr/hal-01819238

[56] A. LADINO LOPEZ, C. CANUDAS DE WIT, A. KIBANGOU, H. FOURATI, M. RODRIGUEZ-VEGA.*Density and flow reconstruction in urban traffic networks using heterogeneous data sources*, in "ECC 2018 - 16th European Control Conference", Limassol, Cyprus, June 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01731356

[57] N. MARTIN, P. FRASCA, C. CANUDAS DE WIT.*A network reduction method inducing scale-free degree distribution*, in "ECC 2018 - 16th European Control Conference", Limassol, Cyprus, June 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01632482

[58] N. MARTIN, P. FRASCA, C. CANUDAS DE WIT.*MergeToCure: a New Strategy to Allocate Cure in an Epidemic over a Grid-like network Using a Scale-Free Abstraction*, in "NecSys", Groningen, Netherlands, IFAC-PapersOnLine, August 2018, vol. 51, n^{o} 23, p. 34-39, https://hal.archives-ouvertes.fr/hal-01820672

[59] S. MOLLIER, M. L. DELLE MONACHE, C. CANUDAS DE WIT.*2D-LWR in large-scale network with space dependent fundamental diagram*, in "ITSC 2018 - 21st IEEE International Conference on Intelligent Transportation Systems", Maui, United States, IEEE, November 2018, p. 1-8, https://hal.archives-ouvertes.fr/hal-01866959

[60] S. MOLLIER, M. L. DELLE MONACHE, C. CANUDAS DE WIT.*A step towards a multidirectional 2D model for large scale traffic networks*, in "Transportation Research Board 2019", Washington, United States, January 2019, https://hal.archives-ouvertes.fr/hal-01948466

[61] O. L. QUINTERO MONTOYA, C. CANUDAS DE WIT. *Estimation of fundamental diagrams in large-scale traffic networks with scarce sensor measurements*, in "ITSC 2018 - 21st IEEE International Conference on Intelligent Transportation Systems", Maui, United States, November 2018, p. 1-21, https://hal.archives-ouvertes.fr/hal-01874562

[62] M. RODRIGUEZ-VEGA, C. CANUDAS DE WIT, H. FOURATI. *Joint Flow and Density Reconstruction in Large Traffic Networks UsingPartial Turning Ratio Information*, in "CDC 2018 - 57th IEEE Conference on Decision and Control", Miami, United States, IEEE, December 2018, p. 1-6, https://hal.inria.fr/hal-01928699

[63] W. S. ROSSI, P. FRASCA. *Asynchronous opinion dynamics on the k-nearest-neighbors graph*, in "CDC 2018 - 57th IEEE Conference on Decision and Control", Miami Beach, United States, 57th IEEE Conference on Decision and Control, IEEE, December 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01877618

[64] W. S. ROSSI, P. FRASCA. *Effects of Network Communities and Topology Changes in Message-Passing Computation of Harmonic Influence in Social Networks*, in "CPHS 2018 - 2nd IFAC Conference on Cyber-Physical & Human Systems", Miami, United States, 2nd IFAC Conference on Cyber-Physical & Human Systems, IEEE, December 2018, p. 1-4, https://hal.archives-ouvertes.fr/hal-01877617

[65] Z. ZHANG, Z. ZHOU, J. WU, S. DU, H. FOURATI. *Fast Linear Attitude Estimation and Angular Rate Generation*, in "IPIN 2018 - International Conference on Indoor Positioning and Indoor Navigation", Nantes, France, IEEE, September 2018, p. 1-7 [*DOI :* 10.1109/IPIN.2018.8533736], https://hal.inria.fr/hal-01928684

Scientific Books (or Scientific Book chapters)

[66] F. CERAGIOLI, P. FRASCA. *Discontinuities, generalized solutions and (dis)agreement in opinion dynamics*, in "Control Subject to Computational and Communication Constraints", S. TARBOURIECH, A. GIRARD, L. HETEL (editors), Lecture Notes in Control and Information Sciences, Springer, June 2018, vol. 475, p. 287-309 [*DOI :* 10.1007/978-3-319-78449-6_14], https://hal.archives-ouvertes.fr/hal-01739420

Scientific Popularization

[67] D. E. QUEVEDO, P. FRASCA, S. YUKSEL, G. COMO. *Technical Committee on Networks and Communication Systems [Technical Activities]*, in "IEEE Control Systems", August 2018, vol. 38, n° 4, p. 18-19 [*DOI :* 10.1109/MCS.2018.2830024], https://hal.archives-ouvertes.fr/hal-01946084

Other Publications

[68] M. L. DELLE MONACHE, K. CHI, Y. CHEN, P. GOATIN, K. HAN, J.-M. QIU, B. PICCOLI. *A three-phase fundamental diagram from three-dimensional traffic data*, August 2018, working paper or preprint, https://hal.inria.fr/hal-01864628

[69] M. L. DELLE MONACHE, T. LIARD, B. PICCOLI, R. E. STERN, D. WORK. *Traffic reconstruction using autonomous vehicles*, September 2018, working paper or preprint, https://hal.inria.fr/hal-01882309

[70] M. L. DELLE MONACHE, T. LIARD, A. RAT, R. STERN, R. BADHANI, B. SEIBOLD, J. SPRINKLE, D. WORK, B. PICCOLI. *Feedback control algorithms for the dissipation of traffic waves with autonomous vehicles*, 2018, working paper or preprint, https://hal.inria.fr/hal-01930724

[71] T. LIARD, B. PICCOLI.*Well-posedness for scalar conservation laws with moving flux constraints*, January 2018, https://arxiv.org/abs/1801.04814 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01684506

[72] A. MALKOVA, J. ARBEL, M. L. DELLE MONACHE.*DATASAFE: understanding Data Accidents for TrAffic SAFEty Acknowledgments*, September 2018, 1, Bayesian learning theory for complex data modelling Workshop, Poster, https://hal.archives-ouvertes.fr/hal-01950663

[73] S. MOLLIER, M. L. DELLE MONACHE, C. CANUDAS DE WIT, B. SEIBOLD.*Two-dimensional macroscopic model for large scale traffic networks*, June 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01819013

Project-Team NUMED

Numerical Medicine

IN COLLABORATION WITH: Unité de Mathématiques Pures et Appliquées

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Lyon

Université Claude Bernard (Lyon 1)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Modeling and Control for Life Sciences

Table of contents

Project-Team NUMED

Creation of the Project-Team: 2009 January 01

Keywords:

Computer Science and Digital Science:
A6. - Modeling, simulation and control
A6.1. - Methods in mathematical modeling
A6.2. - Scientific computing, Numerical Analysis & Optimization
A6.3. - Computation-data interaction

Other Research Topics and Application Domains:
B1. - Life sciences
B1.1. - Biology
B2. - Health
B2.2. - Physiology and diseases
B2.2.2. - Nervous system and endocrinology
B2.2.3. - Cancer
B2.2.4. - Infectious diseases, Virology
B2.4.1. - Pharmaco kinetics and dynamics
B2.4.2. - Drug resistance
B2.6.1. - Brain imaging

1. Team, Visitors, External Collaborators

Research Scientist
Helene Leman [Inria, Researcher, from Oct 2018]

Faculty Members
Emmanuel Grenier [Team leader, Ecole Normale Supérieure Lyon, Professor, HDR]
Arthur Marly [Ecole Normale Supérieure Lyon, Associate Professor, from Sep 2018]
Paul Vigneaux [Ecole Normale Supérieure Lyon, Associate Professor, HDR]
Arthur Marly [Ecole Normale Supérieure Lyon, until Aug 2018]

Technical Staff
David Coulette [CNRS, from Mar 2018]

Administrative Assistant
Sylvie Boyer [Inria]

2. Overall Objectives

2.1. Overall Objectives

The purpose of Numed is to develop new numerical methods and tools to simulate and parametrize complex systems arising in biology and medicine. Numed focuses on two axes:

- Thema 1: Modeling using complex models: how to deal with multiple spatial or temporal scales (theoretical study, numerical simulations)?

This covers several aims: design of models of propagation taking into account the microscopic phenomena and starting from small scale description, importance of mechanics in the growth of tissues, peculiarities of tumor tissues, nonlinear rheology, evolutionary perspectives.

- Thema 2: Parametrization of complex models: how to find parameters for complex models, with particular emphasis on population approaches and on computationally expensive models.

and on main axe of applications, namely cancer.

The aim is to develop models of cancer growth in close link with clinical data.

3. Research Program

3.1. Design of complex models

3.1.1. Project team positioning

The originality of our work is the quantitative description of phenomena accounting for several time and spatial scales. Here, propagation has to be understood in a broad sense. This includes propagation of invasive species, chemotactic waves of bacteira, evoluation of age structures populations ... Our main objectives are the quantitative calculation of macroscopic quantities as the rate of propagation, and microscopic distributions at the edge and the back of the front. These are essential features of propagation which are intimately linked in the long time dynamics.

3.1.2. Recent results

- Population models.

  H. Leman works at the interface between mathematics and biology, thanks to probabilist and determinist studies of models of populations. More precisely, she studies and develops probabilistic models, called agent models that described the population at an individual level. Each individual is characterized by one or more phenotypic traits and by its position, which may influence at the same time its ecological behavior and its motion. From a biological point of view these models are particularly interesting since they allow to include a large variety of interactions between individuals. These processes may also be studied in details to obtain theoretical results which may be simulated thanks to exact algorithms. To get quantitative results H. Leman uses changes of scales in space and time (large population, rare mutations, long time), following various biological assumptions.

  In a first study , H. Leman tries to understand the interactions between sexual preference mechanisms and evolutive forces inside spatially structured populations. Recently she got interesting in the description of necessary conditions to facilitate the emergence of such preferences by individuals.

  As a second example, H. Leman is also interested in the modeling and study of cooperative bacterias and tries to understand the impact of spatial structures in the eco - evolutions of these bacterias. Space seems to be an essential factor to facilitate the emergence of cooperation between bacteries.

- Inviscid limit of Navier Stokes equations.

  The question of the behavior of solutions of Navier Stokes equations in a bounded domain as the viscosity goes to 0 is a classical and highly difficult open question in Fluid Mechanics. A small boundary layer, called Prandtl layer, appears near the boundary, which turns out to be unstable if the viscosity is small enough. The stability analysis of this boundary layer is highly technical and remained open since the first formal analysis in the 1940's by physicists like Orr, Sommerfeld, Tollmien, Schlichting or Lin. E. Grenier recently made a complete mathematical analysis of this spectral problem, in collaboration with T. Nguyen and Y. Guo. We rigorously proved that any shear layer is spectrally and linearly unstable if the viscosity is small enough, which is the first mathematical result in that field. We also get some preliminary nonlinear results. A book on this subject is in preparation, already accepted by Springer.

- Numerical analysis of complex fluids: the example of avalanches.

  This deals with the development of numerical schemes for viscoplastic materials (namely with Bingham or Herschell-Bulkley laws). Recently, with other colleagues, Paul Vigneaux finished the design of the first 2D well-balanced finite volume scheme for a shallow viscoplastic model. It is illustrated on the famous Taconnaz avalanche path in the Mont-Blanc, Chamonix, in the case of dense snow avalanches. The scheme deals with general Digital Elevation Model (DEM) topographies, wet/dry fronts and is designed to compute precisely the stopping state of avalanches, a crucial point of viscoplastic flows which are able to rigidify [21].

3.1.3. Collaborations

- Ecology: Orsay (C. Coron), Toulouse (IMT, M. Costa), MNHM Paris (V. Llaurens), LISC Paris (C. Smadi), ENS Paris (R. Ferrière, E. Abs), CIMAT (Mexique, J. C. P. Millan).
- Inviscid limit of Navier Stokes equations: Brown University (Y. Guo, B. Pausader), Penn State University (T. Nguyen), Orsay University (F. Rousset).
- Numerical analysis of complex fluids: Enrique D. Fernandez - Nieto (Univ. de Sevilla, Spain), Jose Maria Gallardo (Univ. de Malaga, Spain).

3.2. Parametrization of complex systems

3.2.1. Project-team positioning

Clinical data are often sparse: we have few data per patient. The number of data is of the order of the number of parameters. In this context, a natural way to parametrize complex models with real world clinical data is to use a Bayesian approach, namely to try to find the distribution of the model parameters in the population, rather than to try to identify the parameters of every single patient. This approach has been pioneered in the 90's by the Nonmem software, and has been much improved thanks to Marc Lavielle in the 2000's. Refined statistical methods, called SAEM, have been tuned and implemented in commercial softwares like Monolix.

3.2.2. Recent results

The main problem when we try to parametrize clinical data using complex systems is the computational time. One single evaluation of the model can be costly, in particular if this model involves partial differential equations, and SAEM algorithm requires hundreds of thousands of single evaluations. The time cost is then too large, in particular because SAEM may not be parallelized.

To speed up the evaluation of the complex model, we replace it by an approximate one, or so called metamodel, constructed by interpolation of a small number of its values. We therefore combine the classical SAEM algorithm with an interpolation step, leading to a strong acceleration. Interpolation can be done through a precomputation step on a fixed grid, or through a more efficient kriging step. The interpolation grid or the kriging step may be improved during SAEM algorithm in an iterative way in order to get accurate evaluations of the complex system only in the domain of interest, namely near the clinical values [14],[15].

We applied these new algorithms to synthetic data and are currently using them on glioma data. We are also currently trying to prove the convergence of the corresponding algorithms. We will develop glioma applications in the next section.

Moreover E. Ollier in his phD developed new strategies to distinguish various populations within a SAEM algorithm [23].

We have two long standing collaborations with Sanofi and Servier on parametrization issues:

- Servier: during a four years contract, we modelled the pkpd of new drugs and also study the combination and optimization of chimiotherapies.
- Sanofi: during a eight years contract, Emmanuel Grenier wrote a complete software devoted to the study of the degradation of vaccine. This software is used worldwide by Sanofi R&D teams in order to investigate the degradation of existing or new vaccines and to study their behavior when they are heated. This software has been used on flu, dengue and various other diseases.

3.2.3. Collaborations

- Academic collaborations: A. Leclerc Samson (Grenoble University)
- Medical collaborations: Dr Ducray (Centre Léon Bérard, Lyon) and Dr Sujobert (Lyon Sud Hospital)
- Industrial contracts: we used parametrization and treatment improvement techniques for Servier (four years contract, on cancer drug modeling and optimization) and Sanofi (long standing collaboration)

3.3. Multiscale models in oncology

3.3.1. Project-team positioning

Cancer modeling is the major topic of several teams in France and Europe, including Mamba, Monc and Asclepios to quote only a few Inria teams. These teams try to model metastasis, tumoral growth, vascularisation through angiogenesis, or to improve medical images quality. Their approaches are based on dynamical systems, partial differential equations, or on special imagery techniques.

Numed focuses on the link between very simple partial differential equations models, like reaction diffusion models, and clinical data.

3.3.2. Results

During 2018 we developed new collaborations with the Centre Léon Bérard (Lyon), in particular on the following topics

- Barcoding of cells: thanks to recent techniques, it is possible to mark each cell with an individual barcode, and to follow its division and descendance. The analysis of such data requires probabilistic models, in particular to model experimental bias.
- Apoptosis: the question is to investigate whether the fate of neighboring cells influence the evolution of a given cell towards apoptosis, starting from videos of in vitro drug induced apoptosis.
- Dormance: Study of the dynamics of cells under immunotherapy, starting from experimental in vitro data.
- Colorectal cancer: In vitro study of the role of stem cells in drug resistance, in colorectal cancer.

3.3.3. Collaborations

- Centre Léon Bérard (in particular: Pr Puisieux, G. Ichim, M. Plateroni, S. Ortiz).

4. New Software and Platforms

4.1. Bingham flows

FUNCTIONAL DESCRIPTION: A 1D and 2D code with a new method for the computation of viscoplatic flows with free-surface. It essentially couples Optimization methods and Well-Balanced Finite-Volumes schemes for viscous shallow-water equations (induced by the viscoplastic nature of the fluid). Currently applied to avalanches of dense snow, it is a private code currently actively developed (in C++). One of the key feature is that its well-balanced property allows to obtained the stationary states which are linked to the stopping of the snow avalanche for this highly non-linear type of fluid.

- Contact: Emmanuel Grenier

4.2. OptimChemo

FUNCTIONAL DESCRIPTION: OptimChemo is a userfriendly software designed to study numerically the effect of multiple chemotherapies on simple models of tumour growth and to optimize chemotherapy schedules.

- Participants: Ehouarn Maguet, Emmanuel Grenier, Paul Vigneaux and Violaine Louvet
- Contact: Emmanuel Grenier

4.3. SETIS

KEYWORDS: Health - DICOM - Medical imaging - Drug development

FUNCTIONAL DESCRIPTION: SETIS software is a GUI allowing to treat DICOM medical images to extract pathological data. These data can then be exported and used in a SAEM software (including Monolix (Inria & Lixoft)) for the parameters' estimation of models in the context of population approaches. As an example SETIS can be used to segment and compute the tumor size of a patients from MRI scans taken at different times. The software is sufficiently general to be used in various situations by clinicians (already done by colleagues in Lyon Hospital).

- Participants: Ehouarn Maguet and Paul Vigneaux
- Partner: ENS Lyon
- Contact: Paul Vigneaux

4.4. SIMPHYT

KEYWORDS: Bioinformatics - Cancer - Drug development

FUNCTIONAL DESCRIPTION: SimPHyt is an implementation in Python of the low grad glioma model. The aim is to predict the evolution of the glioma size of patients.

- Participant: Benjamin Ribba
- Contact: Benjamin Ribba

4.5. SITLOG

- Participants: Benjamin Ribba and Morgan Martinet
- Contact: Emmanuel Grenier

4.6. VAXSIMSTAB

KEYWORDS: Bioinformatics - Health - Drug development

FUNCTIONAL DESCRIPTION: VAXSIMSTAB is a modeler stability prediction of vaccine software.

- Participants: Benjamin Ribba, Emmanuel Grenier and Vincent Calvez
- Contact: Benjamin Ribba

5. Partnerships and Cooperations

5.1. National Initiatives

5.1.1. ANR

CNRS InFIniti, 2017-2018 (P. Vigneaux): 12ke in 2018

6. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] E. GRENIER, F. HAMEL.*Large time monotonicity of solutions of reaction-diffusion equations in* R^N, in "Journal de Mathématiques Pures et Appliquées", 2018, vol. 112, p. 89-117, https://arxiv.org/abs/1606.00176 [*DOI :* 10.1016/J.MATPUR.2017.09.015], https://hal.archives-ouvertes.fr/hal-01324533

[2] E. GRENIER, C. HELBERT, V. LOUVET, A. SAMSON, P. VIGNEAUX.*Population parametrization of costly black box models using iterations between SAEM algorithm and kriging*, in "Computational and Applied Mathematics", March 2018, vol. 37, n^{o} 1, p. 161-173, Accepted March, 24, 2016. Published online April, 16, 2016. Journal issue available Feb 23, 2018. [*DOI :* 10.1007/s40314-016-0337-5], https://hal.archives-ouvertes.fr/hal-01224004

[3] E. GRENIER, T. T. NGUYEN.*Sublayer of Prandtl boundary layers*, in "Arch. Ration. Mech. Anal.", 2018, vol. 229, n^{o} 3, p. 1139–1151, https://doi.org/10.1007/s00205-018-1235-3

Project-Team PERCEPTION

Interpretation and Modelling of Images and Videos

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Vision, perception and multimedia interpretation

Table of contents

Project-Team PERCEPTION

Creation of the Team: 2006 September 01, updated into Project-Team: 2008 January 01

Keywords:

Computer Science and Digital Science:

A3.4. - Machine learning and statistics
A5.1. - Human-Computer Interaction
A5.3. - Image processing and analysis
A5.4. - Computer vision
A5.7. - Audio modeling and processing
A5.10.2. - Perception
A5.10.5. - Robot interaction (with the environment, humans, other robots)
A9.2. - Machine learning
A9.5. - Robotics

Other Research Topics and Application Domains:

B5.6. - Robotic systems

1. Team, Visitors, External Collaborators

Research Scientists

Radu Patrice Horaud [Team leader, Inria, Senior Researcher, HDR]
Xavier Alameda-Pineda [Inria, Researcher]
Xiaofei Li [Inria, Starting Research Position]
Pablo Mesejo Santiago [Inria, Starting Research Position, until March 2018]

Faculty Member

Laurent Girin [Institut polytechnique de Grenoble, Professor, HDR]

Technical Staff

Soraya Arias [Inria]
Bastien Mourgue [Inria]
Guillaume Sarrazin [Inria]

PhD Students

Yutong Ban [Inria]
Guillaume Delorme [Inria]
Sylvain Guy [Univ Grenoble Alpes]
Stephane Lathuiliere [Inria, until May 2018]
Benoit Masse [Univ Grenoble Alpes, until November 2018]
Yihong Xu [Inria, since September 2018]

Post-Doctoral Fellows

Simon Leglaive [Inria, since February 2018]
Mostafa Sadeghi [Inria, since August 2018]

Visiting Scientists

Christine Evers [Imperial College London, from Jul 2018 until Aug 2018]
Sharon Gannot [Bar Ilan University, from Jan 2018 until Feb 2018]
Tomislav Pribanic [University of Zagreb, from Apr 2018 until Aug 2018]

Administrative Assistant

Nathalie Gillot [Inria]

2. Overall Objectives

2.1. Audio-Visual Machine Perception

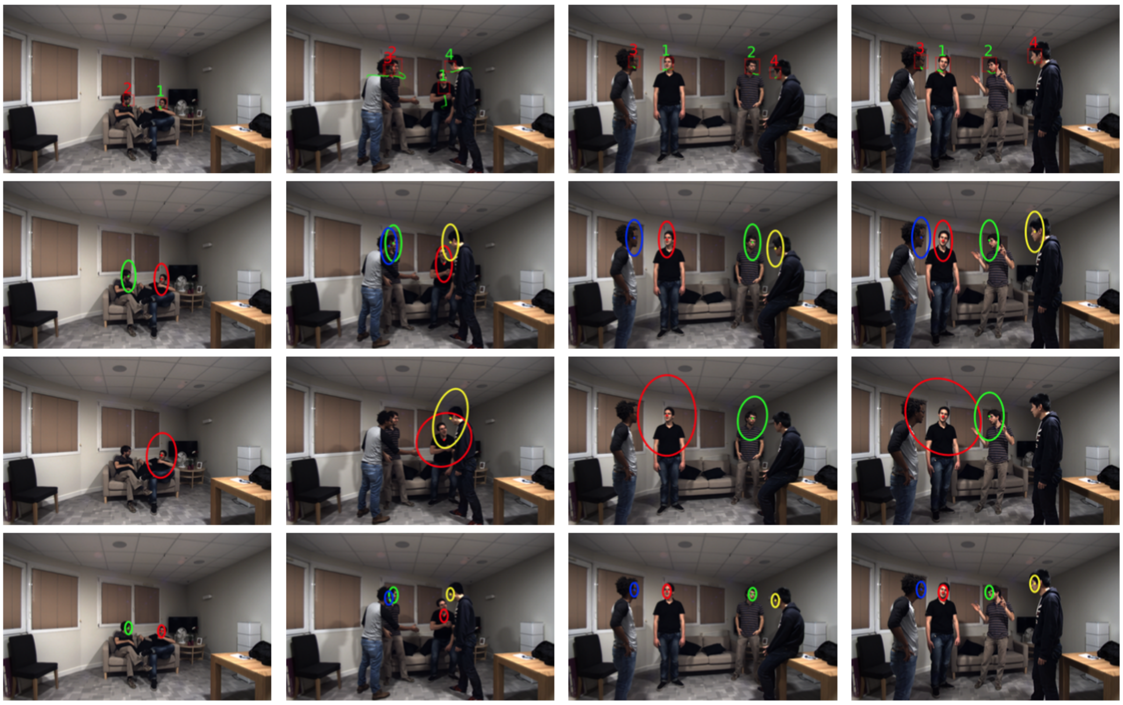


Figure 1. This figure illustrates the audio-visual multiple-person tracking that has been developed by the team [44], [56], [58]. The tracker is based on variational inference [4] and on supervised sound-source localization [9], [26]. Each person is identified with a digit. Green digits denote active speakers while red digits denote silent persons. The next rows show the covariances (uncertainties) associated with the visual (second row), audio (third row) and dynamic (fourth row) contributions for tracking a varying number of persons. Notice the large uncertainty associated with audio and the small uncertainty associated with the dynamics of the tracker. In the light of this example, one may notice the complementary roles played by vision and audio: vision data are more accurate while audio data provide speaker information. These developments have been supported by the European Union via the FP7 STREP project "Embodied Audition for Robots" (EARS) and the ERC advanced grant "Vision and Hearing in Action" (VHIA).

Auditory and visual perception play a complementary role in human interaction. Perception enables people to communicate based on verbal (speech and language) and non-verbal (facial expressions, visual gaze, head movements, hand and body gesturing) communication. These communication modalities have a large degree of overlap, in particular in social contexts. Moreover, the modalities disambiguate each other whenever one of the modalities is weak, ambiguous, or corrupted by various perturbations. Human-computer interaction (HCI) has attempted to address these issues, e.g., using smart & portable devices. In HCI the user is in the loop for decision taking: images and sounds are recorded purposively in order to optimize their quality with respect to the task at hand.

However, the robustness of HCI based on speech recognition degrades significantly as the microphones are located a few meters away from the user. Similarly, face detection and recognition work well under limited lighting conditions and if the cameras are properly oriented towards a person. Altogether, the HCI paradigm cannot be easily extended to less constrained interaction scenarios which involve several users and whenever is important to consider the *social context*.

The PERCEPTION team investigates the fundamental role played by audio and visual perception in human-robot interaction (HRI). The main difference between HCI and HRI is that, while the former is user-controlled, the latter is robot-controlled, namely *it is implemented with intelligent robots that take decisions and act autonomously*. The mid term objective of PERCEPTION is to develop computational models, methods, and applications for enabling non-verbal and verbal interactions between people, analyze their intentions and their dialogue, extract information and synthesize appropriate behaviors, e.g., the robot waves to a person, turns its head towards the dominant speaker, nods, gesticulates, asks questions, gives advices, waits for instructions, etc. The following topics are thoroughly addressed by the team members: audio-visual sound-source separation and localization in natural environments, for example to detect and track moving speakers, inference of temporal models of verbal and non-verbal activities (diarisation), continuous recognition of particular gestures and words, context recognition, and multimodal dialogue.
Video: https://team.inria.fr/perception/demos/lito-video/

3. Research Program

3.1. Audio-Visual Scene Analysis

From 2006 to 2009, R. Horaud was the scientific coordinator of the collaborative European project POP (Perception on Purpose), an interdisciplinary effort to understand visual and auditory perception at the crossroads of several disciplines (computational and biological vision, computational auditory analysis, robotics, and psychophysics). This allowed the PERCEPTION team to launch an interdisciplinary research agenda that has been very active for the last five years. There are very few teams in the world that gather scientific competences spanning computer vision, audio signal processing, machine learning and human-robot interaction. The fusion of several sensorial modalities resides at the heart of the most recent biological theories of perception. Nevertheless, multi-sensor processing is still poorly understood from a computational point of view. In particular and so far, audio-visual fusion has been investigated in the framework of speech processing using close-distance cameras and microphones. The vast majority of these approaches attempt to model the temporal correlation between the auditory signals and the dynamics of lip and facial movements. Our original contribution has been to consider that audio-visual localization and recognition are equally important. We have proposed to take into account the fact that the audio-visual objects of interest live in a three-dimensional physical space and hence we contributed to the emergence of *audio-visual scene analysis* as a scientific topic in its own right. We proposed several novel statistical approaches based on supervised and unsupervised mixture models. The *conjugate mixture model* (CMM) is an unsupervised probabilistic model that allows to cluster observations from different modalities (e.g., vision and audio) living in different mathematical spaces [22], [2]. We thoroughly investigated CMM, provided practical resolution algorithms and studied their convergence properties. We developed several methods for sound localization using two or more microphones [1]. The *Gaussian locally-linear model* (GLLiM) is a partially supervised mixture model that allows to map high-dimensional observations (audio, visual, or concatenations of audio-visual vectors) onto low-dimensional manifolds with a partially known structure [8]. This model is particularly well suited for perception because it encodes both observable and unobservable phenomena. A variant of this model, namely *probabilistic piecewise affine mapping* has also been proposed and successfully applied to the problem of sound-source localization and separation [7]. The European projects HUMAVIPS (2010-2013) coordinated by R. Horaud and EARS (2014-2017), applied audio-visual scene analysis to human-robot interaction.

3.2. Stereoscopic Vision

Stereoscopy is one of the most studied topics in biological and computer vision. Nevertheless, classical approaches of addressing this problem fail to integrate eye/camera vergence. From a geometric point of view, the integration of vergence is difficult because one has to re-estimate the epipolar geometry at every new eye/camera rotation. From an algorithmic point of view, it is not clear how to combine depth maps obtained with different eyes/cameras relative orientations. Therefore, we addressed the more general problem of binocular vision that combines the low-level eye/camera geometry, sensor rotations, and practical algorithms based on global optimization [16], [28]. We studied the link between mathematical and computational approaches to stereo (global optimization and Markov random fields) and the brain plausibility of some of these approaches: indeed, we proposed an original mathematical model for the complex cells in visual-cortex areas V1 and V2 that is based on steering Gaussian filters and that admits simple solutions [17]. This addresses the fundamental issue of how local image structure is represented in the brain/computer and how this structure is used for estimating a dense disparity field. Therefore, the main originality of our work is to address both computational and biological issues within a unifying model of binocular vision. Another equally important problem that still remains to be solved is how to integrate binocular depth maps over time. Recently, we have addressed this problem and proposed a semi-global optimization framework that starts with sparse yet reliable matches and proceeds with propagating them over both space and time. The concept of seed-match propagation has then been extended to TOF-stereo fusion [11].

3.3. Audio Signal Processing

Audio-visual fusion algorithms necessitate that the two modalities are represented in the same mathematical space. Binaural audition allows to extract sound-source localization (SSL) information from the acoustic signals recorded with two microphones. We have developed several methods, that perform sound localization in the temporal and the spectral domains. If a direct path is assumed, one can exploit the *time difference of arrival* (TDOA) between two microphones to recover the position of the sound source with respect to the position of the two microphones. The solution is not unique in this case, the sound source lies onto a 2D manifold. However, if one further assumes that the sound source lies in a horizontal plane, it is then possible to extract the azimuth. We used this approach to predict possible sound locations in order to estimate the direction of a speaker [2]. We also developed a geometric formulation and we showed that with four non-coplanar microphones the azimuth and elevation of a single source can be estimated without ambiguity [1]. We also investigated SSL in the spectral domain. This exploits the filtering effects of the head related transfer function (HRTF): there is a different HRTF for the left and right microphones. The interaural spectral features, namely the ILD (interaural level difference) and IPD (interaural phase difference) can be extracted from the short-time Fourier transforms of the two signals. The sound direction is encoded in these interaural features but it is not clear how to make SSL explicit in this case. We proposed a supervised learning formulation that estimates a mapping from interaural spectral features (ILD and IPD) to source directions using two different setups: audio-motor learning [7] and audio-visual learning [9].

3.4. Visual Reconstruction With Multiple Color and Depth Cameras

For the last decade, one of the most active topics in computer vision has been the visual reconstruction of objects, people, and complex scenes using a multiple-camera setup. The PERCEPTION team has pioneered this field and by 2006 several team members published seminal papers in the field. Recent work has concentrated onto the robustness of the 3D reconstructed data using probabilistic outlier rejection techniques combined with algebraic geometry principles and linear algebra solvers [31]. Subsequently, we proposed to combine 3D representations of shape (meshes) with photometric data [29]. The originality of this work was to represent photometric information as a scalar function over a discrete Riemannian manifold, thus *generalizing image analysis to mesh and graph analysis*. Manifold equivalents of local-structure detectors and descriptors were developed [30]. The outcome of this pioneering work has been twofold: the formulation of a new research topic now addressed by several teams in the world, and allowed us to start a three year collaboration with Samsung Electronics. We developed the novel concept of *mixed camera systems* combining high-resolution color cameras with low-resolution depth cameras [18], [14],[13]. Together with our start-up company 4D Views Solutions and with Samsung, we developed the first practical depth-color multiple-camera multiple-PC system and the first algorithms to reconstruct high-quality 3D content [11].

3.5. Registration, Tracking and Recognition of People and Actions

The analysis of articulated shapes has challenged standard computer vision algorithms for a long time. There are two difficulties associated with this problem, namely how to represent articulated shapes and how to devise robust registration and tracking methods. We addressed both these difficulties and we proposed a novel kinematic representation that integrates concepts from robotics and from the geometry of vision. In 2008 we proposed a method that parameterizes the occluding contours of a shape with its intrinsic kinematic parameters, such that there is a direct mapping between observed image features and joint parameters [23]. This deterministic model has been motivated by the use of 3D data gathered with multiple cameras. However, this method was not robust to various data flaws and could not achieve state-of-the-art results on standard dataset. Subsequently, we addressed the problem using probabilistic generative models. We formulated the problem of articulated-pose estimation as a maximum-likelihood with missing data and we devised several tractable algorithms [21], [20]. We proposed several expectation-maximization procedures applied to various articulated shapes: human bodies, hands, etc. In parallel, we proposed to segment and register articulated shapes represented with graphs by embedding these graphs using the spectral properties of graph Laplacians [6]. This turned out to be a very original approach that has been followed by many other researchers in computer vision and computer graphics.

4. Highlights of the Year

4.1. Highlights of the Year

- As an ERC Advanced Grant holder, Radu Horaud was awarded a Proof of Concept grant for his project Vision and Hearing in Action Laboratory (VHIALab). The project started in February 2018 for a duration of 12 months. Software packages enabling companion robots to robustly interact with multiple users are being developed.
  Website: https://team.inria.fr/perception/projects/poc-vhialab/
- The 2018 winner of the prestigious ACM Special Interest Group on Multimedia (SIGMM) Rising Star Award is Perception team member Dr. Xavier Alameda-Pineda. The award is given in recognition of Xavier's contributions to multimodal social behavior understanding.
  Website: http://sigmm.org/news/sigmm_rising_star_award_2018
- A book was published by Academic Press (Elsevier), entitled "Multimodal Behavior Analysis in the Wild", co-edited by Xavier Alameda Pineda, Elisa Ricci (Fondazione Bruno Kessler and University of Trento) and Nicu Sebe (University of Trento). The book gathers 20 chapters written by 75 researchers from all over the world [53].

5. New Software and Platforms

5.1. ECMPR

Expectation Conditional Maximization for the Joint Registration of Multiple Point Sets
FUNCTIONAL DESCRIPTION: Rigid registration of two or several point sets based on probabilistic matching between point pairs and a Gaussian mixture model

- Participants: Florence Forbes, Manuel Yguel and Radu Horaud
- Contact: Patrice Horaud
- URL: https://team.inria.fr/perception/research/jrmpc/

5.2. Mixcam

Reconstruction using a mixed camera system

KEYWORDS: Computer vision - 3D reconstruction

FUNCTIONAL DESCRIPTION: We developed a multiple camera platform composed of both high-definition color cameras and low-resolution depth cameras. This platform combines the advantages of the two camera types. On one side, depth (time-of-flight) cameras provide coarse low-resolution 3D scene information. On the other side, depth and color cameras can be combined such as to provide high-resolution 3D scene reconstruction and high-quality rendering of textured surfaces. The software package developed during the period 2011-2014 contains the calibration of TOF cameras, alignment between TOF and color cameras, TOF-stereo fusion, and image-based rendering. These software developments were performed in collaboration with the Samsung Advanced Institute of Technology, Seoul, Korea. The multi-camera platform and the basic software modules are products of 4D Views Solutions SAS, a start-up company issued from the PERCEPTION group.

- Participants: Clément Ménier, Georgios Evangelidis, Michel Amat, Miles Hansard, Patrice Horaud, Pierre Arquier, Quentin Pelorson, Radu Horaud, Richard Broadbridge and Soraya Arias
- Contact: Patrice Horaud
- URL: https://team.inria.fr/perception/mixcam-project/

5.3. NaoLab

Distributed middleware architecture for interacting with NAO

FUNCTIONAL DESCRIPTION: This software provides a set of librairies and tools to simply the control of NAO robot from a remote machine. The main challenge is to make easy prototuping applications for NAO ising C++ and Matlab programming environments. Thus NaoLab provides a prototyping-friendly interface to retrieve sensor date (video and sound streams, odometric data...) and to control the robot actuators (head, arms, legs...) from a remote machine.This interface is available on Naoqi SDK, developed by Aldebarab company, Naoqi SDK is needed as it provides the tools to acess the embedded NAO services (low-level motor command, sensor data access...)

- Authors: Fabien Badeig, Quentin Pelorson and Patrice Horaud
- Contact: Patrice Horaud
- URL: https://team.inria.fr/perception/research/naolab/

5.4. Stereo matching and recognition library

KEYWORD: Computer vision

FUNCTIONAL DESCRIPTION: Library providing stereo matching components to rectify stereo images, to retrieve faces from left and right images, to track faces and method to recognise simple gestures

- Participants: Jan Cech, Jordi Sanchez-Riera, Radu Horaud and Soraya Arias
- Contact: Soraya Arias
- URL: https://code.humavips.eu/projects/stereomatch

5.5. Platforms

5.5.1. Audio-Visual Head Popeye+

In 2016 our audio-visual platform was upgraded from Popeye to Popeye+. Popeye+ has two high-definition cameras with a wide field of view. We also upgraded the software libraries that perform synchronized acquisition of audio signals and color images. Popeye+ has been used for several datasets.

Websites:

https://team.inria.fr/perception/projects/popeye/

https://team.inria.fr/perception/projects/popeye-plus/
https://team.inria.fr/perception/avtrack1/
https://team.inria.fr/perception/avdiar/

5.5.2. NAO Robots

The PERCEPTION team selected the companion robot NAO for experimenting and demonstrating various audio-visual skills as well as for developing the concept of social robotics that is able to recognize human presence, to understand human gestures and voice, and to communicate by synthesizing appropriate behavior. The main challenge of our team is to enable human-robot interaction in the real world.

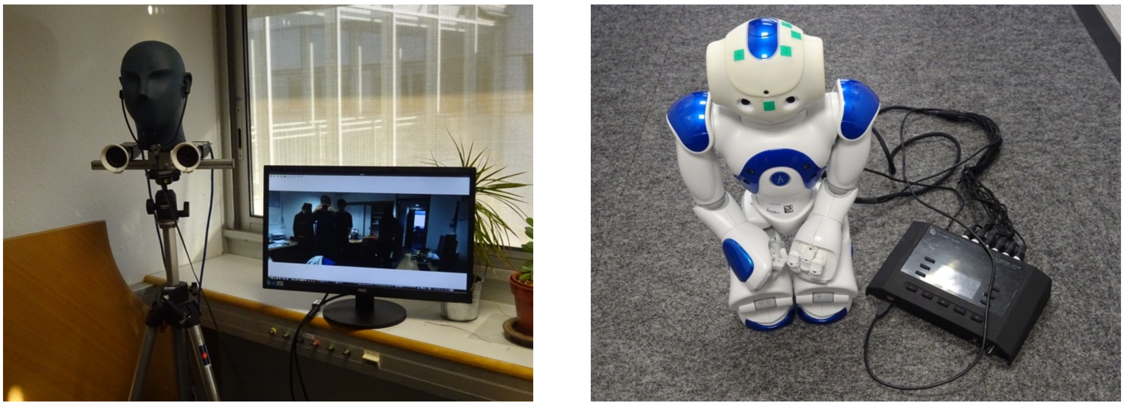


Figure 2. The Popeye+ audio-visual platform (left) delivers high-quality, high-resolution and wide-angle images at 30FPS. The NAO prototype used by PERCEPTION in the EARS STREP project has a twelve-channel spherical microphone array synchronized with a stereo camera pair.

The humanoid robot NAO is manufactured by SoftBank Robotics Europe. Standing, the robot is roughly 60 cm tall, and 35cm when it is sitting. Approximately 30 cm large, NAO includes two CPUs. The first one, placed in the torso, together with the batteries, controls the motors and hence provides kinematic motions with 26 degrees of freedom. The other CPU is placed in the head and is in charge of managing the proprioceptive sensing, the communications, and the audio-visual sensors (two cameras and four microphones, in our case). NAO's on-board computing resources can be accessed either via wired or wireless communication protocols.

NAO's commercially available head is equipped with two cameras that are arranged along a vertical axis: these cameras are neither synchronized nor a significant common field of view. Hence, they cannot be used in combination with stereo vision. Within the EU project HUMAVIPS, Aldebaran Robotics developed a binocular camera system that is arranged horizontally. It is therefore possible to implement stereo vision algorithms on NAO. In particular, one can take advantage of both the robot's cameras and microphones. The cameras deliver VGA sequences of image pairs at 12 FPS, while the sound card delivers the audio signals arriving from all four microphones and sampled at 48 kHz. Subsequently, Aldebaran developed a second binocular camera system to go into the head of NAO v5.

In order to manage the information flow gathered by all these sensors, we implemented several middleware packages. In 2012 we implemented Robotics Services Bus (RSB) developed by the University of Bielefeld. Subsequently (2015-2016) the PERCEPTION team developed NAOLab, a middleware for hosting robotic applications in C, C++, Python and Matlab, using the computing power available with NAO, augmented with a networked PC. In 2017 we abandoned RSB and NAOLab and converted all our robotics software packages to ROS (Robotic Operating System).

Websites:
https://team.inria.fr/perception/nao/
https://team.inria.fr/perception/research/naolab/

6. New Results

6.1. Multichannel Speech Separation and Enhancement Using the Convolutive Transfer Function

We addressed the problem of speech separation and enhancement from multichannel convolutive and noisy mixtures, *assuming known mixing filters*. We proposed to perform the speech separation and enhancement tasks in the short-time Fourier transform domain, using the convolutive transfer function (CTF) approximation [39]. Compared to time-domain filters, CTF has much less taps, consequently it has less near-common zeros among channels and less computational complexity. The work proposes three speech-source recovery methods, namely: (i) the multichannel inverse filtering method, i.e. the multiple input/output inverse theorem (MINT), is exploited in the CTF domain, and for the multi-source case, (ii) a beamforming-like multichannel inverse filtering method applying single source MINT and using power minimization, which is suitable whenever the source CTFs are not all known, and (iii) a constrained Lasso method, where the sources are recovered by minimizing the ℓ_1-norm to impose their spectral sparsity, with the constraint that the ℓ_2-norm fitting cost, between the microphone signals and the mixing model involving the unknown source signals, is less than a tolerance. The noise can be reduced by setting a tolerance onto the noise power. Experiments under various acoustic conditions are carried out to evaluate the three proposed methods. The comparison between them as well as with the baseline methods is presented.

6.2. Speech Dereverberation and Noise Reduction Using the Convolutive Transfer Function

We address the problems of blind multichannel identification and equalization for *joint speech dereverberation and noise reduction*. The standard time-domain cross-relation methods are hardly applicable for blind room impulse response identification due to the near-common zeros of the long impulse responses. We extend the cross-relation formulation to the short-time Fourier transform (STFT) domain, in which the time-domain impulse response is approximately represented by the convolutive transfer function (CTF) with much less coefficients. For the oversampled STFT, CTFs suffer from the common zeros caused by the non-flat-top STFT window. To overcome this, we propose to identify CTFs using the STFT framework with oversampled signals and critically sampled CTFs, which is a good trade-off between the frequency aliasing of the signals and the common zeros problem of CTFs. The phases of the identified CTFs are inaccurate due to the frequency aliasing of the CTFs, and thus only their magnitudes are used. This leads to a non-negative multichannel equalization method based on a non-negative convolution model between the STFT magnitude of the source signal and the CTF magnitude. To recover the STFT magnitude of the source signal and to reduce the additive noise, the ℓ_2-norm fitting error between the STFT magnitude of the microphone signals and the non-negative convolution is constrained to be less than a noise power related tolerance. Meanwhile, the ℓ_1-norm of the STFT magnitude of the source signal is minimized to impose the sparsity [38].
Website: https://team.inria.fr/perception/research/ctf-dereverberation/.

6.3. Speech Enhancement with a Variational Auto-Encoder

We addressed the problem of enhancing speech signals in noisy mixtures using a source separation approach. We explored the use of neural networks as an alternative to a popular speech variance model based on supervised non-negative matrix factorization (NMF). More precisely, we use a variational auto-encoder as a speaker-independent supervised generative speech model, highlighting the conceptual similarities that this approach shares with its NMF-based counterpart. In order to be free of generalization issues regarding the noisy recording environments, we follow the approach of having a supervised model only for the target speech signal, the noise model being based on unsupervised NMF. We developed a Monte Carlo expectation-maximization algorithm for inferring the latent variables in the variational auto-encoder and estimating the unsupervised model parameters. Experiments show that the proposed method outperforms a semi-supervised NMF baseline and a state-of-the-art fully supervised deep learning approach.
Website: https://team.inria.fr/perception/research/ieee-mlsp-2018/.

6.4. Audio-Visual Speaker Tracking and Diarization

We are particularly interested in modeling the interaction between an intelligent device and a group of people. For that purpose we develop audio-visual person tracking methods [36]. As the observed persons are supposed to carry out a conversation, we also include speaker diarization into our tracking methodology. We cast the diarization problem into a tracking formulation whereby the active speaker is detected and tracked over time. A probabilistic tracker exploits the spatial coincidence of visual and auditory observations and infers a single latent variable which represents the identity of the active speaker. Visual and auditory observations are fused using our recently developed weighted-data mixture model [12], while several options for the speaking turns dynamics are fulfilled by a multi-case transition model. The modules that translate raw audio and visual data into image observations are also described in detail. The performance of the proposed method are tested on challenging datasets that are available from recent contributions which are used as baselines for comparison [36].
Websites:
https://team.inria.fr/perception/research/wdgmm/,
https://team.inria.fr/perception/research/speakerloc/,
https://team.inria.fr/perception/research/speechturndet/, and
https://team.inria.fr/perception/research/avdiarization/.

6.5. Tracking Eye Gaze and of Visual Focus of Attention

The visual focus of attention (VFOA) has been recognized as a prominent conversational cue. We are interested in estimating and tracking the VFOAs associated with multi-party social interactions. We note that in this type of situations the participants either look at each other or at an object of interest; therefore their eyes are not always visible. Consequently both gaze and VFOA estimation cannot be based on eye detection and tracking. We propose a method that exploits the correlation between eye gaze and head movements. Both VFOA and gaze are modeled as latent variables in a Bayesian switching state-space model (also named switching Kalman filter). The proposed formulation leads to a tractable learning method and to an efficient online inference procedure that simultaneously tracks gaze and visual focus. The method is tested and benchmarked using two publicly available datasets, Vernissage and LAEO, that contain typical multi-party human-robot and human-human interactions [42].
Website: https://team.inria.fr/perception/research/eye-gaze/.

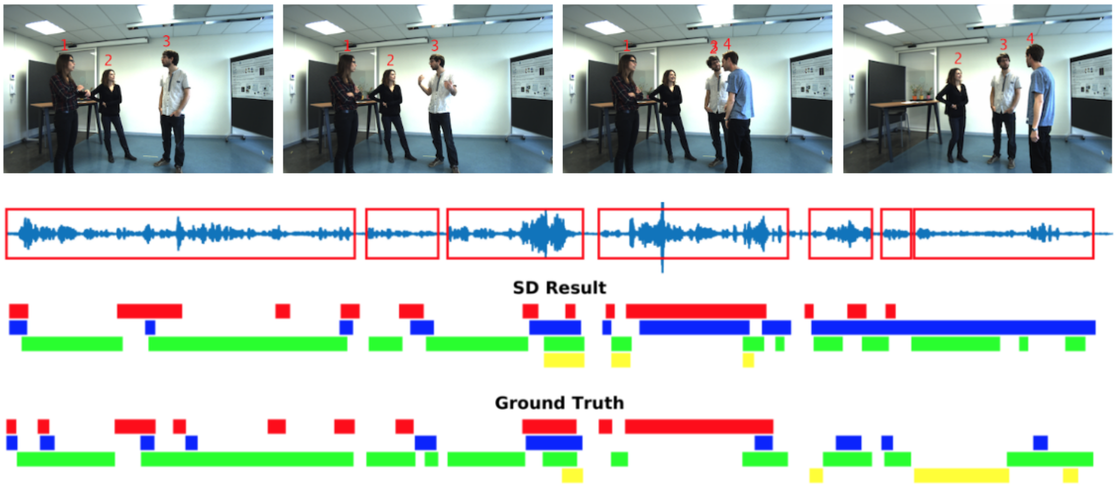


Figure 3. This figure illustrates the audiovisual tracking and diarization method that we have recently developed. First row: A number is associated with each tracked person. Second row: diarization result. Third row: the ground truth diarization. Fourth row: acoustic signal recorded by one of the two microphones.

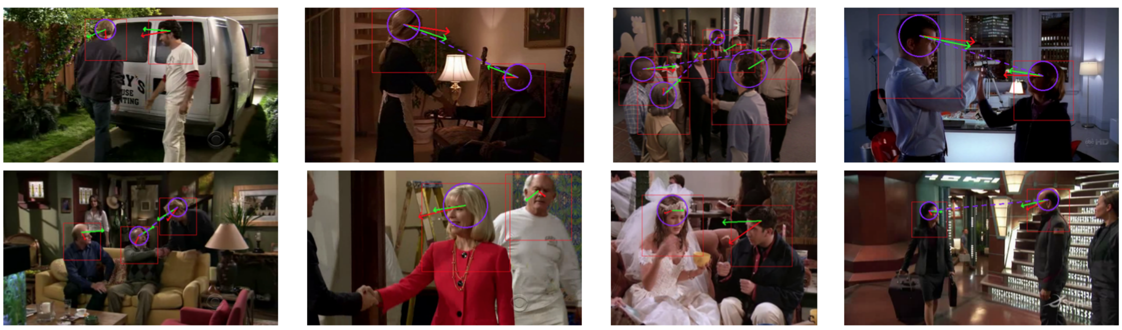


Figure 4. This figure shows some results obtained with the LAEO dataset. The top row shows results obtained with coarse head orientation and the bottom row shows results obtained with fine head orientation. Head orientations are shown with red arrows. The algorithm infers gaze directions (green arrows) and VFOAs (blue circles). People looking at each others are shown with a dashed blue line.

6.6. Variational Bayesian Inference of Multiple-Person Tracking

We addressed the problem of tracking multiple speakers using audio information or via the fusion of visual and auditory information. We proposed to exploit the complementary nature of these two modalities in order to accurately estimate smooth trajectories of the tracked persons, to deal with the partial or total absence of one of the modalities over short periods of time, and to estimate the acoustic status – either speaking or silent – of each tracked person along time, e.g. Figure 1. We proposed to cast the problem at hand into a generative audio-visual fusion (or association) model formulated as a latent-variable temporal graphical model. This may well be viewed as the problem of maximizing the posterior joint distribution of a set of continuous and discrete latent variables given the past and current observations, which is intractable. We propose a variational inference model which amounts to approximate the joint distribution with a factorized distribution. The solutions take the form of closed-form expectation maximization procedures using Gaussian distributions [44], [58], [56] or the von Mises distribution for circular variables [55]. We described in detail the inference algorithms, we evaluate their performance and we compared them with several baseline methods. These experiments show that the proposed audio and audio-visual trackers perform well in informal meetings involving a time-varying number of people.
Websites:
https://team.inria.fr/perception/research/var-av-track/,
https://team.inria.fr/perception/research/audiotrack-vonm/.

6.7. High-Dimensional and Deep Regression

One of the most important achievements for the last years has been the development of high-dimensional to low-dimensional regression methods. The motivation for investigating this problem raised from several problems that appeared both in audio signal processing and in computer vision. Indeed, often the task in data-driven methods is to recover low-dimensional properties and associated parameterizations from high-dimensional observations. Traditionally, this can be formulated as either an unsupervised method (dimensionality reduction of manifold learning) or a supervised method (regression). We developed a learning methodology at the crossroads of these two alternatives: the output variable can be either fully observed or partially observed. This was cast into the framework of linear-Gaussian mixture models in conjunction with the concept of inverse regression. It gave rise to several closed-form and approximate inference algorithms [8]. The method is referred to as *Gaussian locally linear mapping*, or GLLiM. As already mentioned, high-dimensional regression is useful in a number of data processing tasks because the sensory data often lies in high-dimensional spaces. Each one of these tasks required a special-purpose version of our general framework. Sound-source localization was the first to benefit from our formulation. Nevertheless, the sparse nature of speech spectrograms required the development of a GLLiM version that is able to with full-spectrum sounds and to test with sparse-spectrum ones [9]. This could be immediately applied to audio-visual alignment and to sound-source separation and localization [7].

In conjunction with our computer vision work, high-dimensional regression is a very useful methodology since visual features, obtained either by hand-crafted feature extraction methods or using convolutional neural networks, lie in high-dimensional spaces. Such properties as object pose lie in low-dimensional spaces and must be extracted from features. We took such an approach and proposed a head pose estimator [10]. Visual tracking can also benefit from GLLiM. Indeed, it is not practical to track objects based on high-dimensional features. We therefore combined GLLiM with switching linear dynamic systems. In 2018 we proposed a robust deep regression method [46]. In parallel we thoroughly benchmarked and analyzed deep regression tasks using several CNN architectures [57].

6.8. Human-Robot Interaction

Audio-visual fusion raises interesting problems whenever it is implemented onto a robot. Robotic platforms have their own hardware and software constraints. In addition, commercialized robots have economical constraints which leads to the use of cheap components. A robot must be reactive to changes in its environment and hence it must take fast decisions. This often implies that most of the computing resources must be onboard of the robot.

Over the last decade we have tried to do our best to take these constraints into account. Starting from our scientific developments, we put a lot of efforts into robotics implementations. For example, the audio-visual fusion method described in [2] used a specific robotic middleware that allowed fast communication between the robot and an external computing unit. Subsequently we developed a powerful software package that enables distributed computing. We also put a lot of emphasis on the implementation of low-level audio and visual processing algorithms. In particular, our single- and multiple audio source methods were implemented in real time onto the humanoid robot NAO [25], [50]. The multiple person tracker [4] was also implemented onto our robotic platforms [5], e.g. Figure 5.

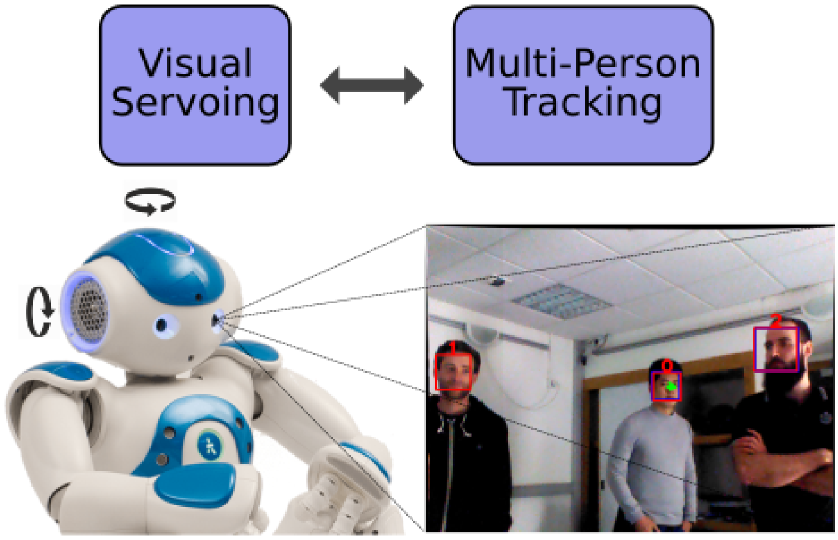


Figure 5. The multi-person tracking method is combined with a visual servoing module. The latter estimates the optimal robot commands and the expected impact of the tracked person locations. The multi-person tracking module refines the locations of the persons with the new observations and the information provided by the visual servoing.

More recently, we investigated the use of reinforcement learning (RL) as an alternative to sensor-based robot control [45], [37]. The robotic task consists of turning the robot head (gaze control) towards speaking people. The method is more general in spirit than visual (or audio) servoing because it can handle an arbitrary number of speaking or non speaking persons and it can improve its behavior online, as the robot experiences new situations. An overview of the proposed method is shown in Fig. 6. The reinforcement learning formulation enables a robot to learn where to look for people and to favor speaking people via a trial-and-error strategy.

Past, present and future HRI developments require datasets for training, validation, test as well as for benchmarking. HRI datasets are challenging because it is not easy to record realistic interactions between a robot and users. RL avoids systematic recourse to annotated datasets for training. In [45], [37] we proposed the use of a simulated environment for pre-training the RL parameters, thus avoiding spending hours of tedious interaction.
Websites:
https://team.inria.fr/perception/research/deep-rl-for-gaze-control/,
https://team.inria.fr/perception/research/mot-servoing/.

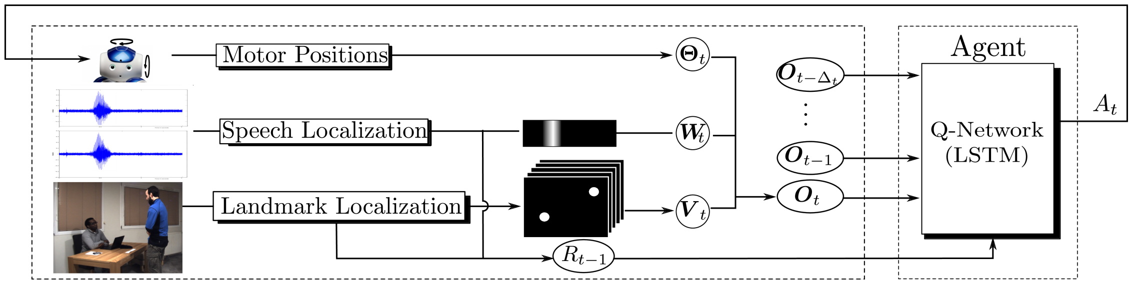


Figure 6. Overview of the proposed deep RL method for controlling the gaze of a robot. At each time index t, audio and visual data are represented as binary maps which, together with motor positions, form the set of observations $\mathbf{O}_t$. A motor action A_t (rotate the head left, right, up, down, or stay still) is selected based on past and present observations via maximization of current and future rewards. The rewards R are based on the number of visible persons as well as on the presence of speech sources in the camera field of view. We use a deep Q-network (DQN) model that can be learned both off-line and on-line. Please consult [45], [37] for further details.

6.9. Generation of Diverse Behavioral Data

We target the automatic generation of visual data depicting human behavior, and in particular how to design a method able to learn the generation of *data diversity*. In particular, we focus on smiles, because each smile is unique: one person surely smiles in different ways (e.g. closing/opening the eyes or mouth). We wonder if given one input image of a neutral face, we can generate multiple smile videos with distinctive characteristics. To tackle this one-to-many video generation problem, we propose a novel deep learning architecture named Conditional MultiMode Network (CMM-Net). To better encode the dynamics of facial expressions, CMM-Net explicitly exploits facial landmarks for generating smile sequences. Specifically, a variational auto-encoder is used to learn a facial landmark embedding. This single embedding is then exploited by a conditional recurrent network which generates a landmark embedding sequence conditioned on a specific expression (e.g. spontaneous smile), implemented as a Conditional LSTM. Next, the generated landmark embeddings are fed into a multi-mode recurrent landmark generator, producing a set of landmark sequences still associated to the given smile class but clearly distinct from each other, we call that a Multi-Mode LSTM. Finally, these landmark sequences are translated into face videos. Our experimental results, see Figure 7, demonstrate the effectiveness of our CMM-Net in generating realistic videos of multiple smile expressions [52].

6.10. Registration of Multiple Point Sets

We have also addressed the rigid registration problem of multiple 3D point sets. While the vast majority of state-of-the-art techniques build on pairwise registration, we proposed a generative model that explains jointly registered multiple sets: back-transformed points are considered realizations of a single Gaussian mixture model (GMM) whose means play the role of the (unknown) scene points. Under this assumption, the joint registration problem is cast into a probabilistic clustering framework. We formally derive an expectation-maximization procedure that robustly estimates both the GMM parameters and the rigid transformations that map each individual cloud onto an under-construction reference set, that is, the GMM means. GMM variances carry rich information as well, thus leading to a noise- and outlier-free scene model as a by-product. A second version of the algorithm is also proposed whereby newly captured sets can be registered online. A thorough discussion and validation on challenging data-sets against several state-of-the-art methods confirm the potential of the proposed model for jointly registering real depth data [35].
Website: https://team.inria.fr/perception/research/jrmpc/

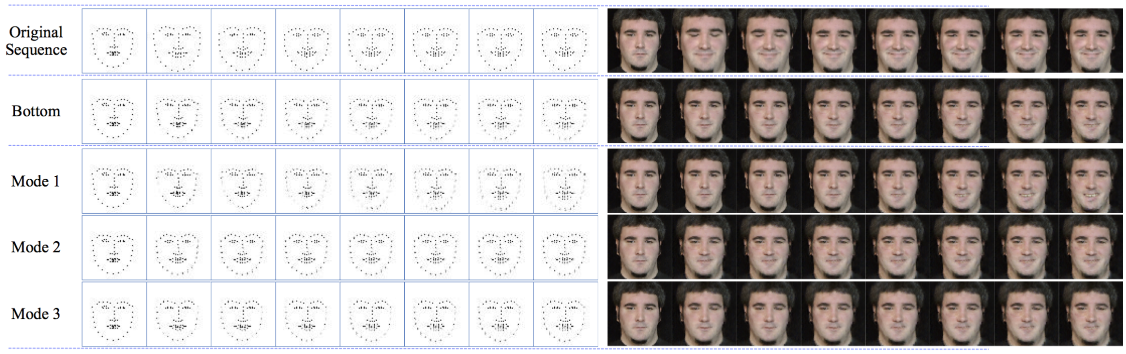


Figure 7. Multi-mode generation example with a sequence: landmarks (left) and associated face images (right) after the landmark-to-image decoding step based on Variational Auto-Encoders. The rows correspond to the original sequence (first), output of the Conditional LSTM (second), and output of the Multi-Mode LSTM (last three rows).

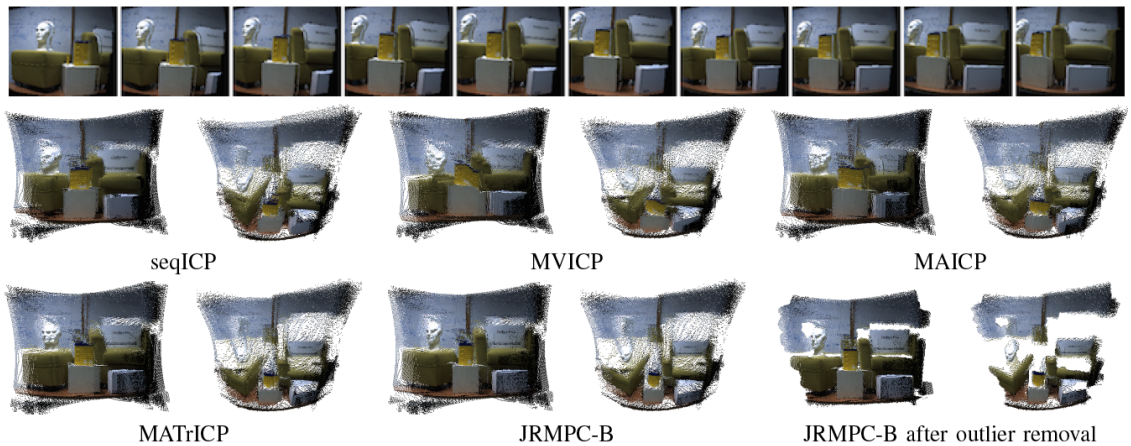


Figure 8. Integrated point clouds from the joint registration of 10 TOF images that record a static scene (EXBI data-set). Top: color images that roughly show the scene content of each range image (occlusions due to cameras baseline may cause texture artefacts). Bottom: front-view and top-view of integrated sets after joint registration. The results obtained with the proposed method (JRMPC-B) are compared with several other methods.

7. Partnerships and Cooperations

7.1. European Initiatives

7.1.1. VHIA

Title: Vision and Hearing in Action

EU framework: FP7

Type: ERC Advanced Grant

Duration: February 2014 - January 2019

Coordinator: Inria

Inria contact: Radu Horaud

'The objective of VHIA is to elaborate a holistic computational paradigm of perception and of perception-action loops. We plan to develop a completely novel twofold approach: (i) learn from mappings between auditory/visual inputs and structured outputs, and from sensorimotor contingencies, and (ii) execute perception-action interaction cycles in the real world with a humanoid robot. VHIA will achieve a unique fine coupling between methodological findings and proof-of-concept implementations using the consumer humanoid NAO manufactured in Europe. The proposed multimodal approach is in strong contrast with current computational paradigms influenced by unimodal biological theories. These theories have hypothesized a modular view, postulating quasi-independent and parallel perceptual pathways in the brain. VHIA will also take a radically different view than today's audiovisual fusion models that rely on clean-speech signals and on accurate frontal-images of faces; These models assume that videos and sounds are recorded with hand-held or head-mounted sensors, and hence there is a human in the loop who intentionally supervises perception and interaction. Our approach deeply contradicts the belief that complex and expensive humanoids (often manufactured in Japan) are required to implement research ideas. VHIA's methodological program addresses extremely difficult issues: how to build a joint audiovisual space from heterogeneous, noisy, ambiguous and physically different visual and auditory stimuli, how to model seamless interaction, how to deal with high-dimensional input data, and how to achieve robust and efficient human-humanoid communication tasks through a well-thought tradeoff between offline training and online execution. VHIA bets on the high-risk idea that in the next decades, social robots will have a considerable economical impact, and there will be millions of humanoids, in our homes, schools and offices, which will be able to naturally communicate with us.
Website: https://team.inria.fr/perception/projects/erc-vhia/

7.1.2. VHIALab

Title: Vision and Hearing in Action Laboratory

EU framework: H2020

Type: ERC Proof of Concept

Duration: February 2018 - January 2019

Coordinator: Inria

Inria contact: Radu Horaud

The objective of VHIALab is the development and commercialization of software packages enabling a robot companion to easily and naturally interact with people. The methodologies developed in ERC VHIA propose state of the art solutions to human-robot interaction (HRI) problems in a general setting and based on audio-visual information. The ambitious goal of VHIALab will be to build software packages based on VHIA, thus opening the door to commercially available multi-party multi-modal human-robot interaction. The methodology investigated in VHIA may well be viewed as a generalization of existing single-user spoken dialog systems. VHIA enables a robot (i) to detect and to locate speaking persons, (ii) to track several persons over time, (iii) to recognize their behavior, and (iv) to extract the speech signal of each person for subsequent speech recognition and face-to-face dialog. These methods will be turned into software packages compatible with a large variety of companion robots. VHIALab will add a strong valorization potential to VHIA by addressing emerging and new market sectors. Industrial collaborations set up in VHIA will be strengthened.

7.2. International Research Visitors

7.2.1. Visits of International Scientists

- Professor Sharon Gannot, Bar Ilan University, Tel Aviv, Israel.
- Professor Tomislav Pribanic, University of Zagreb, Zagreb, Croatia.
- Doctor Christine Evers, Imperial College, London, United Kingdom.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. Member of the Organizing Committees

Xavier Alameda-Pineda organized several workshops in conjunction with IEEE CVPR'18, ECCV'18, and ACM Multimedia'18.

8.1.1.2. Reviewer

Xavier Alameda-PIneda was a reviewer for IEEE CVPR'18, NIPS'18, IEEE ICASSP'18, ACM Multimedia'18 and IEEE ICRA'18.

8.1.2. Journal

8.1.2.1. Member of the Editorial Boards

Radu Horaud is associated editor for the International Journal of Computer Vision and for the IEEE Robotics and Automation Letters.

8.1.2.2. Guest Editor

Xavier Alameda-Pineda was co-guest editor of a special issue of the ACM Transactions on Multimedia Computing Communications and Applications on “Multimodal Understanding of Social, Affective, and Subjective Attributes".

8.1.3. Invited Talks

- Radu Horaud gave an invited talk at the Multimodal Machine Perception Workshop, Google, San Francisco, and at SRI International, Menlo Park, USA, on "Audio-Visual Machine Perception for Human-Robot Interaction".
- Xavier Alameda-Pineda was invited to give a seminar at the University in May 2018 on “Audio-Visual Multiple Speaker Tracking with Robotic Platforms", and
- Xavier Alameda-Pineda gave an invited talk at the SOUND Workshop at Bar-Ilan University, Israel, December 2018 on “Multi-modal Automatic Detection of Social Attractors in Crowded Meetings".

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

- Laurent Girin is professor at Grenoble National Polytechnic Institute (G-INP) where he teaches signal processing and machine learning on the basis of a full professor service (192 hours/year)
- Xavier Alameda-Pineda is involved with the M2 course of the MSIAM Masters Program: “Fundamentals of probabilistic data mining, modeling seminars and projects”, for the practical sessions. Xavier is also preparing a doctoral course on “Learning with Multi-Modal data for Scene Understanding and Human-Robot Interaction” to be taught in spring 2019.

8.2.2. Supervision

- Radu Horaud has supervised the following PhD students: Israel Dejene-Gebru [32], Stéphane Lathuilière [33], Benoît Massé [34], Yutong Ban, Guillaume Delorme and Sylvain Guy.
- Xavier Alameda-PIneda has co-supervised Israel Dejene-Gebru, Yutong Ban, Guillaume Delorme and has supervised Yihong Xu.

8.2.3. Juries

Xavier Alameda-Pineda was reviewer and examiner of the PhD dissertations of Wei Wang (now post-doctoral fellow at EPFL) and of Dr. Dan Xu (now post-doctoral fellow at U. Oxford), both at University of Trento, Italy.

9. Bibliography

Major publications by the team in recent years

[1] X. ALAMEDA-PINEDA, R. HORAUD.*A Geometric Approach to Sound Source Localization from Time-Delay Estimates*, in "IEEE Transactions on Audio, Speech and Language Processing", June 2014, vol. 22, n^{o} 6, p. 1082-1095 [*DOI :* 10.1109/TASLP.2014.2317989], https://hal.inria.fr/hal-00975293

[2] X. ALAMEDA-PINEDA, R. HORAUD.*Vision-Guided Robot Hearing*, in "International Journal of Robotics Research", April 2015, vol. 34, n^{o} 4-5, p. 437-456 [*DOI :* 10.1177/0278364914548050], https://hal.inria.fr/hal-00990766

[3] N. ANDREFF, B. ESPIAU, R. HORAUD.*Visual Servoing from Lines*, in "International Journal of Robotics Research", 2002, vol. 21, n^{o} 8, p. 679–700, http://hal.inria.fr/hal-00520167

[4] S. BA, X. ALAMEDA-PINEDA, A. XOMPERO, R. HORAUD.*An On-line Variational Bayesian Model for Multi-Person Tracking from Cluttered Scenes*, in "Computer Vision and Image Understanding", December 2016, vol. 153, p. 64–76 [*DOI :* 10.1016/J.CVIU.2016.07.006], https://hal.inria.fr/hal-01349763

[5] Y. BAN, X. ALAMEDA-PINEDA, F. BADEIG, S. BA, R. HORAUD.*Tracking a Varying Number of People with a Visually-Controlled Robotic Head*, in "IEEE/RSJ International Conference on Intelligent Robots and Systems", Vancouver, Canada, September 2017, https://hal.inria.fr/hal-01542987

[6] F. CUZZOLIN, D. MATEUS, R. HORAUD.*Robust Temporally Coherent Laplacian Protrusion Segmentation of 3D Articulated Bodies*, in "International Journal of Computer Vision", March 2015, vol. 112, n^{o} 1, p. 43-70 [*DOI :* 10.1007/s11263-014-0754-0], https://hal.archives-ouvertes.fr/hal-01053737

[7] A. DELEFORGE, F. FORBES, R. HORAUD.*Acoustic Space Learning for Sound-Source Separation and Localization on Binaural Manifolds*, in "International Journal of Neural Systems", February 2015, vol. 25, n^{o} 1, 21 [*DOI :* 10.1142/S0129065714400036], https://hal.inria.fr/hal-00960796

[8] A. DELEFORGE, F. FORBES, R. HORAUD.*High-Dimensional Regression with Gaussian Mixtures and Partially-Latent Response Variables*, in "Statistics and Computing", September 2015, vol. 25, n^{o} 5, p. 893-911 [*DOI :* 10.1007/S11222-014-9461-5], https://hal.inria.fr/hal-00863468

[9] A. DELEFORGE, R. HORAUD, Y. Y. SCHECHNER, L. GIRIN.*Co-Localization of Audio Sources in Images Using Binaural Features and Locally-Linear Regression*, in "IEEE Transactions on Audio, Speech and Language Processing", April 2015, vol. 23, n^{o} 4, p. 718-731 [*DOI :* 10.1109/TASLP.2015.2405475], https://hal.inria.fr/hal-01112834

[10] V. DROUARD, R. HORAUD, A. DELEFORGE, S. BA, G. EVANGELIDIS.*Robust Head-Pose Estimation Based on Partially-Latent Mixture of Linear Regressions*, in "IEEE Transactions on Image Processing", March 2017, vol. 26, n^{o} 3, p. 1428 - 1440 [*DOI :* 10.1109/TIP.2017.2654165], https://hal.inria.fr/hal-01413406

[11] G. EVANGELIDIS, M. HANSARD, R. HORAUD.*Fusion of Range and Stereo Data for High-Resolution Scene-Modeling*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", November 2015, vol. 37, n^{o} 11, p. 2178 - 2192 [*DOI :* 10.1109/TPAMI.2015.2400465], https://hal.archives-ouvertes.fr/hal-01110031

[12] I. D. GEBRU, X. ALAMEDA-PINEDA, F. FORBES, R. HORAUD.*EM Algorithms for Weighted-Data Clustering with Application to Audio-Visual Scene Analysis*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", December 2016, vol. 38, n^{o} 12, p. 2402 - 2415 [*DOI :* 10.1109/TPAMI.2016.2522425], https://hal.inria.fr/hal-01261374

[13] M. HANSARD, G. EVANGELIDIS, Q. PELORSON, R. HORAUD.*Cross-Calibration of Time-of-flight and Colour Cameras*, in "Computer Vision and Image Understanding", April 2015, vol. 134, p. 105-115 [*DOI :* 10.1016/J.CVIU.2014.09.001], https://hal.inria.fr/hal-01059891

[14] M. HANSARD, R. HORAUD, M. AMAT, G. EVANGELIDIS.*Automatic Detection of Calibration Grids in Time-of-Flight Images*, in "Computer Vision and Image Understanding", April 2014, vol. 121, p. 108-118 [*DOI :* 10.1016/J.CVIU.2014.01.007], https://hal.inria.fr/hal-00936333

[15] M. HANSARD, R. HORAUD.*Cyclopean geometry of binocular vision*, in "Journal of the Optical Society of America A", September 2008, vol. 25, n^{o} 9, p. 2357-2369 [*DOI :* 10.1364/JOSAA.25.002357], http://hal.inria.fr/inria-00435548

[16] M. HANSARD, R. HORAUD.*Cyclorotation Models for Eyes and Cameras*, in "IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics", March 2010, vol. 40, n^{o} 1, p. 151-161 [*DOI :* 10.1109/TSMCB.2009.2024211], http://hal.inria.fr/inria-00435549

[17] M. HANSARD, R. HORAUD.*A Differential Model of the Complex Cell*, in "Neural Computation", September 2011, vol. 23, n^{o} 9, p. 2324-2357 [*DOI :* 10.1162/NECO_A_00163], http://hal.inria.fr/inria-00590266

[18] M. HANSARD, S. LEE, O. CHOI, R. HORAUD.*Time of Flight Cameras: Principles, Methods, and Applications*, Springer Briefs in Computer Science, Springer, October 2012, 95, http://hal.inria.fr/hal-00725654

[19] R. HORAUD, G. CSURKA, D. DEMIRDJIAN.*Stereo Calibration from Rigid Motions*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", December 2000, vol. 22, n^{o} 12, p. 1446–1452 [*DOI :* 10.1109/34.895977], http://hal.inria.fr/inria-00590127

[20] R. HORAUD, F. FORBES, M. YGUEL, G. DEWAELE, J. ZHANG.*Rigid and Articulated Point Registration with Expectation Conditional Maximization*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", March 2011, vol. 33, n^{o} 3, p. 587-602 [*DOI :* 10.1109/TPAMI.2010.94], http://hal.inria.fr/inria-00590265

[21] R. HORAUD, M. NISKANEN, G. DEWAELE, E. BOYER.*Human Motion Tracking by Registering an Articulated Surface to 3-D Points and Normals*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", January 2009, vol. 31, n^{o} 1, p. 158-163 [*DOI :* 10.1109/TPAMI.2008.108], http://hal.inria.fr/inria-00446898

[22] V. KHALIDOV, F. FORBES, R. HORAUD.*Conjugate Mixture Models for Clustering Multimodal Data*, in "Neural Computation", February 2011, vol. 23, n^{o} 2, p. 517-557 [*DOI :* 10.1162/NECO_A_00074], http://hal.inria.fr/inria-00590267

[23] D. KNOSSOW, R. RONFARD, R. HORAUD.*Human Motion Tracking with a Kinematic Parameterization of Extremal Contours*, in "International Journal of Computer Vision", September 2008, vol. 79, n^{o} 3, p. 247-269 [*DOI :* 10.1007/S11263-007-0116-2], http://hal.inria.fr/inria-00590247

[24] D. KOUNADES-BASTIAN, L. GIRIN, X. ALAMEDA-PINEDA, S. GANNOT, R. HORAUD.*A Variational EM Algorithm for the Separation of Time-Varying Convolutive Audio Mixtures*, in "IEEE/ACM Transactions on Audio, Speech and Language Processing", August 2016, vol. 24, n^{o} 8, p. 1408-1423 [*DOI :* 10.1109/TASLP.2016.2554286], https://hal.inria.fr/hal-01301762

[25] X. LI, L. GIRIN, F. BADEIG, R. HORAUD.*Reverberant Sound Localization with a Robot Head Based on Direct-Path Relative Transfer Function*, in "IEEE/RSJ International Conference on Intelligent Robots and Systems", Daejeon, South Korea, IEEE, October 2016, p. 2819-2826 [*DOI :* 10.1109/IROS.2016.7759437], https://hal.inria.fr/hal-01349771

[26] X. LI, L. GIRIN, R. HORAUD, S. GANNOT.*Estimation of the Direct-Path Relative Transfer Function for Supervised Sound-Source Localization*, in "IEEE/ACM Transactions on Audio, Speech and Language Processing", November 2016, vol. 24, n^{o} 11, p. 2171 - 2186 [*DOI :* 10.1109/TASLP.2016.2598319], https://hal.inria.fr/hal-01349691

[27] X. LI, L. GIRIN, R. HORAUD, S. GANNOT.*Multiple-Speaker Localization Based on Direct-Path Features and Likelihood Maximization with Spatial Sparsity Regularization*, in "IEEE/ACM Transactions on Audio, Speech and Language Processing", October 2017, vol. 25, n^{o} 10, p. 1997 - 2012, 16 pages, 4 figures, 4 tables [*DOI :* 10.1109/TASLP.2017.2740001], https://hal.inria.fr/hal-01413417

[28] M. SAPIENZA, M. HANSARD, R. HORAUD.*Real-time Visuomotor Update of an Active Binocular Head*, in "Autonomous Robots", January 2013, vol. 34, n^{o} 1, p. 33-45 [*DOI :* 10.1007/S10514-012-9311-2], http://hal.inria.fr/hal-00768615

[29] A. ZAHARESCU, E. BOYER, R. HORAUD.*Topology-Adaptive Mesh Deformation for Surface Evolution, Morphing, and Multi-View Reconstruction*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", April 2011, vol. 33, n^{o} 4, p. 823-837 [*DOI :* 10.1109/TPAMI.2010.116], http://hal.inria.fr/inria-00590271

[30] A. ZAHARESCU, E. BOYER, R. HORAUD.*Keypoints and Local Descriptors of Scalar Functions on 2D Manifolds*, in "International Journal of Computer Vision", October 2012, vol. 100, n^{o} 1, p. 78-98 [*DOI :* 10.1007/S11263-012-0528-5], http://hal.inria.fr/hal-00699620

[31] A. ZAHARESCU, R. HORAUD.*Robust Factorization Methods Using A Gaussian/Uniform Mixture Model*, in "International Journal of Computer Vision", March 2009, vol. 81, n° 3, p. 240-258 [*DOI :* 10.1007/S11263-008-0169-X], http://hal.inria.fr/inria-00446987

Publications of the year

Doctoral Dissertations and Habilitation Theses

[32] I. D. GEBRU.*Audio-Visual Analysis In the Framework of Humans Interacting with Robots*, Université Grenoble Alpes, April 2018, https://hal.inria.fr/tel-01774233

[33] S. LATHUILIÈRE.*Deep Regression Models and Computer Vision Applications for Multiperson Human-Robot Interaction*, Université Grenoble Alpes, May 2018, https://tel.archives-ouvertes.fr/tel-01801807

[34] B. MASSÉ.*Gaze Direction in the context of Social Human-Robot Interaction*, Université Grenoble - Alpes, October 2018, https://hal.inria.fr/tel-01936821

Articles in International Peer-Reviewed Journal

[35] G. EVANGELIDIS, R. HORAUD.*Joint Alignment of Multiple Point Sets with Batch and Incremental Expectation-Maximization*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", June 2018, vol. 40, n° 6, p. 1397 - 1410, https://arxiv.org/abs/1609.01466 [*DOI :* 10.1109/TPAMI.2017.2717829], https://hal.inria.fr/hal-01413414

[36] I. GEBRU, S. BA, X. LI, R. HORAUD.*Audio-Visual Speaker Diarization Based on Spatiotemporal Bayesian Fusion*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", July 2018, vol. 40, n° 5, p. 1086 - 1099, https://arxiv.org/abs/1603.09725 [*DOI :* 10.1109/TPAMI.2017.2648793], https://hal.inria.fr/hal-01413403

[37] S. LATHUILIÈRE, B. MASSÉ, P. MESEJO, R. HORAUD.*Neural Network Based Reinforcement Learning for Audio-Visual Gaze Control in Human-Robot Interaction*, in "Pattern Recognition Letters", May 2018, https://arxiv.org/abs/1711.06834 [*DOI :* 10.1016/J.PATREC.2018.05.023], https://hal.inria.fr/hal-01643775

[38] X. LI, S. GANNOT, L. GIRIN, R. HORAUD.*Multichannel Identification and Nonnegative Equalization for Dereverberation and Noise Reduction based on Convolutive Transfer Function*, in "IEEE/ACM Transactions on Audio, Speech and Language Processing", May 2018, vol. 26, n° 10, p. 1755-1768, https://arxiv.org/abs/1711.07911 [*DOI :* 10.1109/TASLP.2018.2839362], https://hal.inria.fr/hal-01645749

[39] X. LI, L. GIRIN, S. GANNOT, R. HORAUD.*Multichannel Speech Separation and Enhancement Using the Convolutive Transfer Function*, in "IEEE/ACM Transactions on Audio, Speech and Language Processing", January 2019, https://arxiv.org/abs/1711.07911 [*DOI :* 10.1109/TASLP.2019.2892412], https://hal.inria.fr/hal-01799809

[40] X. LI, L. GIRIN, R. HORAUD.*Expectation-Maximization for Speech Source Separation using Convolutive Transfer Function*, in "CAAI Transactions on Intelligent Technologies", January 2019 [*DOI :* 10.1049/TRIT.2018.1061], https://hal.inria.fr/hal-01982250

[41] R. T. MARRIOTT, A. PASHEVICH, R. HORAUD.*Plane-extraction from depth-data using a Gaussian mixture regression model*, in "Pattern Recognition Letters", July 2018, vol. 110, p. 44-50, https://arxiv.org/abs/1710.01925 - 2 figures, 1 table [*DOI :* 10.1016/J.PATREC.2018.03.024], https://hal.inria.fr/hal-01663984

[42] B. MASSÉ, S. BA, R. HORAUD.*Tracking Gaze and Visual Focus of Attention of People Involved in Social Interaction*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", November 2018, vol. 40, n^o 11, p. 2711 - 2724, https://arxiv.org/abs/1703.04727 [*DOI :* 10.1109/TPAMI.2017.2782819], https://hal.inria.fr/hal-01511414

[43] D. XU, X. ALAMEDA-PINEDA, J. SONG, E. RICCI, N. SEBE.*Cross-Paced Representation Learning with Partial Curricula for Sketch-based Image Retrieval*, in "IEEE Transactions on Image Processing", September 2018, vol. 27, n^o 9, p. 4410-4421 [*DOI :* 10.1109/TIP.2018.2837381], https://hal.inria.fr/hal-01803694

International Conferences with Proceedings

[44] Y. BAN, X. LI, X. ALAMEDA-PINEDA, L. GIRIN, R. HORAUD.*Accounting for Room Acoustics in Audio-Visual Multi-Speaker Tracking*, in "ICASSP 2018 - IEEE International Conference on Acoustics, Speech and Signal Processing", Calgary, Alberta, Canada, IEEE, April 2018, p. 6553-6557 [*DOI :* 10.1109/ICASSP.2018.8462100], https://hal.inria.fr/hal-01718114

[45] S. LATHUILIÈRE, B. MASSÉ, P. MESEJO, R. HORAUD.*Deep Reinforcement Learning for Audio-Visual Gaze Control*, in "IROS 2018 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Madrid, Spain, October 2018, p. 1-8 [*DOI :* 10.1109/IROS.2018.8594327], https://hal.inria.fr/hal-01851738

[46] S. LATHUILIÈRE, P. MESEJO, X. ALAMEDA-PINEDA, R. HORAUD.*DeepGUM: Learning Deep Robust Regression with a Gaussian-Uniform Mixture Model*, in "ECCV 2018 - European Conference on Computer Vision", Munich, Germany, September 2018, p. 1-16, https://hal.inria.fr/hal-01851511

[47] S. LEGLAIVE, L. GIRIN, R. HORAUD.*A variance modeling framework based on variational autoencoders for speech enhancement*, in "MSLP 2018 - IEEE International Workshop on Machine Learning for Signal Processing", Aalborg, Denmark, IEEE, September 2018, p. 1-6 [*DOI :* 10.1109/MLSP.2018.8516711], https://hal.inria.fr/hal-01832826

[48] X. LI, Y. BAN, L. GIRIN, X. ALAMEDA-PINEDA, R. HORAUD.*A Cascaded Multiple-Speaker Localization and Tracking System*, in "Proceedings of the LOCATA Challenge Workshop - a satellite event of IWAENC 2018", Tokyo, Japan, September 2018, p. 1-5, https://hal.inria.fr/hal-01957137

[49] X. LI, S. GANNOT, L. GIRIN, R. HORAUD.*Multisource MINT Using the Convolutive Transfer Function*, in "ICASSP 2018 - IEEE International Conference on Acoustics, Speech and Signal Processing", Calgary, Alberta, Canada, IEEE, April 2018, p. 756-760 [*DOI :* 10.1109/ICASSP.2018.8462607], https://hal.inria.fr/hal-01718106

[50] X. LI, B. MOURGUE, L. GIRIN, S. GANNOT, R. HORAUD.*Online Localization of Multiple Moving Speakers in Reverberant Environments*, in "10th IEEE Workshop on Sensor Array and Multichannel Signal Processing (SAM 2018)", Sheffield, United Kingdom, IEEE, July 2018, p. 405-409 [*DOI :* 10.1109/SAM.2018.8448423], https://hal.inria.fr/hal-01795462

[51] A. SIAROHIN, E. SANGINETO, S. LATHUILIÈRE, N. SEBE.*Deformable GANs for Pose-based Human Image Generation*, in "IEEE Conference on Computer Vision and Pattern Recognition", Salt Lake City, United States, June 2018, p. 3408-3416, https://arxiv.org/abs/1801.00055 , https://hal.archives-ouvertes.fr/hal-01761539

[52] W. WANG, X. ALAMEDA-PINEDA, D. XU, P. FUA, E. RICCI, N. SEBE.*Every Smile is Unique: Landmark-Guided Diverse Smile Generation*, in "IEEE Conference on Computer Vision and Pattern Recognition", Salk

Lake City, United States, June 2018, p. 7083-7092, https://arxiv.org/abs/1802.01873 , https://hal.inria.fr/hal-01759335

Scientific Books (or Scientific Book chapters)

[53] X. ALAMEDA-PINEDA, E. RICCI, N. SEBE.*Multimodal behavior analysis in the wild: Advances and challenges*, Academic Press (Elsevier), December 2018, https://hal.inria.fr/hal-01858395

[54] L. GIRIN, S. GANNOT, X. LI.*Audio source separation into the wild*, in "Multimodal Behavior Analysis in the Wild", Computer Vision and Pattern Recognition, Academic Press (Elsevier), November 2018, p. 53-78 [*DOI :* 10.1016/B978-0-12-814601-9.00022-5], https://hal.inria.fr/hal-01943375

Other Publications

[55] Y. BAN, X. ALAMEDA-PINEDA, C. EVERS, R. HORAUD.*Tracking Multiple Audio Sources with the von Mises Distribution and Variational EM*, December 2018, https://arxiv.org/abs/1812.08246 - Paper submitted to IEEE Signal Processing Letters, https://hal.inria.fr/hal-01969050

[56] Y. BAN, X. ALAMEDA-PINEDA, L. GIRIN, R. HORAUD.*Variational Bayesian Inference for Audio-Visual Tracking of Multiple Speakers*, December 2018, https://arxiv.org/abs/1809.10961 - Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence, https://hal.inria.fr/hal-01950866

[57] S. LATHUILIÈRE, P. MESEJO, X. ALAMEDA-PINEDA, R. HORAUD.*A Comprehensive Analysis of Deep Regression*, March 2018, https://arxiv.org/abs/1803.08450 - Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence, https://hal.inria.fr/hal-01754839

[58] X. LI, Y. BAN, L. GIRIN, X. ALAMEDA-PINEDA, R. HORAUD.*Online Localization and Tracking of Multiple Moving Speakers in Reverberant Environment*, July 2018, Submitted to Journal on Selected Topics in Signal Processing, https://hal.inria.fr/hal-01851985

[59] X. LI, L. GIRIN, S. GANNOT, R. HORAUD.*Multichannel Online Dereverberation based on Spectral Magnitude Inverse Filtering*, December 2018, https://arxiv.org/abs/1812.08471 - Paper submitted to IEEE/ACM Transactions on Audio, Speech and Language Processing, https://hal.inria.fr/hal-01969041

[60] A. SIAROHIN, G. ZEN, C. MAJTANOVIC, X. ALAMEDA-PINEDA, E. RICCI, N. SEBE.*Increasing Image Memorability with Neural Style Transfer*, August 2018, working paper or preprint, https://hal.inria.fr/hal-01858389

Project-Team PERVASIVE

Pervasive interaction with smart objects and environments

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

IN PARTNERSHIP WITH:
CNRS
Institut polytechnique de Grenoble
Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Robotics and Smart environments

Table of contents

Project-Team PERVASIVE

Creation of the Project-Team: 2017 November 01

Keywords:

Computer Science and Digital Science:

A1.4. - Ubiquitous Systems
A1.6. - Green Computing
A3.4.5. - Bayesian methods
A3.4.6. - Neural networks
A3.4.8. - Deep learning
A3.5.2. - Recommendation systems
A5.1.7. - Multimodal interfaces
A5.1.9. - User and perceptual studies
A5.4. - Computer vision
A5.6. - Virtual reality, augmented reality
A5.7. - Audio modeling and processing
A5.10.2. - Perception
A5.10.3. - Planning
A5.10.4. - Robot control
A5.10.5. - Robot interaction (with the environment, humans, other robots)
A5.11. - Smart spaces
A9. - Artificial intelligence

Other Research Topics and Application Domains:

B1.2.2. - Cognitive science
B2.1. - Well being
B2.5.3. - Assistance for elderly
B6.4. - Internet of things
B6.6. - Embedded systems
B8.1. - Smart building/home
B8.1.1. - Energy for smart buildings
B8.1.2. - Sensor networks for smart buildings
B9.1.1. - E-learning, MOOC

1. Team, Visitors, External Collaborators

Research Scientists

Eric Castelli [CNRS, Researcher, from Feb 2018, HDR]
Sabine Coquillart [Inria, Senior Researcher, HDR]
Thierry Fraichard [Inria, Researcher, HDR]
Emmanuel Mazer [CNRS, Senior Researcher, HDR]

Faculty Members

James Crowley [Team leader, Institut polytechnique de Grenoble, Professor, HDR]
Philippe Dessus [Univ Grenoble Alpes, Professor, HDR]

Patrick Reignier [Institut polytechnique de Grenoble, Professor, HDR]
Dominique Vaufreydaz [Univ Grenoble Alpes, Associate Professor, HDR]

External Collaborators

Jeremy Belot [Univ Grenoble Alpes, from Nov 2018]
Joelle Coutaz [Emeritus, Univ. Grenoble Alpes]
Augustin Lux [Emeritus, Institut polytechnique de Grenoble]
Laura Lassance de Oliveira Morais [Univ Grenoble Alpes Student, until Jul 2018]
Didier Piau [Univ Grenoble Alpes, from Apr 2018]

Technical Staff

Giovanni Balestrieri [Inria, until Nov 2018]
Raffaella Balzarini [Inria]
Maxime Belgodere [Inria]
Nicolas Bonnefond [Inria]
Stanislaw Borkowski [Inria, until Sep 2018]
Mustapha El Majihi [Inria, until May 2018]
Benjamin Girard [Inria, from Oct 2018]
Samuel Heidmann [Inria, from Sep 2018]
Remi Pincent [Inria, until Oct 2018]
Anthony Chavoutier [Inria, from Oct 2018]

PhD Students

Nachwa Abou Bakr [Univ. Grenoble Alpes]
Amr Al-Zouhri Al-Yafi [Univ Grenoble Alpes]
Matteo Ciocca [Univ. Grenoble Alpes]
Rémi Paulin [Inria, until March 2018]
Muhammad Bouguerra [Univ. Annaba (DZ)]
Julien Cumin [Orange Labs, until Sep 2018]
Jose Da Silva Filho [CNPq]
Raphael Frisch [Univ. Grenoble Alpes]
Thomas Guntz [Inria]
Amine Awada [Allied Engineering Group]
Liliya Tsvetanona [Patnering Robotics]
Etienne Balit [Univ Grenoble Alpes, until Jun 2018]

Post-Doctoral Fellows

Marvin Faix [Univ. Grenoble Alpes]
Ivana Schenk [Inria, from Nov 2018]

Administrative Assistant

Alexandra Fitzgerald [Inria]

2. Overall Objectives

2.1. Overall Objectives

Pervasive Interaction develops theories and models for context aware, sociable interaction with systems and services that are dynamically composed from collections of interconnected smart objects. The project uses of situation models as a technological foundation for situated behavior for smart objects and services.

The research program for Pervasive Interaction is designed to respond to the following four research questions:

- Q1: What are the most appropriate computational techniques for acquiring and using situation models for situated behavior by smart objects?
- Q2: What perception and action techniques are most appropriate for situated interaction with smart objects?
- Q3: Can we use situation modelling as a foundation for sociable interaction with smart objects?
- Q4: Can we use situated smart objects as a form of immersive media?

The Pervasive Interaction team was initially formed as a provisional project team of the Inria Grenoble Rhone-Alpes Research Center in April 2016. In November 2017, Pervasive Interaction has been officially designated as an Inria project team. For technical reasons, some publications and results from November and December 2017 are excluded from this report.

3. Research Program

3.1. Situation Models

Situation Modelling, Situation Awareness, Probabilistic Description Logistics

The objectives of this research area are to develop and refine new computational techniques that improve the reliability and performance of situation models, extend the range of possible application domains, and reduce the cost of developing and maintaining situation models. Important research challenges include developing machine-learning techniques to automatically acquire and adapt situation models through interaction, development of techniques to reason and learn about appropriate behaviors, and the development of new algorithms and data structures for representing situation models.

Pervasive Interaction will address the following research challenges:

Techniques for learning and adapting situation models: Hand crafting of situation models is currently an expensive process requiring extensive trial and error. We will investigate combination of interactive design tools coupled with supervised and semi-supervised learning techniques for constructing initial, simplified prototype situation models in the laboratory. One possible approach is to explore developmental learning to enrich and adapt the range of situations and behaviors through interaction with users.

Reasoning about actions and behaviors: Constructing systems for reasoning about actions and their consequences is an important open challenge. We will explore integration of planning techniques for operationalizing actions sequences within behaviors, and for constructing new action sequences when faced with unexpected difficulties. We will also investigate reasoning techniques within the situation modeling process for anticipating the consequences of actions, events and phenomena.

Algorithms and data structures for situation models: In recent years, we have experimented with an architecture for situated interaction inspired by work in human factors. This model organises perception and interaction as a cyclic process in which directed perception is used to detect and track entities, verify relations between entities, detect trends, anticipate consequences and plan actions. Each phase of this process raises interesting challenges questions algorithms and programming techniques. We will experiment alternative programming techniques representing and reasoning about situation models both in terms of difficulty of specification and development and in terms of efficiency of the resulting implementation. We will also investigate the use of probabilistic graph models as a means to better accommodate uncertain and unreliable information. In particular, we will experiment with using probabilistic predicates for defining situations, and maintaining likelihood scores over multiple situations within a context. Finally, we will investigate the use of simulation as technique for reasoning about consequences of actions and phenomena.

Probabilistic Description Logics: In our work, we will explore the use of probabilistic predicates for representing relations within situation models. As with our earlier work, entities and roles will be recognized using multi-modal perceptual processes constructed with supervised and semi-supervised learning [Brdiczka 07], [Barraquand 12]. However, relations will be expressed with probabilistic predicates. We will explore learning based techniques to probabilistic values for elementary predicates, and propagate these through probabilistic representation for axioms using Probabilistic Graphical Models and/or Bayesian Networks.

The challenges in this research area will be addressed through three specific research actions covering situation modelling in homes, learning on mobile devices, and reasoning in critical situations.

3.1.1. Learning Routine patterns of activity in the home.

The objective of this research action is to develop a scalable approach to learning routine patterns of activity in a home using situation models. Information about user actions is used to construct situation models in which key elements are semantic representations of time, place, social role and actions. Activities are encoded as sequences of situations. Recurrent activities are detected as sequences of activities that occur at a specific time and place each day. Recurrent activities provide routines what can be used to predict future actions and anticipate needs and services. An early demonstration has been to construct an intelligent assistant that can respond to and filter communications.

This research action is carried out as part of the doctoral research of Julien Cumin in cooperation with researchers at Orange labs, Meylan. Results are to be published at Ubicomp, Ambient intelligence, Intelligent Environments and IEEE Transactions on System Man and Cybernetics. Julien Cumin will complete and defend his doctoral thesis in 2018.

3.1.2. Learning Patterns of Activity with Mobile Devices

The objective of this research action is to develop techniques to observe and learn recurrent patterns of activity using the full suite of sensors available on mobile devices such as tablets and smart phones. Most mobile devices include seven or more sensors organized in 4 groups: Positioning Sensors, Environmental Sensors, Communications Subsystems, and Sensors for Human-Computer Interaction. Taken together, these sensors can provide a very rich source of information about individual activity.

In this area we explore techniques to observe activity with mobiles devices in order to learn daily patterns of activity. We will explore supervised and semi-supervised learning to construct systems to recognize places and relevant activities. Location and place information, semantic time of day, communication activities, interpersonal interactions, and travel activities (walking, driving, riding public transportation, etc.) are recognized as probabilistic predicates and used to construct situation models. Recurrent sequences of situations will be detected and recorded to provide an ability to predict upcoming situations and anticipate needs for information and services.

Our goal is to develop a theory for building context aware services that can be deployed as part of the mobile applications that companies such as SNCF and RATP use to interact with clients. For example, a current project concerns systems that observe daily travel routines for the Paris region RATP metro and SNCF commuter trains. This system learns individual travel routines on the mobile device without the need to divulge information about personal travel to a cloud based system. The resulting service will consult train and metro schedules to assure that planned travel is feasible and to suggest alternatives in the case of travel disruptions. Similar applications are under discussion for the SNCF inter-city travel and Air France for air travel.

This research action is conducted in collaboration with the Inria Startup Situ8ed. The current objective is to deploy and evaluate a first prototype App during 2017. Techniques will be used commercially by Situ8ed for products to be deployed as early as 2019.

3.1.3. Bibliography

[Brdiczka 07] O. Brdiczka, "Learning Situation Models for Context-Aware Services", Doctoral Thesis of the INPG, 25 may 2007.

[Barraquand 12] R. Barraquand, "Design of Sociable Technologies", Doctoral Thesis of the University Grenoble Alps, 2 Feb 2012.

3.2. Perception of People, Activities and Emotions

Machine perception is fundamental for situated behavior. Work in this area will concern construction of perceptual components using computer vision, acoustic perception, accelerometers and other embedded sensors. These include low-cost accelerometers [Bao 04], gyroscopic sensors and magnetometers, vibration sensors, electromagnetic spectrum and signal strength (wifi, bluetooth, GSM), infrared presence detectors, and bolometric imagers, as well as microphones and cameras. With electrical usage monitoring, every power switch can be used as a sensor [Fogarty 06], [Coutaz 16]. We will develop perceptual components for integrated vision systems that combine a low-cost imaging sensors with on-board image processing and wireless communications in a small, low-cost package. Such devices are increasingly available, with the enabling manufacturing technologies driven by the market for integrated imaging sensors on mobile devices. Such technology enables the use of embedded computer vision as a practical sensor for smart objects.

Research challenges to be addressed in this area include development of practical techniques that can be deployed on smart objects for perception of people and their activities in real world environments, integration and fusion of information from a variety of sensor modalities with different response times and levels of abstraction, and perception of human attention, engagement, and emotion using visual and acoustic sensors.

Work in this research area will focus on three specific Research Actions

3.2.1. Multi-modal perception and modeling of activities

The objective of this research action is to develop techniques for observing and scripting activities for common household tasks such as cooking and cleaning. An important part of this project involves acquiring annotated multi-modal datasets of activity using an extensive suite of visual, acoustic and other sensors. We are interested in real-time on-line techniques that capture and model full body movements, head motion and manipulation actions as 3D articulated motion sequences decorated with semantic labels for individual actions and activities with multiple RGB and RGB-D cameras.

We will explore the integration of 3D articulated models with appearance based recognition approaches and statistical learning for modeling behaviors. Such techniques provide an important enabling technology for context aware services in smart environments [Coutaz 05], [Crowley 15], investigated by Pervasive Interaction team, as well as research on automatic cinematography and film editing investigated by the Imagine team [Gandhi 13] [Gandhi 14] [Ronfard 14] [Galvane 15]. An important challenge is to determine which techniques are most appropriate for detecting, modeling and recognizing a large vocabulary of actions and activities under different observational conditions.

We will explore representations of behavior that encodes both temporal-spatial structure and motion at multiple levels of abstraction. We will further propose parameters to encode temporal constraints between actions in the activity classification model using a combination of higher-level action grammars [Pirsiavash 14] and episodic reasoning [Santofimia 14] [Edwards 14].

Our method will be evaluated using long-term recorded dataset that contains recordings of activities in home environments. This work is carried out in the doctoral research of Nachwa Abou Bakr in cooperation with Remi Ronfard of the Imagine Team of Inria.

3.2.2. Perception with low-cost integrated sensors

In this research action, we will continue work on low-cost integrated sensors using visible light, infrared, and acoustic perception. We will continue development of integrated visual sensors that combine micro-cameras and embedded image processing for detecting and recognizing objects in storage areas. We will combine visual and acoustic sensors to monitor activity at work-surfaces. Low cost real-time image analysis procedures will be designed that acquire and process images directly as they are acquired by the sensor.

Bolometric image sensors measure the Far Infrared emissions of surfaces in order to provide an image in which each pixel is an estimate of surface temperature. Within the European MIRTIC project, Grenoble startup, ULIS has created a relatively low-cost Bolometric image sensor (Retina) that provides small images of 80 by 80 pixels taken from the Far-infrared spectrum. Each pixel provides an estimate of surface temperature. Working with Schneider Electric, engineers in the Pervasive Interaction team had developed a small, integrated sensor that combines the MIRTIC Bolometric imager with a microprocessor for on-board image processing. The package has been equipped with a fish-eye lens so that an overhead sensor mounted at a height of 3 meters has a field of view of approximately 5 by 5 meters. Real-time algorithms have been demonstrated for detecting, tracking and counting people, estimating their trajectories and work areas, and estimating posture.

Many of the applications scenarios for Bolometric sensors proposed by Schneider Electric assume a scene model that assigns pixels to surfaces of the floor, walls, windows, desks or other items of furniture. The high cost of providing such models for each installation of the sensor would prohibit most practical applications. We have recently developed a novel automatic calibration algorithm that determines the nature of the surface under each pixel of the sensor.

Work in this area will continue to develop low-cost real time infrared image sensing, as well as explore combinations of far-infrared images with RGB and RGBD images.

3.2.3. Observing and Modelling Competence and Awareness from Eye-gaze and Emotion

Humans display awareness and emotions through a variety of non-verbal channels. It is increasingly possible to record and interpret such information with available technology. Publicly available software can be used to efficiently detect and track face orientation using web cameras. Concentration can be inferred from changes in pupil size [Kahneman 66]. Observation of Facial Action Units [Ekman 71] can be used to detect both sustained and instantaneous (micro-expressions) displays of valence and excitation. Heart rate can be measured from the Blood Volume Pulse as observed from facial skin color [Poh 11]. Body posture and gesture can be obtained from low-cost RGB sensors with depth information (RGB+D) [Shotton 13] or directly from images using detectors learned using deep learning [Ramakrishna 14]. Awareness and attention can be inferred from eye-gaze (scan path) and fixation using eye-tracking glasses as well as remote eye tracking devices [Holmqvist 11]. Such recordings can be used to reveal awareness of the current situation and to predict ability to respond effectively to opportunities and threats.

This work is supported by the ANR project CEEGE in cooperation with the department of NeuroCognition of Univ. Bielefeld. Work in this area includes the Doctoral research of Thomas Guntz to be defended in 2019.

3.2.4. Bibliography

[Bao 04] L. Bao, and S. S. Intille. "Activity recognition from user-annotated acceleration data.", IEEE Pervasive computing. Springer Berlin Heidelberg, pp1-17, 2004.

[Fogarty 06] J. Fogarty, C. Au and S. E. Hudson. "Sensing from the basement: a feasibility study of unobtrusive and low-cost home activity recognition." In Proceedings of the 19th annual ACM symposium on User interface software and technology, UIST 2006, pp. 91-100. ACM, 2006.

[Coutaz 16] J. Coutaz and J.L. Crowley, A First-Person Experience with End-User Development for Smart Homes. IEEE Pervasive Computing, 15(2), pp.26-39, 2016.

[Coutaz 05] J. Coutaz, J.L. Crowley, S. Dobson, D. Garlan, "Context is key", Communications of the ACM, 48 (3), 49-53, 2005.

[Crowley 15] J. L. Crowley and J. Coutaz, "An Ecological View of Smart Home Technologies", 2015 European Conference on Ambient Intelligence, Athens, Greece, Nov. 2015.

[Gandhi 13] Vineet Gandhi, Remi Ronfard. "Detecting and Naming Actors in Movies using Generative Appearance Models", Computer Vision and Pattern Recognition, 2013.

[Gandhi 14] Vineet Gandhi, Rémi Ronfard, Michael Gleicher. "Multi-Clip Video Editing from a Single Viewpoint", European Conference on Visual Media Production, 2014

[Ronfard 14] R. Ronfard, N. Szilas. "Where story and media meet: computer generation of narrative discourse". Computational Models of Narrative, 2014.

[Galvane 15] Quentin Galvane, Rémi Ronfard, Christophe Lino, Marc Christie. "Continuity Editing for 3D Animation". AAAI Conference on Artificial Intelligence, Jan 2015.

[Pirsiavash 14] Hamed Pirsiavash , Deva Ramanan, "Parsing Videos of Actions with Segmental Grammars", Computer Vision and Pattern Recognition, p.612-619, 2014.

[Edwards 14] C. Edwards. 2014, "Decoding the language of human movement". Commun. ACM 57, 12, 12-14, November 2014.

[Kahneman 66] D. Kahneman and J. Beatty, Pupil diameter and load on memory. Science, 154(3756), 1583-1585, 1966.

[Ekman 71] P. Ekman and W.V. Friesen, Constants across cultures in the face and emotion. Journal of Personality and Social Psychology, 17(2), 124, 1971.

[Poh 11] M. Z. Poh, D. J. McDuff, and R. W. Picard, Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Trans. Biomed. Eng., 58, 7–11, 2011.

[Shotton 13] J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore, Real-time human pose recognition in parts from single depth images. Commun. ACM, 56, 116–124, 2013.

[Ramakrishna 14] V. Ramakrishna, D. Munoz, M. Hebert, J. A. Bagnell and Y. Sheikh, Pose machines: Articulated pose estimation via inference machines. In European Conference on Computer Vision (ECCV 2016), pp. 33-47, Springer, 2014.

[Cao 17] Z. Cao, T. Simon, S. E. Wei and Y. Sheikh, Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2017), IEEE Press, pp. 1302-1310, July, 2017.

[Holmqvist 11] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer, Eye Tracking: A Comprehensive Guide to Methods and Measures, OUP Oxford: Oxford, UK, 2011.

3.3. Sociable Interaction with Smart Objects

Reeves and Nass argue that a social interface may be the truly universal interface [Reeves 98]. Current systems lack ability for social interaction because they are unable to perceive and understand humans or to learn from interaction with humans. One of the goals of the research to be performed in Pervasive Interaction is to provide such abilities.

Work in research area RA3 will demonstrate the use of situation models for sociable interaction with smart objects and companion robots. We will explore the use of situation models as a representation for sociable interaction. Our goal in this research is to develop methods to endow an artificial agent with the ability to acquire social common sense using the implicit feedback obtained from interaction with people. We believe that such methods can provide a foundation for socially polite man-machine interaction, and ultimately for other forms of cognitive abilities. We propose to capture social common sense by training the appropriateness of behaviors in social situations. A key challenge is to employ an adequate representation for social situations.

Knowledge for sociable interaction will be encoded as a network of situations that capture both linguistic and non-verbal interaction cues and proper behavioral responses. Stereotypical social interactions will be represented as trajectories through the situation graph. We will explore methods that start from simple stereotypical situation models and extending a situation graph through the addition of new situations and the splitting of existing situations. An important aspect of social common sense is the ability to act appropriately in social situations. We propose to learn the association between behaviors and social situation using reinforcement learning. Situation models will be used as a structure for learning appropriateness of actions and behaviors that may be chosen in each situation, using reinforcement learning to determine a score for appropriateness based on feedback obtained by observing partners during interaction.

Work in this research area will focus on four specific Research Actions

3.3.1. Moving with people

Our objective in this area is to establish the foundations for robot motions that are aware of human social situation that move in a manner that complies with the social context, social expectations, social conventions and cognitive abilities of humans. Appropriate and socially compliant interactions require the ability for real time perception of the identity, social role, actions, activities and intents of humans. Such perception can be used to dynamically model the current situation in order to understand the situation and to compute the appropriate course of action for the robot depending on the task at hand.

To reach this objective, we propose to investigate three interacting research areas:

- Modeling the context and situation of human activities for motion planning
- Planning and acting in a social context.
- Identifying and modeling interaction behaviors.

In particular, we will investigate techniques that allow a tele-presence robot, such as the BEAM system, to autonomously navigate in crowds of people as may be found at the entry to a conference room, or in the hallway of a scientific meeting.

3.3.2. Understanding and communicating intentions from motion

This research area concerns the communication through motion. When two or more people move as a group, their motion is regulated by implicit rules that signal a shared sense of social conventions and social roles. For example, moving towards someone while looking directly at them signals an intention for engagement. In certain cultures, subtle rules dictate who passes through a door first or last. When humans move in groups, they implicitly communicate intentions with motion. In this research area, we will explore the scientific literature on proxemics and the social sciences on such movements, in order to encode and evaluate techniques for socially appropriate motion by robots.

3.3.3. Socially aware interaction

This research area concerns socially aware man-machine interaction. Appropriate and socially compliant interaction requires the ability for real time perception of the identity, social role, actions, activities and intents of humans. Such perception can be used to dynamically model the current situation in order to understand the context and to compute the appropriate course of action for the task at hand. Performing such interactions in manner that respects and complies with human social norms and conventions requires models for social roles and norms of behavior as well as the ability to adapt to local social conventions and individual user preferences. In this research area, we will complement research area 3.2 with other forms of communication and interaction, including expression with stylistic face expressions rendered on a tablet, facial gestures, body motions and speech synthesis. We will experiment with use of commercially available tool for spoken language interaction in conjunction with expressive gestures.

3.3.4. Stimulating affection and persuasion with affective devices.

This research area concerns technologies that can stimulate affection and engagement, as well as induce changes in behavior. When acting as a coach or cooking advisor, smart objects must be credible and persuasive. One way to achieve this goal is to express affective feedbacks while interacting. This can be done using sound, light and/or complex moves when the system is composed of actuators.

Research in this area will address 3 questions:

1. How do human perceive affective signals expressed by smart objects (including robots)?
2. How does physical embodiment effect perception of affect by humans?
3. What are the most effective models and tools for animation of affective expression?

Both the physical form and the range of motion have important impact on the ability of a system to inspire affection. We will create new models to propose a generic animation model, and explore the effectiveness of different forms of motion in stimulating affect.

3.3.5. Bibliography

[Reeves 98] B. Reeves and C. Nass, The Media Equation: how People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press, 1998.

3.4. Interaction with Pervasive Smart Objects and Displays

Currently, the most effective technologies for new media for sensing, perception and experience are provided by virtual and augmented realities [Van Krevelen 2010]. At the same time, the most effective means to augment human cognitive abilities are provided by access to information spaces such as the world-wide-web using graphical user interfaces. A current challenge is to bring these two media together.

Display technologies continue to decrease exponentially, driven largely by investment in consumer electronics as well as the overall decrease in cost of microelectronics. A consequence has been an increasing deployment of digital displays in both public and private spaces. This trend is likely to accelerate, as new technologies and growth in available communications bandwidth enable ubiquitous low-cost access to information and communications.

The arrival of pervasive displays raises a number of interesting challenges for situated multi-modal interaction. For example:

1. Can we use perception to detect user engagement and identify users in public spaces?
2. Can we replace traditional pointing hardware with gaze and gesture based interaction?
3. Can we tailor information and interaction for truly situated interaction, providing the right information at the right time using the right interaction modality?
4. How can we avoid information overload and unnecessary distraction with pervasive displays?

It is increasingly possible to embed sensors and displays in clothing and ordinary devices, leading to new forms of tangible and wearable interaction with information. This raises challenges such as

1. What are the tradeoffs between large-scale environmental displays and wearable displays using technologies such as e-textiles and pico-projector?
2. How can we manage the tradeoffs between implicit and explicit interaction with both tangible and wearable interaction?
3. How can we determine the appropriate modalities for interaction?
4. How can we make users aware of interaction possibilities without creating distraction?

In addition to display and communications, the continued decrease in microelectronics has also driven an exponential decrease in cost of sensors, actuators, and computing resulting in an exponential growth in the number of smart objects in human environments. Current models for systems organization are based on centralized control, in which a controller or local hub, orchestrates smart objects, generally in connection with cloud computing. This model creates problems with privacy and ownership of information. An alternative is to organize local collections of smart objects to provide distributed services without the use of a centralized controller. The science of ecology can provide an architectural model for such organization.

This approach raises a number of interesting research challenges for pervasive interaction:

1. Can we devise distributed models for multi-modal fusion and interaction with information on heterogeneous devices?
2. Can we devise models for distributed interaction that migrates over available devices as the user changes location and task?
3. Can we manage migration of interaction over devices in a manner that provides seamless immersive interaction with information, services and media?
4. Can we provide models of distributed interaction that conserve the interaction context as services migrate?

Research Actions for Interaction with Pervasive Smart Objects for the period 2017 - 2020 include

3.4.1. Situated interaction with pervasive displays

The emergence of low-cost interactive displays will enable a confluence of virtual and physical environments. Our goal in this area is to go beyond simple graphical user interfaces in such environments to provide immersive multi-sensorial interaction and communication. A primary concern will be interaction technologies that blend visual with haptic/tactile feedback and 3D interaction and computer vision. We will investigate the use of visual-tactile feedback as well as vibratory signals to augment multi-sensorial interaction and communication. The focus will be on the phenomena of immersive interaction in real worlds that can be made possible by the blending of physical and virtual in ordinary environments.

3.4.2. Wearable and tangible interaction with smart textiles and wearable projectors

Opportunities in this area result from the emergence of new forms of interactive media using smart objects. We will explore the use of smart objects as tangible interfaces that make it possible to experience and interact with information and services by grasping and manipulating objects. We will explore the use of sensors and actuators in clothing and wearable devices such as gloves, hats and wrist bands both as a means of unobtrusively sensing human intentions and emotional states and as a means of stimulating human senses through vibration and sound. We will explore the new forms of interaction and immersion made possible by deploying interactive displays over large areas of an environment.

3.4.3. Pervasive interaction with ecologies of smart objects in the home

In this research area, we will explore and evaluate interaction with ecologies of smart objects in home environments. We will explore development of a range of smart objects that provide information services, such as devices for Episodic Memory for work surfaces and storage areas, devices to provide energy efficient control of environmental conditions, and interactive media that collect and display information. We propose to develop a new class of socially aware managers that coordinate smart objects and manage logistics in functional areas such as the kitchen, living rooms, closets, bedrooms, bathroom or office.

3.4.4. Bibliography

[Van Krevelen 10] D. W. F. Van Krevelen and R. Poelman, A survey of augmented reality technologies, applications and limitations. International Journal of Virtual Reality, 9(2), 1, 2010

4. Application Domains

4.1. Smart Energy Systems

Participants: Amr Alyafi, Amine Awada, Patrick Reignier Partners: UMR G-SCOP, UMR LIG (Persuasive Interaction, IIHM), CEA Liten, PACTE, Vesta Systems and Elithis.

Work in this area explores techniques for a user centric energy management system, where user needs and tacit knowledge drive the search of solutions. These are calculated using a flexible energy model of the living areas. The system is personified by energy consultants with which building actors such as building owners, building managers, technical operators but also occupants, can interact in order to co-define energy strategies, benefiting of both assets: tacit knowledge of human actors, and measurement with computation capabilities of calculators. Putting actors in the loop, i.e. making energy not only visible but also controllable is the needed step before large deployment of energy management solutions. It is proposed to develop interactive energy consultants for all the actors, which are energy management aided systems embedding models in order to support the decision making processes. MIRROR (interactive monitoring), WHAT-IF (interactive quantitative simulation), EXPLAIN (interactive qualitative simulation), SUGGEST-AND-ADJUST (interactive management) and RECOMMEND (interactive diagnosis) functionalities will be developed.

4.2. E-Textile

Participant: Sabine Coquillart

Partner: LIMSI

Collaboration with the HAPCO team from LIMSI on e-textiles. A patent application has been filed related to this work:

- F. Bimbard, M. Bobin, M. Ammi, S. Coquillart "Procédé de conception d'un capteur de flexion textile piézorésistif à partir de fils fonctionnels", Patent Application, 2017.

4.3. Interaction with Pervasive Media

Participants: Sabine Coquillart, Jingtao Chen

Partners: Inria GRA, GIPSA, G-SCOP

Pseudo-haptic feedback is a technique aiming to simulate haptic sensations without active haptic feedback devices. Peudo-haptic techniques have been used to simulate various haptic feedbacks such as stiffness, torques, and mass. In the framework of Jingtao Chen PhD thesis, a novel pseudo-haptic experiment has been set up. The aim of this experiment is to study the EMG signals during a pseudo-haptic task. A stiffness discrimination task similar to the one published in Lecuyer's PhD thesis has been chosen. The experimental set-up has been developed, as well as the software controlling the experiment. Pre-tests are under way. They will be followed by the tests with subjects.

4.4. Bayesian Reasoning

Participants: Emmanuel Mazer, Raphael Frisch, Augustin Lux, Didier Piau, Marvin Faix, Jeremy Belot

The development of modern computers is mainly based on increase of performances and decrease of size and energy consumption, with no notable modification of the basic principles of computation. In particular, all the components perform deterministic and exact operations on sets of binary signals. These constraints obviously impede further sizable progresses in terms of speed, miniaturization and power consumption. The main goal of the project MicroBayes is to investigate a radically different approach, using stochastic bit streams to perform computations. The aim of this project is to show that stochastic architectures can outperform standard computers to solve complex inference problems both in terms of execution speed and of power consumption. We will demonstrate the feasibility on two applications involving low level information processing from sensor signals, namely sound source localization and separation.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

James Crowley has received the ICMI Sustained Achievements award at the 2018 International Conference on Multimodal Interaction at Boulder Colorado in Oct. 2018.

6. New Software and Platforms

6.1. Platforms

The AmiQual4Home Innovation Factory is an open research facility for innovation and experimentation with human-centered services based on the use of large-scale deployment of interconnected digital devices capable of perception, action, interaction and communication. The Innovation Factory is composed of a collection of workshops for rapid creation of prototypes, surrounded by a collection of living labs and supported by an industrial innovation and transfer service. Creation of the Innovation Factory has been made possible by a grant from French National programme Investissement d'avenir, together with substantial contributions of resources by Grenoble INP, Univ Joseph Fourier, UPMF, CNRS, Schneider Electric and the Communaute de Communes of Montbonnot. The objective is to provide the academic and industrial communities with an open platform to enable research on design, integration and evaluation of systems and services for smart habitats.

The AmiQual4Home Innovation Factory is a unique combination of three different innovation instruments:

1. Workshops for rapid prototyping of devices that embed perception, action, interaction and communication in ordinary objects based on the MIT FabLab model,
2. Facilities for real-world test and evaluation of devices and services organized as open Living Labs,
3. Resources for assisting students, researchers, entrepreneurs and industrial partners in creating new economic activities.

The AmiQual4Home Innovation Factory works with the Inovallee TARMAC technology incubator as well as the SAT Linksium to proved innovation and transfer services to enable students, researchers and local entrepreneurs to create and grow new commercial activities based on smart objects and services.

7. New Results

7.1. Using Attention to Address Human-Robot Motion

Participants: Thierry Fraichard, Rémi Paulin, Patrick Reignier.

To capture the specificity of robot motion among people, we choose the term **Human-Robot Motion** (HRM) [0], to denote the study of how robots should move among people. HRM is about designing robots whose motions are deemed socially **acceptable** from a human point of view while remaining **safe**.

After 15 years of research on HRM, the main concept that has emerged is that of *social spaces*, *i.e.* regions of the environment that people consider as psychologically theirs [33], any intrusion in their social space will be a source of discomfort. Such social spaces are characterized by the position of the person, *i.e.* "Personal Space", or the activity they are currently engaged in, *i.e.* "Interaction Space" and "Activity Space". The most common approach in HRM is to define costmaps on such social spaces: the higher the cost, the less desirable it is to be there. The costmaps are then used for navigation purposes, *e.g.* [37] and [36].

Social spaces are of course relevant to HRM but they have limitations. First, it is not straightforward to define them; what is their shape or size, especially in cluttered environments? Second, it seems obvious that there is more to acceptability than geometry only: the appearance of a robot and its velocity will also influence the way it is perceived by people. Finally, social spaces can be conflicting because when a robot needs to interact with a person, it is very likely that it will have to penetrate a social space.

[0] In reference to Human-Robot Interaction (HRI), *i.e.* the study of the interactions, in the broad sense of the word, between people and robots.

To complement social spaces, we have started to explore whether human attention could be useful to address HRM vis-à-vis the acceptability aspect. Why attention? The answer is straightforward: the acceptability of a robot motion is directly related to the way it is perceived by a person hence our interest in human attention. For a person, attention is a cognitive mechanism for filtering the person's sensory information (to avoid an overwhelming amount of information) [35]. It controls where and to what the person's attentional resources are allocated.

In 2014, we introduced the concept of **attention field**, *i.e.* a predictor of the amount of attention that a person allocates to the robot when the robot is in a given state. In [32], the attention field was computed thanks to a computational model of attention proposed in [34] in the context of ambient applications and pervasive systems. In this model, attentional resources are focused on a single specific area of the person's visual space (as per the zoom lens model [31]). Later studies have demonstrated that the situation is more complex and that attentional resources can be distributed over multiple objects in the visual space [35].

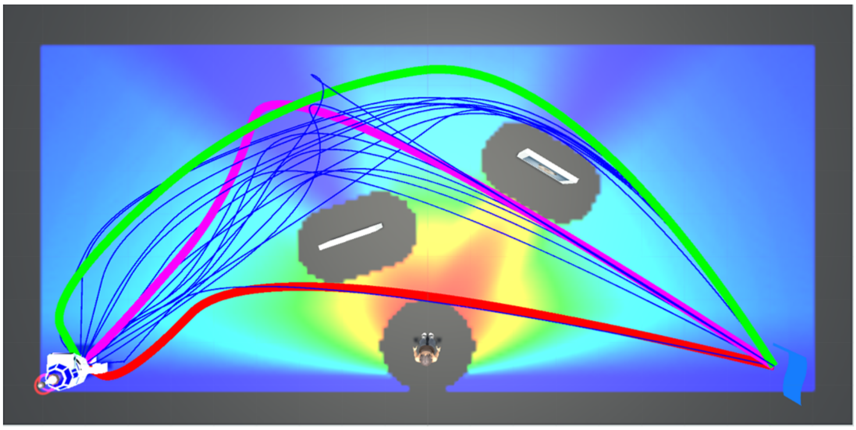


Figure 1. Motions with different attentional properties in a scenario featuring a person watching at paintings in a museum and a robot which is to travel from left to right: less distracting (green) vs. shortest (red) motions are depicted. The purple motion is a trade-off solution.

In 2018, we have developed a novel **computational model of attention** that takes this property into account. This model is used to compute the attention field for a robot. The attention field is then used to define different **attentional properties** for the robot's motions such as distraction or surprise. The relevance of the attentional properties for HRM have been demonstrated on a proof-of-concept **acceptable motion planner** on various case studies where a robot is assigned different tasks. The multi-criteria nature of motion planning in the context of HRM led to the design of an acceptable motion planner based upon a state-of-the-art many-objective optimization algorithm. It shows how to compute acceptable motions that are non-distracting and non-surprising, but also motions that convey the robot's intention to interact with a person. All these contributions have been presented in the PhD of Rémi Paulin [6] and the conference article [26].

7.2. Simulating Haptic Sensations

Participants: Jingtao Chen, Sabine Coquillart

Partners: Inria GRA, LIG, GIPSA, G-SCOP

Pseudo-haptic feedback is a technique aiming to simulate haptic sensations without active haptic feedback devices. Peudo-haptic techniques have been used to simulate various haptic feedbacks such as stiffness, torques, and mass. In the framework of the Persyval project, a novel pseudo-haptic experiment has been set up. The aim of this experiment is to study the force and EMG signals during a pseudo-haptic task. A stiffness discrimination task similar to the one published in Lecuyer's PhD thesis has been chosen. The experimental set-up has been developed, as well as the software controlling the experiment. Pre-tests have been conducted. They have been followed by formal tests with subjects.

7.3. Observing and Modeling Awareness and Expertise During Problem Solving

Participants: Thomas Guntz, Dominique Vaufreydaz, James Crowley, Philippe Dessus, Raffaella Balzarini

7.3.1. Observing and Modelling Competence and Awareness from Eye-gaze and Emotion

We have constructed an instrument for capturing and interpreting multimodal signals of humans engaged in solving challenging problems. Our instrument captures eye gaze, fixations, body postures, and facial expressions signals from humans engaged in interactive tasks on a touch screen. We use a 23 inch Touch-Screen computer, a Kinect 2.0 mounted 35 cm above the screen to observe the subject, a 1080p Webcam for a frontal view, a Tobii Eye-Tracking bar (Pro X2-60 screen-based) and two adjustable USB-LED for lighting condition control. A wooden structure is used to rigidly mount the measuring equipment in order to assure identical sensor placement and orientation for all recordings.

As a pilot study, we observed expert chess players engaged in solving problems of increasing difficulty [Guntz et al 18a]. Our initial hypothesis was that we could directly detect awareness of significant configurations of chess pieces (chunks) from eye-scan and physiological measurements of emotion in reaction to game situation. The pilot experiment demonstrated that this initial hypothesis was overly simplistic.

In order to better understand the phenomena observed in our pilot experiment, we have constructed a model of the cognitive processes involved, using theories from cognitive science and classic (symbolic) artificial intelligence. This model is a very partial description that allows us to ask questions and make predictions to guide future experiments. Our model posits that experts reason with a situation model that is strongly constrained by limits to the number of entities and relations that may be considered at a time. This limitation forces subjects to construct abstract concepts (chunks) to describe game play, in order to explore alternative moves. Expert players retain associations of situations with emotions in long-term memory. The rapid changes in emotion correspond to recognition of previously encountered situations during exploration of the game tree. Recalled emotions guide selection of situation models for reasoning. This hypothesis is in accordance with Damasio's Somatic Marker hypothesis, which posits that emotions guide behavior, particularly when cognitive processes are overloaded [Damasio 91].

Our hypothesis is that the subject uses the evoked emotions to select from the many possible situations for reasoning about moves during orientation and exploration. With this interpretation, the player rapidly considers partial descriptions as situations composed of a limited number of perceived chunks. Recognition of situations from experience evokes emotions that are displayed as face expressions and body posture.

With this hypothesis, valence, arousal and dominance are learned from experience and associated with chess situations in long-term memory to guide reasoning in chess. Dominance corresponds to the degree of experience with the recognized situation. As players gain experience with alternate outcomes for a situation, they become more assured in their ability to spot opportunities and avoid dangers. Valence corresponds to whether the situation is recognized as favorable (providing opportunities) or unfavorable (creating threats). Arousal corresponds to the imminence of a threat or opportunity. A defensive player will give priority to reasoning about unfavorable situations and associated dangers. An aggressive player will seek out high valence situations. All players will give priority to situations that evoke strong arousal. The amount of effort that player will expend exploring a situation can determined by dominance.

In 2019 we will conduct an additional experiment designed confirm and explore this hypothesis. Results will be reported in a journal paper (under preparation) as well as in the doctoral thesis of Thomas Guntz, to be defended in late 2019.

7.3.2. Bibliography

[Damasio 91] Damasio, A., Somatic Markers and the Guidance of Behavior. New York: Oxford University Press. pp. 217–299, 1991.

[Guntz et al. 18a] T. Guntz, R. Balzarini, D. Vaufreydaz, and J.L. Crowley, "Multimodal Observation and Classification of People Engaged in Problem Solving: Application to Chess Players". Multimodal Technologies and Interaction, Vol 2 No. 2, p11, 2018.

[Guntz et al. 18b] T. Guntz, J.L. Crowley, D. Vaufreydaz, R. Balzarini, P. Dessus, The Role of Emotion in Problem Solving: first results from observing chess, Workshop on Modeling Cognitive Processes from Multimodal Data, at the 2018 ACM International Conference on Multimodal Interaction, ICMI 2018, Oct 2018.

7.4. Learning Routine Patterns of Activity in the Home

Participants: Julien Cumin, James Crowley

Other Partners: Fano Ramparany, Greg Lefevre (Orange Labs)

During the month of February 2017, we have collected 4 weeks of data on daily activities within the Amiqual4Home Smart Home Living lab apartment. This dataset was presented at the international Conference on Ubiquitous Computing and Ambient Intelligence, UCAmI 2017, at Bethlehem PA, in Nov 2017 and is currently available for download from the Amiqual4Home web server (http://amiqual4home.inria.fr/en/orange4home/)

The objective of this research action is to develop a scalable approach to learning routine patterns of activity in a home using situation models. Information about user actions is used to construct situation models in which key elements are semantic time, place, social role, and actions. Activities are encoded as sequences of situations. Recurrent activities are detected as sequences of activities that occur at a specific time and place each day. Recurrent activities provide routines that can be used to predict future actions and anticipate needs and services. An early demonstration has been to construct an intelligent assistant that can respond to and filter inter-personal communications.

7.5. Bayesian Reasoning

Participants: Emmanuel Mazer, Raphael Frisch, Marvin Faix, Augustin Lux, Didier Piau, Jeremy Belot.

To overcome the ever growing needs in computing power, alternative computing paradigms have been developed such as stochastic architectures. These latter have found substantial interests for energy efficient implementations in artificial intelligence. In particular, mixing stochastic computing with Bayesian models makes a promising paradigm for non-conventional computational architectures dedicated to Bayesian inference. The ability to deal with uncertainty and adapt its computational accuracy is some of the advantages of these computing approaches.

During 2018 we have designed a first hardware prototype to localize a sound source with a stochastic machine. The goal of this project was to provide a proof of concept of stochastic machines by implementing an autonomous platform of sound source localization. It includes an sound acquisition module, a pre-processing circuit, and the stochastic machine. The platform has been implemented on an Altera Cyclone V FPGA and validated functionally with digital simulations. Several optimization to improve size and power consumption have been proposed. Results in terms of computation time, power and used FPGA resources allowed to assess their impact on future design. The same architecture of stochastic machine was also analyzed in simulation to provide design guidelines for our next design [25].

Further, we have proposed a way to reduce the memory needs of our architecture by sharing a memory between the processing units (in collaboration with TIMA and C2M -Université Paris Sud). This optimization reduces the area and the cost of our architecture. However, its impact on power consumption is not obvious. Therefore, we designed an integrated circuit (ASIC) with our original and optimized proposals. We synthesized the VHDL description of the circuit in the FDSOI 28nm technology from STMicroelectronics. Notice that the memory has been implemented thanks to a SRAM memory compiler. The results highlight that the optimized machine significantly reduces both the circuit area (by 30%) and the power consumption (by 35%). Nevertheless, the simulations showed that, in the optimized version, the memory represents nearly 60 % of our circuit area and more than 55% of the power consumption. According to the latest literature, the Magnetic Random Access Memory (MRAM) technology provides some promising features and would approximately reduce by a factor of 20 the memory area. Moreover, this feature should drastically impact the power consumption. Thus, our future works will focus on the implementation of Bayesian machines using MRAM instead of SRAM. A poster describing this work was presented at the International Conference on rebooting Computing.

We have proposed (in collaboration with ISIR - Université Paris Sorbonne)a new way to localize several sound sources using a Bayesian model. This multi-source localization algorithm is fast and can readily be implemented on our stochastic machine (Paper submitted at ICASSP 2019). The Figure 2 shows the location of the source and of the microphones in the simulated environment. The Figure 3 shows the posterior distribution of the location of one source using a short frame and the Figure 4 shows the result using fifty frames. As the frame are very short the localization of the two sources is readily obtained and it is used as a bootstrap for the source separation algorithm .

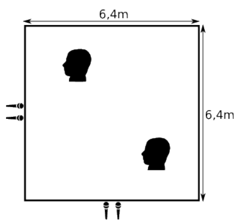
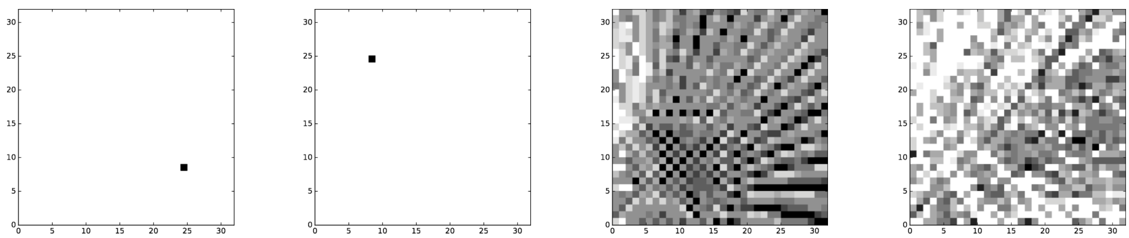


Figure 2. Simulated room setup.



(a) Frame 1 (b) Frame 3 (c) Frame 29 (d) Frame 46

Figure 3. Posterior distribution maps for a single source obtained for 4 very short time-frames of a given 50-frame bloc.

We devised and successfully tested a Bayesian model for the source separation problem. The model assumes the localization of the sources are known. The inference - retrieving the sound emitted by each source from the

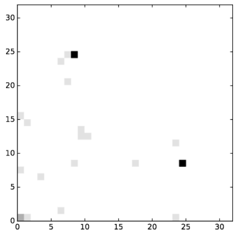


Figure 4. Final distribution map after fusion over 50 frames. The two black squares correspond to the actual positions of the two sources.

mixed signals obtained with several microphones - takes place in a very high dimensional space. Nevertheless, the Gibbs algorithm is well suited to solve the problem when the location of the sources are known. A very efficient implementation of this algorithm was tested with a realistic sound simulator using human voices. The algorithm can be implemented on a sampling machine and the corresponding stochastic architecture has been devised. It is currently implemented on an FPGA.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Pervasive has a contract with Orange Labs, Meylan, for supervision of the doctoral research of Julien Cumin for Learning daily routines by observing activity in a smart home

Members of the Pervasive interaction team are work with Orange Labs on techniques for observing activity and learning routines in a smart home. Activity is observed by monitoring use of electrical appliances and Communications media (Telephone, Television, Internet). Activities are described using Bayesian Situation Modeling techniques demonstrated in earlier projects. A log of daily activities is used to discover daily routines expressed as temporal sequences of contexts, where each context is expressed as a network of situations. Experiments will be performed using the Smart home living lab that has been constructed as part of the EquipEx Amiqual4home.

8.1.1. Toutilo project

Participants: Stan Borkoswki, Dominique Vaufreydaz, Joelle Coutaz, James Crowley, Giovanni Balestrieri, Anthony Chavoutier
Partners: Inria, Touti Terre

Touti Terre is a pioneer startup in the use of agricultural robotics for market gardening, developing innovative solutions to make working the land easier and farms sustainable. The Toutirobo-2 Inria innovation lab proposes the design of on overall IT solution for their cobot solution: the Toutilo robot. This project aims at providing significant time and productivity gains for its users. Thanks to the support of the experimentation and prototyping platform Amiqual4Home, members of the Pervasive team contribute to this project on several innovation topics: farm and vehicle management, autonomous guidance, navigation and planning, and interaction systems adapted to farm jobs. boisi

8.1.2. IRT Silver Economy

Participants: James Crowley, Maxime Belgodere

Partners: CEA, Schneider Electric.

Members of the Pervasive Interaction team are working with the CEA and Schneider Electric to develop environmental sensors that can detect when a hospital patient or elderly person has fallen and is unable to get up. The project uses an infrared Bolometric image sensor to observe human activity. Image processing and fall detection logic are to be performed by an embedded image processor on board.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. LabEx Persyval, Project RHUM, "Robots in Human Environments"

Participants: Thierry Fraichard, Rémi Paulin, Patrick Reignier.

Partners: GIPSA, Inria, LIG, LJK and TIMC.
Dates:*[Sep. 15-Dec. 19].*

The RHUM project from the LabEx Persyval (ANR-11-LABX-0025-01) brings together ten teams from different labs from the Grenoble academic scene: GIPSA, Inria, LIG, LJK and TIMC. Its goal is to tackle scientific problems related to active perception, navigation in human environments, learning and adaptation of robots behaviors for social interaction. PERVASIVE contributes to the navigation in human environments aspects.

9.1.2. ANR Project Involved

Participants: Amr Al-Zhouri Al-Yafi, Patrick Reignier.

Other Partners: UMR G-SCOP, UMR LIG (Persuasive Interaction, IIHM, Getalp), CEA Liten, PACTE, Vesta Systems and Elithis.

Dates: Jan 2015 to Dec 2018

The ANR project Involved focuses on bringing solutions to building actors for upcoming challenges in energy management in residential buildings. The project explores a user centric energy management system, where user needs and tacit knowledge drive the search of solutions. These are calculated using a flexible energy model of the living areas. The system is personified by energy consultants with which building actors such as building owners, building managers, technical operators but also occupants, can interact with in order to co-define energy strategies, benefiting of both assets: tacit knowledge of human actors, and measurement with computation capabilities of calculators. Putting actors in the loop, i.e. making energy not only visible but also controllable is the needed step before large deployment of energy management solutions.

The project will develop interactive energy consultants for all the actors, providing energy management aided systems embedding models in order to support the decision making processes. MIRROR (interactive monitoring), WHAT-IF (interactive quantitative simulation), EXPLAIN (interactive qualitative simulation), SUGGEST- AND-ADJUST (interactive management) and RECOMMEND (interactive diagnosis) functionalities will be developed.

9.1.3. ANR Project CEEGE: Chess Expertise from Eye Gaze and Emotion

Participants: James Crowley, Dominique Vaufreydaz, Rafaellea Balzarini, Thomas Guntz

Other Partners: Dept of NeuroCognition, CITEN, Bielefeld University

Dates: Jan 2016 to Dec 2019

CEEGE is a multidisciplinary scientific research project conducted by the Inria PRIMA team in cooperation with the Dept of Cognitive Neuroscience at the University of Bielefeld. The primary impacts will be improved scientific understanding in the disciplines of Computer Science and Cognitive NeuroScience. The aim of this project is to experimentally evaluate and compare current theories for mental modelling for problem solving and attention, as well as to refine and evaluate techniques for observing the physiological reactions of humans to situation that inspire pleasure, displeasure, arousal, dominance and fear.

In this project, we will observe the visual attention, physiological responses and mental states of subject with different levels of expertise solving classic chess problems, and participating in chess matches. We will observe chess players using eye-tracking, sustained and instantaneous face-expressions (micro-expressions), skin conductivity, blood flow (BVP), respiration, posture and other information extracted from audio-visual recordings and sensor readings of players. We will use the recorded information to estimate the mental constructs with which the players understand the game situation. Information from visual attention as well as physiological reactions will be used to determine and model the degree to which a player understands the game situation in terms of abstract configurations of chess pieces. This will provide a structured environment that we will use for experimental evaluation of current theories of mental modeling and emotional response during problem solving and social interaction.

The project is organized in three phases. During the first phase, we will observe individual players of different levels of chess expertise solving known chess problems. We will correlate scan-path from eye tracking and other information about visual attention to established configurations of pieces and known solutions to chess problems. This will allow us to construct a labeled corpus of chess play that can be used to evaluate competing techniques for estimating mental models and physiological responses. In a second phase, we will observe the attention and face expressions of pairs of players of different levels of chess ability during game play. In particular, we will seek to annotate and segment recordings with respect to the difficulty of the game situation as well as situations that elicit particularly strong physiological reactions. In the final phase, we will use these recordings to evaluate the effectiveness of competing techniques for mental modeling and observation of emotions in terms of their abilities to predict the chess abilities of players, game outcomes and individual moves and player self reports. Results of our work will be published in scientific conferences and journals concerned with cognitive science and cognitive neuroscience as well as computer vision, multimodal interaction, affective computing and pervasive computing. Possible applications include construction of systems that can monitor the cognitive abilities and emotional reactions of users of interactive systems to provide assistance that is appropriate but not excessive, companion systems that can aid with active healthy ageing, and tutoring systems that can assist users in developing skills in a variety of domains including chess.

9.1.4. CDP EcoSesa - Cross Disciplinary Project of the ComUE UGA

Participants: James Crowley, Patrick Reignier, Rafaellea Balzarini Dates: Jan 2017 to Dec 2020

Cities and their energy systems are undergoing profound transformations. Electric Power networks are being transformed from centralized, high capacity, generating plants, dimensioned to meet peak loads to decentralized, local, production based on intermittent renewable sources. This transformation is made possible by integration of information and energy technologies, new energy materials and components, and the rapid spread of pervasive computing. The result is a change in the socio-economics of energy distribution, and a change in the role of users from passive consumers to active participants in a dynamically fluctuating energy market. Many cities worldwide have initiated research projects and experiments to accelerate the spread of clean technologies. However, these initiatives generally focus on a specific issue that depends on the priorities and preferences of the local decision makers and stakeholders. At the same time, academic research has generally been confined to specialized silos in energy materials and management systems, in Social Sciences as well as in Information and Communication Technologies (ICT), resulting in piecemeal knowledge.

The vision of Eco-SESA is to address the problems resulting from the transition to clean decentralized energy production based on renewable sources with a holistic integrated humansystem approach. The project will address the development of Safe, Efficient, Sustainable and Accessible energy systems, from the individual end-user to dynamic communities of stakeholders at the district and grid levels.

Pervasive is involved in two research front of the project :

- Interactive systems to involve occupants of buildings
- Emerging behaviors from individual to communities

9.1.5. ANR VALET

Participant: Dominique Vaufreydaz.

Partners: Inria (Pervasive and Chroma teams for Inria Rhône-Alpes, RITS in Paris), Ircyyn (Nantes), AKKA (Paris)
Dates:*[2016-2018].*

The ANR VALET project investigates two aspects of car sharing. In the first one, a novel approach for solving vehicle redistribution problem is proposed by managing an autonomous platoons guided by professional drivers. The second aspect concerns autonomous parking of shared cars when they arrived at their destination parking lot. In this project, our researches address the prediction of pedestrians' behaviors during urban fleet movements and during parking phases. The PhD student (Pavan Vashista) recruited in this project focus on integrating models of human behaviors to evaluate the risk that surrounding pedestrians encounter the trajectory of the VALET vehicles. His PhD thesis started in February 2016 is co-supervised by Anne Spalanzani (Chroma team) and Dominique Vaufreydaz.

9.1.6. ANR HIANIC

Participant: Dominique Vaufreydaz.

Partners: ARMEN and PACCE teams from LS2N laboratory (Nantes), Inria (Pervasive and Chroma teams for Inria Rhône-Alpes, RITS in Paris), MAGMA from LIG laboratory (Grenoble).
Dates:*[2018-2021].*

The HIANIC project proposes to endow autonomous vehicles with smart behaviors (cooperation, negotiation, socially acceptable movements) to address problems that arise when autonomous cars are mixed with pedestrians in urban shared environment. It aims at developing new technologies in term of autonomous navigation in dense and human populated traffic. In order to contribute to urban safety and intelligent mobility, the HIANIC project also explores the complex problem of sociable interactions between pedestrians and cars while sharing the same urban environment.

In this project, Dominique Vaufreydaz works jointly with the Chroma team on perceiving pedestrians and their behaviors around autonomous cars and on interaction between autonomous vehicles and pedestrians.

9.1.7. LabEx Persyval - Project MicroBayes: Probabilistic Machines for Low-level Sensor Interpretation

Participants: Emmanuel Mazer, Raphael Frisch Other Partners: Laurent Girin (GIPSA Lab), Didier Piau (L'Institut Fourier)

Dates: Nov 2016 to Nov 2019

The project MicroBayes builds on results of the recently completed EC FET Open project BAMBI to explore a new technique for Blind source separation and acoustic signal location using a new form of Bayesian Computer. The techniques have recently been demonstrated using a software simulation. Current plans are to implement and demonstrate the Bayesian computer using an FPGA. By the end of the project we expect to produce a hardware implementation suitable for use in low-cost low-power applications.

9.1.8. Competitivity Clusters

James Crowley is on the scientific committee for the Minalogic Competitivity Cluster. Minalogic is the global innovation cluster for digital technologies serving France's Auvergne-Rhône-Alpes region. The Scientific Committee advises the pole of strategy, advises local industry in proposal preparation, reviews FUI project proposals, and makes recommendations about labelling and support of project proposals.

9.2. European Initiatives

9.2.1. H2020 Project AI4EU - ICT-26-2018 Artificial Intelligence

From February 2018 to Sept 2018, James Crowley has participated in the core writing team for the H2020 proposal AI4EU submitted to the call ICT-26-2018 Artificial Intelligence. The project proposal was submitted in April 2018. The consortium has been notified in September 2018 that the project has been accepted for funding, and will begin on 1 January 2019.

AI4EU will bring together European researchers, educators, entrepreneurs and socio-economic innovators around a shared, crowd-sourced, innovation ecosystem that lowers barriers for education, research and innovation through AI. This ecosystem will be constructed by federating existing national innovation platforms and their user communities wherever possible, and by completing this federation with new components, new services and new enabling technologies that respond to opportunities for innovation.

9.2.2. H2020 FET Flagship Humane AI

James Crowley has participated as part of the core team for the proposal to create a FET Flagship named Humane AI. The Humane AI Flagship will develop the scientific and technological foundations needed to shape the AI revolution in a direction that is beneficial to humans on both individual and social level and strictly adheres to European ethical values and social norms. The core concept is that of AI systems that understand and adapt to complex dynamic environments and social settings in order enhance human capabilities and empower people as individuals and the society as whole.

Following a successful 1st stage proposal submitted in 2017, the consortium was invited to submit a 2nd stage proposal in Sept. 2018. We have been notified in November that this 2nd stage proposal has been accepted for funding. The project is start date is proposed for March 2019.

9.3. International Initiatives

9.3.1. Participation in Other International Programs

Vietnam

International partnership with **HUST** (Hanoi University of Science and Technology), Vietnam Joint lab unit between Grenoble INP and HUST, with the support of CNRS: **International Research Institute MICA** (Multimedia, Information, Communication and Applications) – UMI 2954 of CNRS from January 2006 to March 2018.

- Eric Castelli: French director of UMI 2954 “MICA Institute”, Vietnam, from 01 September 2001 to 5 February 2018
- Eric Castelli: now Adjunct Member of International Research Institute MICA, Vietnam (from June 2018)
- Eric Castelli: Responsible (and co-founder) of the International MASTER degree ACMI (Ambient Computing, Multimedia & Interactions), Hanoi University of Science and Technology (from January 2014 to June 2018)
- Eric Castelli: International scientific expert for the Vietnamese agency for research development NAFOSTED (National Foundation for Science and Technology Development), Ministry of Science and Technology, Vietnam (from 2015 to now)
- Eric Castelli: active participant to the bilateral French-Vietnam program PFIEV (Programme de Formation d’Ingénieurs d’Excellence au Vietnam), Grenoble INP is one of the main French partners.

Cambodia

International partnership with **ITC** (Institut de Technologie du Cambodge), Phnom Penh, Cambodia

- Eric Castelli: Member of the International Consortium of “Institut de Technologie du Cambodge (ITC)”, Phnom Penh, Cambodia, representative of Hanoi University of Science and Technology (from 2008 to March 2018)
- Eric Castelli: Elected Member, representative of the International Consortium at the Administration Council of the “Institut de Technologie du Cambodge” (ITC), Phnom Penh, Cambodia (from 2014 to March 2018)

International partnership with **NIPTICT** (National Institute of Post and Telecoms, and Information Communication Technologies), Phnom Penh, Cambodia. NIPTICT Institute is under the authority of the Ministry of Posts and Telecommunications of Cambodia

- Eric Castelli: Scientific advisor for the Ministry of Posts and Telecommunications of Cambodia, for the creation of the research center CSSD (Computer Sciences for Social Development, a new research lab of NIPTICT)
- Eric Castelli: cowriter of the MELISSA international project, submitted to French AFD Agency (with NIPTICT (leader), NUOL, and HUST partners) in 2018 (1st submission) and 2019 (2nd submission)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Sabine Coquillart was Conference co-chair for ICAT-EGVE 2018, Cyprus, Nov. 7-9, 2018.

10.1.1.2. Chair of Conference Program Committees

- Thierry Fraichard was Program Co-Chair for the IEEE Int. Conf. on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAR), Brisbane (AT), *May 2018*.

10.1.1.3. Member of Conference Program Committees

- Patrick Reignier was a member of the Program Committee of Smart Objects 2018 (satellite workshop to CHI 2018, Montreal (CA).
- Eric Castelli: co-organisator and "board member" of the international conference serie STLU’xx (Spoken Language Technologies for Under-resourced languages) since it-s creation. For 2018: participation of SLTU’18 (Gurugam, India) (as reminder: SLTU’16 (Jakarta, Indonesia) - SLTU’14 (St Petersburg, Russia) - SLTU’12 (Cape Town, South Africa) - STLU’10 (Penang, Malaysia) - Organisator and “co-chair” of the first edition STLU’08 (Hanoi, Vietnam))
- Sabine Coquillart was a member of the Program Committee for 3DCVE’18, CENTRIC’18, GRAPP’18, ICGI’18, ICMI’18, IEEE VR’18 Journal Papers, IEEE VR’18 Conference Papers, ICAT-EGVE’18, VRST’18, WSCG’18,

10.1.1.4. Reviewer

- Thierry Fraichard reviewed papers for the IEEE/RSJ IROS conference.
- Patrick Reignier reviewed papers for Smart Objects 2018.
- Dominique Vaufreydaz reviewed articles for IV2018, IUadapt2018, RO-MAN2018, ICARCV2018, AVEC2018, Workshop on Modeling Cognitive Processes from Multimodal Data at ICMI 2018, HRI 2019, MAPR2018.

- Eric Castelli reviewed papers of 14 international conferences and workshops: SPIN 2018 (Noida, Dehly, India) - 3ICT18 (Bahrain) – IALP 2018 (Bandung, Indonesia) – NICS 2018 (Ho Chi Minh City, Vietnam) - SCS 2018 (Bahrain) – SPECOM 2018 (Leipzig, Germany) – SigTelCom 2018 (Ho Chi Minh Citiy, Vietnam) - STLU'18 (Gurugram, India) – ATC 2018 (Ho Chi Minh City, Vietnam) – ICVES 2018 (Madrid, Spain) – SCS 2019 (Bahrain) – SigTelCom 2019 (Hanoi, Vietnam) – ICMSAO 2019 (Bahrain) – SPECOM 2019 (Istanbul, Turkey) – MAPR 2019 (Ho Chi Minh City, Vietnam)
- James Crowley Reviewed papers for ACCV 2018, ICMI 2019, and CVPR 2018.
- Sabine Coquillart Reviewed papers for Eurographics'18, Eurohaptics'18.

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

- Sabine Coquillart is a member of the Scientific Committee of the Journal of Virtual Reality and Broadcasting.
- Sabine Coquillart is Academic Editor for PeerJ Computer Science Journal.
- Sabine Coquillart is Review Editor for Frontiers in Virtual Environments Journal.
- Sabine Coquillart is Associate Editor for Presence: Teleoperators and Virtual Environments.
- Thierry Fraichard is an Asssociate Editor for IEEE Robotics and Automation Letters (RA-L).
- Patrick Reignier is a member of the editorial board of the Modeling and Using Context Journal.
- Patrick Reignier served as a guest editor for a special issue of the Revue d'Intelligence Artificielle on Smart Homes.
- Dominique Vaufreydaz served as a guest editor for a special issue Special Issue "Human Behavior, Emotion and Representation" of Multimodal Technologies and Interaction.

10.1.2.2. Reviewer - Reviewing Activities

- Thierry Fraichard reviewed articles for IEEE Trans. Intelligent Vehicles (TIV), IEEE Robotics and Automation Letters (RA-L).
- Thierry Fraichard reviewed proposals for the IDEX Université Grenoble Alpes and the Italian Research Agency.
- Thierry Fraichard served as a reviewer for the H2020 ILIAD European project.
- Eric Castelli reviewed article for the "Romanian Human Computer Interaction" Journal (RRIOC)
- Dominique Vaufreydaz reviewed articles for the journal of Interactive Technology and Smart Education

10.1.3. Invited Talks

- Eric Castelli: *The future of scientific international cooperation in ASEAN area*. Seminar of International Research Institute MICA, Hanoi, Vietnam, 9 & 10 July 2018
- James Crowley: *Put That There: 30 Years of Research on Multi-Modal Interaction*, Invited Plenary Presentation, 2018 International Conference on Multimodal Interaction, ICMI 2018, Boulder Co. 16 Oct 2018.
- James Crowley: *Artificial Intelligence*, Invited Plenary Presentation, 2018 Colloque annuel des directeurs des écoles d'ingénieures, Marseilles, 31 May 2018.
- Sabine Coquillart: *Pseudo-haptics : résultats de recherche et perspectives d'applications*, Invited presentation, Séminaire "Sensorimotricité, intersensorialité et réalité virtuelle", Grenoble, 28 janvier 2018.

10.1.4. Leadership within the Scientific Community

- James Crowley is a member of the Steering Committee for the ACM Int. Conf. on Multimodal Interaction.
- Sabine Coquillart is serving as member of the steering committee for the ICAT Int. Conf. on Artificial Reality and Telexistence.
- Sabine Coquillart is chairing the steering committee for the EGVE Working Group – EUROGRAPHICS Working group on Virtual Environments.
- Sabine Coquillart is an elected member of the EUROGRAPHICS Executive Committee.
- Sabine Coquillart is a member of the EUROGRAPHICS Working Group and Workshop board.

10.1.5. Scientific Expertise

- James L. Crowley served on the ANR committee CE 33 Interaction, Robotics.
- Sabine Coquillart served as an expert reviewer for the Austrian Science Fund.
- Thierry Fraichard served as an expert reviewer for the European Commission and the Italian Research Agency.
- Patrick Reignier is a member of the Scientific Council of the Amiqual4Home EquipEx.
- Dominique Vaufreydaz served as an evaluator for Initiatives de Recherche Stratégiques (IRS) call of the Grenoble Idex.
- Dominique Vaufreydaz served as an evaluator for Icelandic Research Fund.
- Eric Castelli served as an international expert for the Vietnamese agency NAFOSTED (National Foundation for Science and Technology Development), Ministry of Science and Technology of Vietnam.
- Eric Castelli served as an international expert for the Cambodian Ministry of Posts and Telecommunications
- Sabine Coquillart served as member for the Best Paper Award Committee for IEEE VR 2018.
- Sabine Coquillart served as a member of the Scientific Committee of "Challenges of IoT in the Digital Tools and Uses Congress", 2018.

10.1.6. Research Administration

- James Crowley is director for the Amiqual4Home Innovation Platform (EquipEx).
- James Crowley serves on the Administrative Committee (Bureau) for the Laboratoire Informatique de Grenoble.
- James Crowley served on the Scientific Committee (CoS) for the CRI Inria Grenoble-Rhone Alpes.
- James Crowley served on the orientation committee for the Competitivety Pole Minalogic
- Patrick Reignier is head of the engineering support group of the Laboratoire d'Informatique de Grenoble (13 members). He is currently supervising the moving of the Domus Living lab to a new site on Campus.
- Patrick Reignier serves on the Administrative Office (Bureau) for the Laboratoire Informatique de Grenoble.
- Patrick Reignier is at the head of the Domus Living Lab
- Patrick Reignier is a member of the Comité Executif of the Amiqual4Home Equipex
- Patrick Reignier is a member of the Comité de pilotage of the MACI (Maison de la Création et l'Innovation)

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

10.2.1.1. James Crowley

James Crowley was co-director of the Master of Science in Informatics at Grenoble (MoSIG) to Sept 2018.

Master : Computer Vision, Course 27h EqTD, M2 year, Master of Science in Informatics at Grenoble

Master 1: Intelligent Systems, Cours 54h EqTD, UFRIM2AG

ENSIMAG 2: Intelligent Systems, Cours 54h EqTD, ENSIMAG

ENSIMAG 3 : Pattern Recognition and Machine Learning, Cours 27h EqTD, ENSIMAG

10.2.1.2. Patrick Reignier

- Patrick Reignier has been elected member of the Conseil des Etudes et de la Vie Universitaire of Grenoble INP
- Patrick Reignier was a member of the consultation group for the proposal of the founding text for the integrated university.
- Patrick Reignier has been nominated as a member of the Conseil de la Formation Continue de Grenoble INP
- Patrick Reignier participated in the editing of a successful Idex Educational program proposal "FromLivingLab".
- Patrick Reignier is co-director of the "formation en apprentissage" of Ensimag (3 years program : 1 year for the Licence and 2 years for the Master)
- Patrick Reignier Supervises the industrial part of the “formation en apprentissage” of the Ensimag engineering school.
- Master: Patrick Reignier, Projet Genie Logiciel, 55h eqTD, M1, Ensimag/Grenoble INP, France.
- Master: Patrick Reignier, Développement d’applications communicantes, 18h eqTD, M2, Ensimag/Grenoble-INP, France
- Master: Patrick Reignier, Introduction aux applications reparties, 18h eqTD, M2, Ensimag/Grenoble-INP, France
- Master: Patrick Reignier, Applications Web et Mobiles , 27h eqTD, M1, Ensimag/Grenoble-INP, France
- Master: Patrick Reignier, Projet Systeme, 12h eq TD, M1, Ensimag/Grenoble-INP, France
- Licence: Patrick Reignier, Projet C, 20h eqTD, L3, Ensimag/Grenoble-INP, France.

10.2.1.3. Dominique Vaufreydaz

- Co-responsibility of the Graphic, Vision and Robotics track of the MOSIG Master program.
- In charge of the transversal numerical competence courses at the Grenoble Faculty of Economics (Grenoble and Valence campuses)
- Licence: Compétences Numériques, 126h eq TD,L1, Université Grenoble Alpes, France.
- Licence: Informatique appliquée à l’économie et à la gestion, 32 h eq TD, enseignement à distance, Licence, Université Grenoble Alpes, France.
- Licence: Pratique avancée du Tableur, 72 h eq TD, L3, Université Grenoble Alpes, France.
- Licence Professionnelle: Enquêtes et traitement d’enquêtes avec le logiciel Sphinx, 12h eq TD, Licence pro Métiers de l’Emploi et de la Formation, Université Grenoble Alpes, France.
- Licence Professionnelle: Administration en environnement hétérogène, 20h eq TD, Licence pro Administration et Sécurité des Systèmes et des Réseaux, Université Grenoble Alpes, France.
- IUT année spéciale: Programmation C++, 18h eq TD, Année Spéciale IUT Informatique, Université Grenoble Alpes, France.
- Master: Pratique avancée du Tableur, 24 h eq TD, M1 économie internationale et stratégies d’acteurs, Université Grenoble Alpes, France.

- Master: Développement Web Mobile, 24h eq TD, M2 Mathématiques et Informatique Appliquées aux Sciences Humaines et Sociales, Université Grenoble Alpes, France

10.2.1.4. Thierry Fraichard

- Master: Thierry Fraichard, Autonomous Robotics, 22.5h eqTD, M2 MOSIG, Univ. Grenoble Alpes.

10.2.2. Supervision - Doctoral theses completed in 2018

- Rémi Paulin, *Human-Robot Motion: an Attention-Based Approach*, 22 March 2018, Patrick Reignier and Thierry Fraichard.
- Nguyen Dinh Van, *Wireless Sensors Networks for Indoor Mapping and Accurate Localization for Low Speed Navigation in Smart Cities*, 5 December 2018, Fawzi Nashashibi and Eric Castelli

10.2.3. Supervision - Current Doctoral Students

- PhD in progress: Jose Grimaldo Da Silva Filho, "Human-Robot Motion, a Shared Effort Approach", Octobre 2015, Thierry Fraichard.
- PhD in progress: Matteo Ciocca, "Safe Robot Motion", Octobre 2016, Thierry Fraichard and Pierre-Brice Wieber.
- PhD in progress : Romain Bregier, Détection et estimation de pose d'instances d'objet rigide pour la manipulation robotisée, Octobre 2014, Frederic Devernay and James Crowley
- PhD in progress : Etienne Balit, expressive social robot design, Octobre 2014, Patrick Reignier and Dominique Vaufreydaz.
- PhD in progress : Interaction-Aware Tracking and Lane Change Detection in Highway Driving, David Sierra Gonzales, Octobre 2014, Dizan Vasquez and Emmanuel Mazer
- PhD in progress : Minh-khoa Nguyen, Robotics-inspired methods for modeling and simulation of large nanosystems, Octobre 2014, Leaonard Jaillet and Emmanuel Mazer
- PhD in progress : Amr Alyafi, explanatory observation of energy usage, Octobre 2015, Stephane Ploix and Patrick Reignier.
- PhD in progress : Amine Awada, activity recognition for smart energy management, Octobre 2015, Stephane Ploix and Patrick Reignier.
- PhD in progress : Liliya Tsetanova, Social Robotics, Octobre 2015, Veronique Aubergé and Patrick Reignier.
- PhD in progress : Thomas Guntz, Multimodal Observation of Subjects Engaged in Problem Solving, Octobre 2016, James Crowley and Dominique Vaufreydaz
- PhD in progress : Pavan Vashista, Situational Awareness of Autonomous Cars in Urban Areas, February 2016, Dominique Vaufreydaz and Anne Spalanzani (Inria Chroma).
- PhD in progress : Nashwa Abou Bakr, Observation of Kitchen Activities, Octobre 2016, James Crowley and Remi Ronfard
- PhD in progress : Raphael Frisch, conception de machine stochastique pour la localisation et la séparation de source sonor, Octobre 2016, Emmanuel Mazer

10.2.4. Juries

- Thierry Fraichard served as an expert evaluator in the PhD Jury of Florent Altché, Ecole Nat. Sup. des Mines de Paris, *Sep. 18*.
- Patrick Reignier served as a jury member of the Doctoral Jury of Hiary Landy Rajaonarivo (Lab-STICC, Brest)
- Patrick Reignier served as a president of the Doctoral Jury of Nathan Ramoly (Telecom Sud Paris)
- Patrick Reignier served as a president of the Doctoral Jury of Jérémy WAMBECKE (University Grenoble Alps)

- Patrick Reignier served as a president of the Doctoral Jury of Paola Gomez (University Grenoble Alps)
- Patrick Reignier served as a reporter of the Doctoral Jury of Mauricio Gomez Morales (Purdue University and Paris Est Créteil)
- Sabine Coquillart served as reporter for the doctoral jury of A. Costes at INSA Rennes
- Sabine Coquillart served as examiner for the doctoral juty of G. Cortes at Univ. Rennes 1.

11. Bibliography

Major publications by the team in recent years

[1] R. Brégier, F. Devernay, L. Leyrit, J. Crowley.*Defining the Pose of any 3D Rigid Object and an Associated Distance*, in "International Journal of Computer Vision", June 2018, vol. 126, n^o 6, p. 571–596, https://arxiv.org/abs/1612.04631 [*DOI :* 10.1007/s11263-017-1052-4], https://hal.inria.fr/hal-01415027

[2] J. Coutaz, J. L. Crowley.*A First-Person Experience with End-User Development for Smart Homes*, in "IEEE Pervasive Computing", May 2016, vol. 15, p. 26 - 39 [*DOI :* 10.1109/MPRV.2016.24], https://hal.inria.fr/hal-01422364

[3] T. Fraichard, R. Paulin, P. Reignier.*Human-Robot Motion: An Attention-Based Navigation Approach*, in "IEEE RO-MAN", Edinburgh (UK), August 2014, Best Paper Award Nominee [*DOI :* 10.1109/ROMAN.2014.6926332], http://hal.inria.fr/hal-01018471

[4] T. Guntz, J. L. Crowley, D. Vaufreydaz, R. Balzarini, P. Dessus.*The Role of Emotion in Problem Solving: First Results from Observing Chess*, in "ICMI 2018 - Workshop at 20th ACM International Conference on Multimodal Interaction", Boulder, Colorado, United States, October 2018, p. 1-13, https://arxiv.org/abs/1810.11094 , https://hal.inria.fr/hal-01886694

[5] P. Vasishta, D. Vaufreydaz, A. Spalanzani.*Building Prior Knowledge: A Markov Based Pedestrian Prediction Model Using Urban Environmental Data*, in "ICARCV 2018 - 15th International Conference on Control, Automation, Robotics and Vision", Singapore, Singapore, November 2018, p. 1-12, Best Student Paper Award, https://hal.inria.fr/hal-01875147

Publications of the year

Doctoral Dissertations and Habilitation Theses

[6] R. Paulin.*Human-Robot Motion: an Attention-Based Approach*, Université Grenoble Alpes, March 2018, https://tel.archives-ouvertes.fr/tel-01864355

[7] D. Vaufreydaz.*Multimodal perception and sociable interaction*, Université Grenoble Alpes (France) ; MSTII, July 2018, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01970420

Articles in International Peer-Reviewed Journal

[8] A. Alzouhri Alyafi, V.-B. Nguyen, Y. Laurillau, P. Reignier, S. Ploix, G. Calvary, J. Coutaz, M. Pal, J.-P. Guilbaud.*From Usable to Incentive Building Energy Management Systems*, in "Modélisation et utilisation du contexte (Modeling and Using Context)", November 2018, vol. 2, n^o 1, p. 1-30 [*DOI :* 10.21494/ISTE.OP.2018.0302], https://hal.inria.fr/hal-01929051

[9] M. Bouguerra, T. Fraichard, M. Fezari. *Viability-Based Guaranteed Safe Robot Navigation*, in "Journal of Intelligent and Robotic Systems", November 2018, p. 1–13 [*DOI :* 10.1007/s10846-018-0955-9], https://hal.inria.fr/hal-01924855

[10] R. Brégier, F. Devernay, L. Leyrit, J. Crowley. *Defining the Pose of any 3D Rigid Object and an Associated Distance*, in "International Journal of Computer Vision", June 2018, vol. 126, n^{o} 6, p. 571–596, https://arxiv.org/abs/1612.04631 [*DOI :* 10.1007/s11263-017-1052-4], https://hal.inria.fr/hal-01415027

[11] J. L. Coutaz, J. L. Crowley. *Appsgate, a Programmable Domestic Eco-system: Learning from "Living in it"*, in "Journal d'Interaction Personne-Système (JIPS)", October 2018, vol. Volume 7, Number 1, n^{o} 1, p. 1-35, https://hal.archives-ouvertes.fr/hal-01898173

[12] T. Guntz, R. Balzarini, D. Vaufreydaz, J. L. Crowley. *Multimodal Observation and Classification of People Engaged in Problem Solving: Application to Chess Players*, in "Multimodal Technologies and Interaction", March 2018, vol. 2, n^{o} 2 [*DOI :* 10.3390/mti2020011], https://hal.inria.fr/hal-01886354

[13] G. Nieto, F. Devernay, J. L. Crowley. *Rendu basé image avec contraintes sur les gradients*, in "Traitement du Signal", 2018, p. 1-26 [*DOI :* 10.3166/HSP.x.1-26], https://hal.archives-ouvertes.fr/hal-01900200

[14] M. Pal, A. Alzouhri Alyafi, S. Ploix, P. Reignier, S. Bandyopadhyay. *Unmasking the causal relationships latent in the interplay between occupant's actions and indoor ambience: a building energy management outlook*, in "Applied Energy", January 2019, https://hal.archives-ouvertes.fr/hal-01984833

Invited Conferences

[15] S. Nabil, R. Balzarini, F. Devernay, J. L. Crowley. *Designing objective quality metrics for panoramic videos based on human perception*, in "IMVIP 2018 - Irish Machine Vision and Image Processing Conference", Belfast, United Kingdom, August 2018, p. 1-4, https://hal.inria.fr/hal-01849043

International Conferences with Proceedings

[16] E. Balit, D. Vaufreydaz, P. Reignier. *PEAR: Prototyping Expressive Animated Robots - A framework for social robot prototyping*, in "HUCAPP 2018 - 2nd International Conference on Human Computer Interaction Theory and Applications", Funchal, Madeira, Portugal, January 2018, 1, https://hal.inria.fr/hal-01698493

[17] J. G. Da Silva Filho, A.-H. Olivier, A. Crétual, J. Pettré, T. Fraichard. *Human inspired effort distribution during collision avoidance in human-robot motion*, in "RO-MAN 2018 - 27th IEEE International Symposium on Robot and Human Interactive Communication", Nanjing and Tai'an, China, Springer, August 2018, p. 1111-1117 [*DOI :* 10.1109/ROMAN.2018.8525623], https://hal.inria.fr/hal-01831757

[18] E. Ferrier-Barbut, D. Vaufreydaz, J.-A. David, J. Lussereau, A. Spalanzani. *Personal space of autonomous car's passengers sitting in the driver's seat*, in "IV'2018 - The 29th IEEE Intelligent Vehicles Symposium", Changshu, Suzhou, China, IEEE, June 2018, p. 2022-2029, https://arxiv.org/abs/1805.03563 [*DOI :* 10.1109/IVS.2018.8500648], https://hal.inria.fr/hal-01786006

[19] T. Guntz, J. L. Crowley, D. Vaufreydaz, R. Balzarini, P. Dessus. *The Role of Emotion in Problem Solving: First Results from Observing Chess*, in "ICMI 2018 - Workshop at 20th ACM International Conference on Multimodal Interaction", Boulder, Colorado, United States, October 2018, p. 1-13, https://arxiv.org/abs/1810.11094 , https://hal.inria.fr/hal-01886694

[20] D.-V. N. NGUYEN, R. DE CHARETTE, T.-K. DAO, E. CASTELLI, F. NASHASHIBI.*WiFi Fingerprinting Localization for Intelligent Vehicles in Car Park*, in "IPIN 2018 : Ninth International Conference on Indoor Positioning and Indoor Navigation", Nantes, France, September 2018, https://hal.inria.fr/hal-01851504

[21] V.-C. TA, T.-K. DAO, D. VAUFREYDAZ, E. CASTELLI.*Smartphone-based user positioning in a multiple-user context with Wi-Fi and Bluetooth*, in "IPIN 2018 - 9th International Conference on Indoor Positioning and Indoor Navigation", Nantes, France, September 2018, p. 1-13, https://arxiv.org/abs/1807.05716 , https://hal.inria.fr/hal-01839574

[22] P. VASISHTA, D. VAUFREYDAZ, A. SPALANZANI.*Building Prior Knowledge: A Markov Based Pedestrian Prediction Model Using Urban Environmental Data*, in "ICARCV 2018 - 15th International Conference on Control, Automation, Robotics and Vision", Singapore, Singapore, November 2018, p. 1-12, https://arxiv.org/abs/1809.06045 , https://hal.inria.fr/hal-01875147

Conferences without Proceedings

[23] J. CUMIN, G. LEFEBVRE, F. RAMPARANY, J. L. CROWLEY.*Inferring Availability for Communication in Smart Homes Using Context*, in "PerCom 2018 - IEEE International Conference on Pervasive Computing and Communications", Athènes, Greece, March 2018, p. 1-6, https://hal.archives-ouvertes.fr/hal-01762137

[24] P. DESSUS, L.-H. AUBINEAU, D. VAUFREYDAZ, J. L. CROWLEY.*A Framework for a Multimodal Analysis of Teaching Centered on Shared Attention and Knowledge Access*, in "Grenoble Workshop on Models and Analysis of Eye Movements", Grenoble, France, June 2018, https://hal.archives-ouvertes.fr/hal-01811092

[25] R. FRISCH, M. FAIX, E. MAZER, L. FESQUET, A. LUX.*A cognitive stochastic machine based on Bayesian inference: a behavioral analysis*, in "ICCI*CC 2018 - 17th IEEE International Conference on Cognitive Informatics and Cognitive Computing", Berkeley, United States, July 2018, p. 1-8, https://hal.archives-ouvertes.fr/hal-01867789

[26] R. PAULIN, T. FRAICHARD, P. REIGNIER.*Human-Robot Motion: Taking Human Attention into Account*, in "IROS 2018- IEEE/RSJ International Conference on Intelligent Robots and Systems; Workshop on Assistance and Service Robotics in a Human Environment", Madrid, Spain, October 2018, p. 1-5, https://hal.inria.fr/hal-01856016

Scientific Books (or Scientific Book chapters)

[27] N. DÍAZ-RODRÍGUEZ, S. GRÖNROOS, F. WICKSTRÖM, J. LILIUS, H. EERTINK, A. BRAUN, P. DILLEN, J. CROWLEY, J. ALEXANDERSSON.*An Ontology for Wearables Data Interoperability and Ambient Assisted Living Application Development*, in "Recent Developments and the New Direction in Soft-Computing Foundations and Applications", L. A. ZADEH, R. R. YAGE, S. N. SHAHBAZOVA, M. Z. REFORMA, V. KREINOVICH (editors), Studies in Fuzziness and Soft Computing, Springer, 2018, vol. 361, p. 559-568, Selected Paper from the 6th World Conference on Soft Computing, May 22-25, 2016, Berkeley, USA [*DOI :* 10.1007/978-3-319-75408-6_43], https://hal.archives-ouvertes.fr/hal-01951531

[28] M. PAL, A. ALZOUHRI ALYAFI, S. BANDYOPADHYAY, S. PLOIX, P. REIGNIER.*Enhancing Comfort of Occupants in Energy Buildings*, in "Operations Research and Optimization", April 2018, p. 1-11, https://hal.archives-ouvertes.fr/hal-01865292

Research Reports

[29] M. Bouguerra, T. Fraichard, M. Fezari.*Viability-Based Guaranteed Safe Robot Navigation*, Inria, October 2018, n^{o} RR-9217, https://hal.inria.fr/hal-01901675

Other Publications

[30] S. Nabil, F. Devernay, J. L. Crowley.*An objective quality metric for panoramic videos*, July 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01849261

References in notes

[31] C. Eriksen, D. James.*Visual attention within and around the field of focal attention: A zoom lens model*, in "Perception & psychophysics", 1986, vol. 40, n^{o} 4

[32] T. Fraichard, R. Paulin, P. Reignier.*Human-Robot Motion: An Attention-Based Navigation Approach*, in "IEEE Robot Hum. Interact. Commun. (RO-MAN)", Edinburgh (GB), August 2014

[33] F. Lindner, C. Eschenbach.*Towards a formalization of social spaces for socially aware robots*, in "Spatial Information Theory", 2011

[34] J. Maisonnasse, N. Gourier, O. Brdiczka, P. Reignier.*Attentional model for perceiving social context in intelligent environments*, in "IFIP Int. Conf. on Artificial Intelligence Applications and Innovations", 2006

[35] B. J. Scholl.*Objects and attention: The state of the art*, in "Cognition", 2001, vol. 80, n^{o} 1

[36] D. Shi, E. Collins, B. Goldiez, A. Donate, X. Liu, D. Dunlap.*Human-aware robot motion planning with velocity constraints*, in "Int. Symp. on Collaborative Technologies and Systems", Irvine (US), May 2008

[37] E. Sisbot, K. Marin-Urias, R. Alami, T. Siméon.*A human aware mobile robot motion planner*, in "IEEE Trans. Robot.", 2007, vol. 23, n^{o} 5

Project-Team POLARIS

Performance analysis and Optimization of LARge Infrastructures and Systems

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

IN PARTNERSHIP WITH:
CNRS

Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Distributed and High Performance Computing

Table of contents

Project-Team POLARIS

Creation of the Team: 2016 January 01, updated into Project-Team: 2018 January 01

Keywords:

Computer Science and Digital Science:

A1.1.1. - Multicore, Manycore
A1.1.2. - Hardware accelerators (GPGPU, FPGA, etc.)
A1.1.4. - High performance computing
A1.1.5. - Exascale
A1.2. - Networks
A1.2.3. - Routing
A1.2.5. - Internet of things
A1.6. - Green Computing
A3.4. - Machine learning and statistics
A3.5.2. - Recommendation systems
A5.2. - Data visualization
A6. - Modeling, simulation and control
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.2.6. - Optimization
A6.2.7. - High performance computing
A8.2. - Optimization
A8.9. - Performance evaluation
A8.11. - Game Theory

Other Research Topics and Application Domains:

B4.4. - Energy delivery
B4.4.1. - Smart grids
B4.5.1. - Green computing
B6.2. - Network technologies
B6.2.1. - Wired technologies
B6.2.2. - Radio technology
B6.4. - Internet of things
B8.3. - Urbanism and urban planning
B9.6.7. - Geography
B9.7.2. - Open data
B9.8. - Reproducibility

1. Team, Visitors, External Collaborators

Research Scientists

Arnaud Legrand [Team leader, CNRS, Senior Researcher, HDR]
Nicolas Gast [Inria, Researcher]

Bruno Gaujal [Inria, Senior Researcher, HDR]
Panayotis Mertikopoulos [CNRS, Researcher]
Patrick Loiseau [Univ Grenoble Alpes then Inria (from Oct 2018), Researcher, HDR]
Bary Pradelski [CNRS, Researcher, from Oct 2018]

Faculty Members

Vincent Danjean [Univ Grenoble Alpes, Associate Professor]
Guillaume Huard [Univ Grenoble Alpes, Associate Professor]
Florence Perronnin [Univ Grenoble Alpes, Associate Professor]
Jean-Marc Vincent [Univ Grenoble Alpes, Associate Professor]
Philippe Waille [Univ Grenoble Alpes, Associate Professor]
Elena-Veronica Belmega [Ecole nationale supérieure de l'électronique et de ses applications, Associate Professor]

PhD Students

Kimon Antonakopoulos [Inria]
Tom Cornebize [Univ Grenoble Alpes]
Bruno de Moura Donassolo [Orange]
Stephane Durand [Univ Grenoble Alpes]
Vitalii Emelianov [Inria, from Sep 2018]
Vinicius Garcia Pinto [Université fédérale du Rio Grande do Sul, until Oct 2018]
Franz Christian Heinrich [Inria]
Alexis Janon [Univ Grenoble Alpes, from Oct 2018]
Baptiste Jonglez [Univ Grenoble Alpes]
Alexandre Marcastel [Univ de Cergy Pontoise, until Sep 2018]
Stephan Plassart [Univ Grenoble Alpes]
Pedro Rocha Bruel [Univ Sao Paolo]
Benjamin Roussillon [Univ Grenoble Alpes, from Oct 2018]
Benoît Vinot [Schneider Electric, until Apr 2018]
Dong Quan Vu [Nokia]

Post-Doctoral Fellows

George Arvanitakis [Univ Grenoble Alpes, until Oct 2018]
Olivier Bilenne [CNRS, from Sep 2018]
Amélie Hélious [Floralis, until Apr 2018]
Takai Eddine Kennouche [Univ Grenoble Alpes]
Mouhcine Mendil [INP SA]

Administrative Assistant

Annie Simon [Inria]

2. Overall Objectives

2.1. Context

Large distributed infrastructures are rampant in our society. Numerical simulations form the basis of computational sciences and high performance computing infrastructures have become scientific instruments with similar roles as those of test tubes or telescopes. Cloud infrastructures are used by companies in such an intense way that even the shortest outage quickly incurs the loss of several millions of dollars. But every citizen also relies on (and interacts with) such infrastructures via complex wireless mobile embedded devices whose nature is constantly evolving. In this way, the advent of digital miniaturization and interconnection has enabled our homes, power stations, cars and bikes to evolve into smart grids and smart transportation systems that should be optimized to fulfill societal expectations.

Our dependence and intense usage of such gigantic systems obviously leads to very high expectations in terms of performance. Indeed, we strive for low-cost and energy-efficient systems that seamlessly adapt to changing environments that can only be accessed through uncertain measurements. Such digital systems also have to take into account both the users' profile and expectations to efficiently and fairly share resources in an online way. Analyzing, designing and provisioning such systems has thus become a real challenge.

Such systems are characterized by their **ever-growing size**, intrinsic **heterogeneity** and **distributedness**, **user-driven** requirements, and an unpredictable variability that renders them essentially **stochastic**. In such contexts, many of the former design and analysis hypotheses (homogeneity, limited hierarchy, omniscient view, optimization carried out by a single entity, open-loop optimization, user outside of the picture) have become obsolete, which calls for radically new approaches. Properly studying such systems requires a drastic rethinking of fundamental aspects regarding the system's **observation** (measure, trace, methodology, design of experiments), **analysis** (modeling, simulation, trace analysis and visualization), and **optimization** (distributed, online, stochastic).

2.2. Objectives

The goal of the POLARIS project is to **contribute to the understanding of the performance of very large scale distributed systems** by applying ideas from diverse research fields and application domains. We believe that studying all these different aspects at once without restricting to specific systems is the key to push forward our understanding of such challenges and to proposing innovative solutions. This is why we intend to investigate problems arising from application domains as varied as large computing systems, wireless networks, smart grids and transportation systems.

The members of the POLARIS project cover a very wide spectrum of expertise in performance evaluation and models, distributed optimization, and analysis of HPC middleware. Specifically, POLARIS' members have worked extensively on:

Experiment design: Experimental methodology, measuring/monitoring/tracing tools, experiment control, design of experiments, and reproducible research, especially in the context of large computing infrastructures (such as computing grids, HPC, volunteer computing and embedded systems).

Trace Analysis: Parallel application visualization (paje, triva/viva, framesoc/ocelotl, ...), characterization of failures in large distributed systems, visualization and analysis for geographical information systems, spatio-temporal analysis of media events in RSS flows from newspapers, and others.

Modeling and Simulation: Emulation, discrete event simulation, perfect sampling, Markov chains, Monte Carlo methods, and others.

Optimization: Stochastic approximation, mean field limits, game theory, discrete and continuous optimization, learning and information theory.

In the rest of this document, we describe in detail our new results in the above areas.

3. Research Program

3.1. Sound and Reproducible Experimental Methodology

Participants: Vincent Danjean, Nicolas Gast, Guillaume Huard, Arnaud Legrand, Patrick Loiseau, Jean-Marc Vincent.

Experiments in large scale distributed systems are costly, difficult to control and therefore difficult to reproduce. Although many of these digital systems have been built by men, they have reached such a complexity level that we are no longer able to study them like artificial systems and have to deal with the same kind of experimental issues as natural sciences. The development of a sound experimental methodology for the evaluation of resource management solutions is among the most important ways to cope with the growing complexity of computing environments. Although computing environments come with their own specific challenges, we believe such general observation problems should be addressed by borrowing good practices and techniques developed in many other domains of science.

This research theme builds on a transverse activity on *Open science and reproducible research* and is organized into the following two directions: (1) *Experimental design* (2) *Smart monitoring and tracing*. As we will explain in more detail hereafter, these transverse activity and research directions span several research areas and our goal within the POLARIS project is foremost to transfer original ideas from other domains of science to the distributed and high performance computing community.

3.2. Multi-Scale Analysis and Visualization

Participants: Vincent Danjean, Guillaume Huard, Arnaud Legrand, Jean-Marc Vincent, Panayotis Mertikopoulos.

As explained in the previous section, the first difficulty encountered when modeling large scale computer systems is to observe these systems and extract information on the behavior of both the architecture, the middleware, the applications, and the users. The second difficulty is to *visualize* and *analyze* such *multi-level traces to understand how the performance* of the application *can be improved*. While a lot of efforts are put into visualizing scientific data, in comparison little effort have gone into to developing techniques specifically tailored for understanding the behavior of distributed systems. Many visualization tools have been developed by renowned HPC groups since decades (e.g., BSC [87], Jülich and TU Dresden [86], [57], UIUC [75], [90], [78] and ANL [103], Inria Bordeaux [63] and Grenoble [105], ...) but most of these tools build on the classical information visualization mantra [95] that consists in always first presenting an overview of the data, possibly by plotting everything if computing power allows, and then to allow users to zoom and filter, providing details on demand. However in our context, the amount of data comprised in such traces is several orders of magnitude larger than the number of pixels on a screen and displaying even a small fraction of the trace leads to harmful visualization artifacts [82]. Such traces are typically made of events that occur at very different time and space scales, which unfortunately hinders classical approaches. Such visualization tools have focused on easing interaction and navigation in the trace (through gantcharts, intuitive filters, pie charts and kiviats) but they are very difficult to maintain and evolve and they require some significant experience to identify performance bottlenecks.

Therefore many groups have more recently proposed in combination to these tools some techniques to help identifying the structure of the application or regions (applicative, spatial or temporal) of interest. For example, researchers from the SDSC [85] propose some segment matching techniques based on clustering (Euclidean or Manhattan distance) of start and end dates of the segments that enables to reduce the amount of information to display. Researchers from the BSC use clustering, linear regression and Kriging techniques [94], [81], [74] to identify and characterize (in term of performance and resource usage) application phases and present aggregated representations of the trace [93]. Researchers from Jülich and TU Darmstadt have proposed techniques to identify specific communication patterns that incur wait states [100], [50]

3.3. Fast and Faithful Performance Prediction of Very Large Systems

Participants: Vincent Danjean, Bruno Gaujal, Arnaud Legrand, Florence Perronnin, Jean-Marc Vincent.

Evaluating the scalability, robustness, energy consumption and performance of large infrastructures such as exascale platforms and clouds raises severe methodological challenges. The complexity of such platforms mandates empirical evaluation but direct experimentation via an application deployment on a real-world testbed is often limited by the few platforms available at hand and is even sometimes impossible (cost, access, early stages of the infrastructure design, ...). Unlike direct experimentation via an application deployment on a real-world testbed, simulation enables fully repeatable and configurable experiments that can often be conducted quickly for arbitrary hypothetical scenarios. In spite of these promises, current simulation practice is often not conducive to obtaining scientifically sound results. To date, most simulation results in the parallel and distributed computing literature are obtained with simulators that are ad hoc, unavailable, undocumented, and/or no longer maintained. For instance, Naicken et al. [49] point out that out of 125 recent papers they surveyed that study peer-to-peer systems, 52% use simulation and mention a simulator, but 72% of them use a custom simulator. As a result, most published simulation results build on throw-away (short-lived and non validated) simulators that are specifically designed for a particular study, which prevents other researchers from building upon it. There is thus a strong need for recognized simulation frameworks by which simulation results can be reproduced, further analyzed and improved.

The *SimGrid* simulation toolkit [61], whose development is partially supported by POLARIS, is specifically designed for studying large scale distributed computing systems. It has already been successfully used for simulation of grid, volunteer computing, HPC, cloud infrastructures and we have constantly invested on the software quality, the scalability [53] and the validity of the underlying network models [51], [98]. Many simulators of MPI applications have been developed by renowned HPC groups (e.g., at SDSC [96], BSC [47], UIUC [104], Sandia Nat. Lab. [99], ORNL [60] or ETH Zürich [76] for the most prominent ones). Yet, to scale most of them build on restrictive network and application modeling assumptions that make them difficult to extend to more complex architectures and to applications that do not solely build on the MPI API. Furthermore, simplistic modeling assumptions generally prevent to faithfully predict execution times, which limits the use of simulation to indication of gross trends at best. Our goal is to improve the quality of SimGrid to the point where it can be used effectively on a daily basis by practitioners to *reproduce the dynamic of real HPC systems*.

We also develop another simulation software, *PSI* (Perfect SImulator) [65], [58], dedicated to the simulation of very large systems that can be modeled as Markov chains. PSI provides a set of simulation kernels for Markov chains specified by events. It allows one to sample stationary distributions through the Perfect Sampling method (pioneered by Propp and Wilson [88]) or simply to generate trajectories with a forward Monte-Carlo simulation leveraging time parallel simulation (pioneered by Fujimoto [69], Lin and Lazowska [80]). One of the strength of the PSI framework is its expressiveness that allows us to easily study networks with finite and infinite capacity queues [59]. Although PSI already allows to simulate very large and complex systems, our main objective is to push its scalability even further and *improve its capabilities by one or several orders of magnitude*.

3.4. Local Interactions and Transient Analysis in Adaptive Dynamic Systems

Participants: Nicolas Gast, Bruno Gaujal, Florence Perronnin, Jean-Marc Vincent, Panayotis Mertikopoulos.

Many systems can be effectively described by stochastic population models. These systems are composed of a set of n entities interacting together and the resulting stochastic process can be seen as a continuous-time Markov chain with a finite state space. Many numerical techniques exist to study the behavior of Markov chains, to solve stochastic optimal control problems [89] or to perform model-checking [48]. These techniques, however, are limited in their applicability, as they suffer from the *curse of dimensionality*: the state-space grows exponentially with n.

This results in the need for approximation techniques. Mean field analysis offers a viable, and often very accurate, solution for large n. The basic idea of the mean field approximation is to count the number of entities that are in a given state. Hence, the fluctuations due to stochasticity become negligible as the number of entities grows. For large n, the system becomes essentially deterministic. This approximation has been originally developed in statistical mechanics for vary large systems composed of more than 10^{20} particles (called entities here). More recently, it has been claimed that, under some conditions, this approximation can be successfully used for stochastic systems composed of a few tens of entities. The claim is supported by various convergence results [70], [79], [102], and has been successfully applied in various domains: wireless networks [52], computer-based systems [73], [84], [97], epidemic or rumour propagation [62], [77] and bike-sharing systems [66]. It is also used to develop distributed control strategies [101], [83] or to construct approximate solutions of stochastic model checking problems [54], [55], [56].

Within the POLARIS project, we will continue developing both the theory behind these approximation techniques and their applications. Typically, these techniques require a homogeneous population of objects where the dynamics of the entities depend only on their state (the state space of each object must not scale with n the number of objects) but neither on their identity nor on their spatial location. Continuing our work in [70], we would like to be able to handle heterogeneous or uncertain dynamics. Typical applications are caching mechanisms [73] or bike-sharing systems [67]. A second point of interest is the use of mean field or large deviation asymptotics to compute the time between two regimes [92] or to reach an equilibrium state. Last, mean-field methods are mostly descriptive and are used to analyse the performance of a given system. We wish to extend their use to solve optimal control problems. In particular, we would like to implement numerical algorithms that use the framework that we developed in [71] to build distributed control algorithms [64] and optimal pricing mechanisms [72].

3.5. Distributed Learning in Games and Online Optimization

Participants: Nicolas Gast, Bruno Gaujal, Arnaud Legrand, Patrick Loiseau, Panayotis Mertikopoulos.

Game theory is a thriving interdisciplinary field that studies the interactions between competing optimizing agents, be they humans, firms, bacteria, or computers. As such, game-theoretic models have met with remarkable success when applied to complex systems consisting of interdependent components with vastly different (and often conflicting) objectives – ranging from latency minimization in packet-switched networks to throughput maximization and power control in mobile wireless networks.

In the context of large-scale, decentralized systems (the core focus of the POLARIS project), it is more relevant to take an inductive, "bottom-up" approach to game theory, because the components of a large system cannot be assumed to perform the numerical calculations required to solve a very-large-scale optimization problem. In view of this, POLARIS' overarching objective in this area is to *develop novel algorithmic frameworks that offer robust performance guarantees when employed by all interacting decision-makers.*

A key challenge here is that most of the literature on learning in games has focused on *static* games with a *finite number of actions* per player [68], [91]. While relatively tractable, such games are ill-suited to practical applications where players pick an action from a continuous space or when their payoff functions evolve over time – this being typically the case in our target applications (e.g., routing in packet-switched networks or energy-efficient throughput maximization in wireless). On the other hand, the framework of online convex optimization typically provides worst-case performance bounds on the learner's *regret* that the agents can attain irrespectively of how their environment varies over time. However, if the agents' environment is determined chiefly by their interactions these bounds are fairly loose, so more sophisticated convergence criteria should be applied.

From an algorithmic standpoint, a further challenge occurs when players can only observe their own payoffs (or a perturbed version thereof). In this bandit-like setting regret-matching or trial-and-error procedures guarantee convergence to an equilibrium in a weak sense in certain classes of games. However, these results apply exclusively to static, finite games: learning in games with continuous action spaces and/or nonlinear payoff functions cannot be studied within this framework. Furthermore, even in the case of finite games, the complexity of the algorithms described above is not known, so it is impossible to decide a priori which algorithmic scheme can be applied to which application.

4. Application Domains

4.1. Large Computing Infrastructures

Supercomputers typically comprise thousands to millions of multi-core CPUs with GPU accelerators interconnected by complex interconnection networks that are typically structured as an intricate hierarchy of network switches. Capacity planning and management of such systems not only raises challenges in term of computing efficiency but also in term of energy consumption. Most legacy (SPMD) applications struggle to benefit from such infrastructure since the slightest failure or load imbalance immediately causes the whole program to stop or at best to waste resources. To scale and handle the stochastic nature of resources, these applications have to rely on dynamic runtimes that schedule computations and communications in an opportunistic way. Such evolution raises challenges not only in terms of programming but also in terms of observation (complexity and dynamicity prevents experiment reproducibility, intrusiveness hinders large scale data collection, ...) and analysis (dynamic and flexible application structures make classical visualization and simulation techniques totally ineffective and require to build on *ad hoc* information on the application structure).

4.2. Next-Generation Wireless Networks

Considerable interest has arisen from the seminal prediction that the use of multiple-input, multiple-output (MIMO) technologies can lead to substantial gains in information throughput in wireless communications, especially when used at a massive level. In particular, by employing multiple inexpensive service antennas, it is possible to exploit spatial multiplexing in the transmission and reception of radio signals, the only physical limit being the number of antennas that can be deployed on a portable device. As a result, the wireless medium can accommodate greater volumes of data traffic without requiring the reallocation (and subsequent re-regulation) of additional frequency bands. In this context, throughput maximization in the presence of interference by neighboring transmitters leads to games with convex action sets (covariance matrices with trace constraints) and individually concave utility functions (each user's Shannon throughput); developing efficient and distributed optimization protocols for such systems is one of the core objectives of Theme 5.

Another major challenge that occurs here is due to the fact that the efficient physical layer optimization of wireless networks relies on perfect (or close to perfect) channel state information (CSI), on both the uplink and the downlink. Due to the vastly increased computational overhead of this feedback – especially in decentralized, small-cell environments – the ongoing transition to fifth generation (5G) wireless networks is expected to go hand-in-hand with distributed learning and optimization methods that can operate reliably in feedback-starved environments. Accordingly, one of POLARIS' application-driven goals will be to leverage the algorithmic output of Theme 5 into a highly adaptive resource allocation framework for next-géneration wireless systems that can effectively "learn in the dark", without requiring crippling amounts of feedback.

4.3. Energy and Transportation

Smart urban transport systems and smart grids are two examples of collective adaptive systems. They consist of a large number of heterogeneous entities with decentralised control and varying degrees of complex autonomous behaviour. We develop an analysis tools to help to reason about such systems. Our work relies on tools from fluid and mean-field approximation to build decentralized algorithms that solve complex optimization problems. We focus on two problems: decentralized control of electric grids and capacity planning in vehicle-sharing systems to improve load balancing.

4.4. Social Computing Systems

Social computing systems are online digital systems that use personal data of their users at their core to deliver personalized services directly to the users. They are omnipresent and include for instance recommendation systems, social networks, online medias, daily apps, etc. Despite their interest and utility for users, these systems pose critical challenges of privacy, security, transparency, and respect of certain ethical constraints such as fairness. Solving these challenges involves a mix of measurement and/or audit to understand and assess issues, and modeling and optimization to propose and calibrate solutions.

5. Highlights of the Year

5.1. Highlights of the Year

- Bruno Gaujal joined the scientific committee of the GDR IM (Informatique Mathématique).
- Arnaud Legrand co-created a MOOC on "Recherche reproductible : principes méthodologiques pour une science transparente" hosted on the FUN platform https://www.fun-mooc.fr/courses/course-v1:inria+41016+session01bis/about.

5.1.1. Awards

- The paper by Nicolas Gast and co-authors received the Best Paper Award at ACM SIGMETRICS 2018.

- The paper by Patrick Loiseau and co-authors was nominated for the Best Paper Award at ACM FAT* 2018.
- The work on “Multi-Agent Online Learning with Imperfect Information” by Panayotis Mertikopoulos and co-authors was shortlisted for the INFORMS George Nicholson Best Student Paper Award.
- Panayotis Mertikopoulos received an Outstanding Reviewer Award at NIPS 2018.
- Benjamin Roussillon was co-laureate of the “Prix de mémoire de master 2018 en RO/AD” (best MSc thesis in operations research) from ROADEF for his Master thesis on “Development of adversarial classifiers using Bayesian games” under the supervision of Patrick Loiseau.

BEST PAPERS AWARDS :

[24]
N. GAST, B. V. HOUDT.*A Refined Mean Field Approximation*, in "ACM SIGMETRICS 2018", Irvine, France, June 2018, 1, https://hal.inria.fr/hal-01891642

[30]
T. SPEICHER, M. ALI, G. VENKATADRI, F. RIBEIRO, G. ARVANITAKIS, F. BENEVENUTO, K. P. GUMMADI, P. LOISEAU, A. MISLOVE.*Potential for Discrimination in Online Targeted Advertising*, in "FAT 2018 - Conference on Fairness, Accountability, and Transparency", New-York, United States, February 2018, vol. 81, p. 1-15, https://hal.archives-ouvertes.fr/hal-01955343

6. New Software and Platforms

6.1. Framesoc

FUNCTIONAL DESCRIPTION: Framesoc is the core software infrastructure of the SoC-Trace project. It provides a graphical user environment for execution-trace analysis, featuring interactive analysis views as Gantt charts or statistics views. It provides also a software library to store generic trace data, play with them, and build other analysis tools (e.g., Ocelotl).

- Participants: Arnaud Legrand and Jean-Marc Vincent
- Contact: Guillaume Huard
- URL: http://soctrace-inria.github.io/framesoc/

6.2. GameSeer

FUNCTIONAL DESCRIPTION: GameSeer is a tool for students and researchers in game theory that uses Mathematica to generate phase portraits for normal form games under a variety of (user-customizable) evolutionary dynamics. The whole point behind GameSeer is to provide a dynamic graphical interface that allows the user to employ Mathematica’s vast numerical capabilities from a simple and intuitive front-end. So, even if you’ve never used Mathematica before, you should be able to generate fully editable and customizable portraits quickly and painlessly.

- Contact: Panayotis Mertikopoulos
- URL: http://mescal.imag.fr/membres/panayotis.mertikopoulos/publications.html

6.3. marmoteCore

Markov Modeling Tools and Environments - the Core

KEYWORDS: Modeling - Stochastic models - Markov model

FUNCTIONAL DESCRIPTION: marmoteCore is a C++ environment for modeling with Markov chains. It consists in a reduced set of high-level abstractions for constructing state spaces, transition structures and Markov chains (discrete-time and continuous-time). It provides the ability of constructing hierarchies of Markov models, from the most general to the particular, and equip each level with specifically optimized solution methods.

This software is developed within the ANR MARMOTE project: ANR-12-MONU-00019.

- Participants: Alain Jean-Marie, Hlib Mykhailenko, Benjamin Briot, Franck Quessette, Issam Rabhi, Jean-Marc Vincent and Jean-Michel Fourneau
- Partner: UVSQ
- Contact: Alain Jean-Marie
- Publications: marmoteCore: a Markov Modeling Platform - marmoteCore: a software platform for Markov modeling
- URL: http://marmotecore.gforge.inria.fr/

6.4. Moca

Memory Organisation Cartography and Analysis
KEYWORDS: High-Performance Computing - Performance analysis

- Contact: David Beniamine
- URL: https://github.com/dbeniamine/MOCA

6.5. Ocelotl

Multidimensional Overviews for Huge Trace Analysis
FUNCTIONAL DESCRIPTION: Ocelotl is an innovative visualization tool, which provides overviews for execution trace analysis by using a data aggregation technique. This technique enables to find anomalies in huge traces containing up to several billions of events, while keeping a fast computation time and providing a simple representation that does not overload the user.

- Participants: Arnaud Legrand and Jean-Marc Vincent
- Contact: Jean-Marc Vincent
- URL: http://soctrace-inria.github.io/ocelotl/

6.6. PSI

Perfect Simulator
FUNCTIONAL DESCRIPTION: Perfect simulator is a simulation software of markovian models. It is able to simulate discrete and continuous time models to provide a perfect sampling of the stationary distribution or directly a sampling of functional of this distribution by using coupling from the past. The simulation kernel is based on the CFTP algorithm, and the internal simulation of transitions on the Aliasing method.

- Contact: Jean-Marc Vincent
- URL: http://psi.gforge.inria.fr/

6.7. SimGrid

KEYWORDS: Large-scale Emulators - Grid Computing - Distributed Applications
SCIENTIFIC DESCRIPTION: SimGrid is a toolkit that provides core functionalities for the simulation of distributed applications in heterogeneous distributed environments. The simulation engine uses algorithmic and implementation techniques toward the fast simulation of large systems on a single machine. The models are theoretically grounded and experimentally validated. The results are reproducible, enabling better scientific practices.

Its models of networks, cpus and disks are adapted to (Data)Grids, P2P, Clouds, Clusters and HPC, allowing multi-domain studies. It can be used either to simulate algorithms and prototypes of applications, or to emulate real MPI applications through the virtualization of their communication, or to formally assess algorithms and applications that can run in the framework.

The formal verification module explores all possible message interleavings in the application, searching for states violating the provided properties. We recently added the ability to assess liveness properties over arbitrary and legacy codes, thanks to a system-level introspection tool that provides a finely detailed view of the running application to the model checker. This can for example be leveraged to verify both safety or liveness properties, on arbitrary MPI code written in C/C++/Fortran.
NEWS OF THE YEAR: There were 3 major releases in 2018: The public API was sanitized (with compatibility wrappers in place). Th documentation was completely overhauled. Our continuous integration was greatly improved (45 Proxy Apps + BigDFT + StarPU + BatSim now tested nightly). Some kernel headers are now installed, allowing external plugins. Allow dynamic replay of MPI apps, controlled by S4U actors. Port the MPI trace replay engine to C++, fix visualization (+ the classical bug fixes and doc improvement).

- Participants: Adrien Lèbre, Arnaud Legrand, Augustin Degomme, Florence Perronnin, Frédéric Suter, Jean-Marc Vincent, Jonathan Pastor, Luka Stanisic and Martin Quinson
- Partners: CNRS - ENS Rennes
- Contact: Martin Quinson
- URL: https://simgrid.org/

6.8. Tabarnac

Tool for Analyzing the Behavior of Applications Running on NUMA ArChitecture
KEYWORDS: High-Performance Computing - Performance analysis - NUMA

- Contact: David Beniamine
- URL: https://dbeniamine.github.io/Tabarnac/

7. New Results

7.1. Design of Experiments

A large amount of resources is spent writing, porting, and optimizing scientific and industrial High Performance Computing applications, which makes autotuning techniques fundamental to lower the cost of leveraging the improvements on execution time and power consumption provided by the latest software and hardware platforms. Despite the need for economy, most autotuning techniques still require a large budget of costly experimental measurements to provide good results, while rarely providing exploitable knowledge after optimization. In [40], we present a user-transparent (white-box) autotuning technique based on Design of Experiments that operates under tight budget constraints by significantly reducing the measurements needed to find good optimizations. Our approach enables users to make informed decisions on which optimizations to pursue and when to stop. We present an experimental evaluation of our approach and show it is capable of leveraging user decisions to find the best global configuration of a GPU Laplacian kernel using half of the measurement budget used by other common autotuning techniques. We show that our approach is also capable of finding speedups of up to $50\times$, compared to gcc's-O3, for some kernels from the SPAPT benchmark suite, using up to $10\times$ less measurements than random sampling.

7.2. Experimenting with Fog Infrastructures

To this day, the Internet of Things (IoT) continues its explosive growth. Nevertheless, with the exceptional evolution of traffic demand, existing infrastructures are struggling to resist. In this context, Fog computing is shaping the future of IoT applications. Fog computing provides computing, storage and communication resources at the edge of the network, near the physical world. This section describes two independent contributions on how to study and develop FOG infrastructures. These contributions take place in the context of the Inria/Orange Labs joint laboratory.

- Despite its several advantages, Fog computing raises new challenges which slow its adoption down. In particular, there are currently few practical solutions allowing to exploit such infrastructure and to evaluate potential strategies. In [42], we propose a prototype orchestration architecture building on both Grid5000 and Fit-IoT lab (SILECS). This experimental testbed allows to realistically and rigorously evaluate orchestration strategies. In [20], we propose FITOR, an orchestration system for IoT applications in the Fog environment, which extends the actor-model based Calvin framework to cope with Fog environments while offering efficient orchestration mechanisms. In order to optimize the provisioning of Fog-Enabled IoT applications, FITOR relies on O-FSP, an optimized fog service provisioning strategy which aims to minimize the provisioning cost of IoT applications, while meeting their requirements. Based on extensive experiments, the results obtained show that O-FSP optimizes the placement of IoT applications and outperforms the related strategies in terms of i) provisioning cost ii) resource usage and iii) acceptance rate.
- End devices nearing the physical world can have interesting properties such as short delays, responsiveness, optimized communications and privacy. However, these end devices have low stability and are prone to failures. There is consequently a need for failure management protocols for IoT applications in the Fog. The design of such solutions is complex due to the specificities of the environment, i.e., (i) dynamic infrastructure where entities join and leave without synchronization, (ii) high heterogeneity in terms of functions, communication models, network, processing and storage capabilities, and, (iii) cyber-physical interactions which introduce non-deterministic and physical world's space and time dependent events. In [29], [37], we present a fault tolerance approach taking into account these three characteristics of the Fog-IoT environment. Fault tolerance is achieved by saving the state of the application in an uncoordinated way. When a failure is detected, notifications are propagated to limit the impact of failures and dynamically reconfigure the application. Data stored during the state saving process are used for recovery, taking into account consistency with respect to the physical world. The approach was validated through practical experiments on a smart home platform.

7.3. HPC Application Analysis and Visualization

- Programming paradigms in High-Performance Computing have been shifting towards task-based models which are capable of adapting readily to heterogeneous and scalable supercomputers. The performance of task-based application heavily depends on the runtime scheduling heuristics and on its ability to exploit computing and communication resources. Unfortunately, the traditional performance analysis strategies are unfit to fully understand task-based runtime systems and applications: they expect a regular behavior with communication and computation phases, while task-based applications demonstrate no clear phases. Moreover, the finer granularity of task-based applications typically induces a stochastic behavior that leads to irregular structures that are difficult to analyze. Furthermore, the combination of application structure, scheduler, and hardware information is generally essential to understand performance issues. The papers [36], [6] presents a flexible framework that enables one to combine several sources of information and to create custom visualization panels allowing to understand and pinpoint performance problems incurred by bad scheduling decisions in task-based applications. Three case-studies using StarPU-MPI, a task-based multi-node runtime system, are detailed to show how our framework can be used to study the performance of the well-known Cholesky factorization. Performance improvements include a better task partitioning among the multi-(GPU,core) to get closer to theoretical lower bounds, improved MPI pipelining in multi-(node,core,GPU) to reduce the slow start, and changes in the runtime system to increase MPI bandwidth, with gains of up to 13% in the total makespan.

- In the context of multi-physics simulations on unstructured and heterogeneous meshes, generating well-balanced partitions is not trivial. The computing cost per mesh element in different phases of the simulation depends on various factors such as its type, its connectivity with neighboring elements or its layout in memory with respect to them, which determines the data locality. Moreover, if different types of discretization methods or computing devices are combined, the performance variability across the domain increases. Due to all these factors, evaluate a representative computing cost per mesh element, to generate well-balanced partitions, is a difficult task. Nonetheless, load balancing is a critical aspect of the efficient use of extreme scale systems since idle-times can represent a huge waste of resources, particularly when a single process delays the overall simulation. In this context, we present in [16] some improvements carried out on an in-house geometric mesh partitioner based on the Hilbert Space-Filling Curve. We have previously tested its effectiveness by partitioning meshes with up to 30 million elements in a few tenths of milliseconds using up to 4096 CPU cores, and we have leveraged its performance to develop an autotuning approach to adjust the load balancing according to runtime measurements. In this paper, we address the problem of having different load distributions in different phases of the simulation, particularly in the matrix assembly and in the solution of the linear system. We consider a multi-partition approach to ensure a proper load balance in all the phases. The initial results presented show the potential of this strategy.

7.4. Energy Optimization and Smart Grids Simulation

Large-scale decentralized photovoltaic (PV) generators are currently being installed in many low-voltage distribution networks. Without grid reinforcements or production curtailment, they might create current and/or voltage issues. In [13], [45], we consider the use the advanced metering infrastructure (AMI) as the basis for PV generation control. We show that the advanced metering infrastructure may be used to infer some knowledge about the underlying network, and we show how this knowledge can be used by a simple feed-forward controller to curtail the solar production efficiently.

We developed a environment for co-simulating electrical networks, telecommunication networks and online learning algorithms [3]. One the outputs of this work was to allow us to perform realistic numerical simulations of active distribution networks. We used this simulator to compare our proposed controller with two other controller structures: open-loop, and feedback P (U) and Q(U). We demonstrate that our feed-forward controller –that requires no prior knowledge of the underlying electrical network– brings significant performance improvements as it can effectively suppress over-voltage and over-current while requiring low energy curtailment. This method can be implemented at low cost and require no specific information about the network on which it is deployed.

Finally, we study demand-Response (DR) programs, whereby users of an electricity network are encouraged by economic incentives to rearrange their consumption in order to reduce production costs. Such mechanisms are envisioned to be a key feature of the smart grid paradigm. Several recent works proposed DR mechanisms and used analytical models to derive optimal incentives. Most of these works, however, rely on a macroscopic description of the population that does not model individual choices of users. In in[4], we conduct a detailed analysis of those models and we argue that the macroscopic descriptions hide important assumptions that can jeopardize the mechanisms' implementation (such as the ability to make personalized offers and to perfectly estimate the demand that is moved from a timeslot to another). Then, we start from a microscopic description that explicitly models each user's decision. We introduce four DR mechanisms with various assumptions on the provider's capabilities. Contrarily to previous studies, we find that the optimization problems that result from our mechanisms are complex and can be solved numerically only through a heuristic. We present numerical simulations that compare the different mechanisms and their sensitivity to forecast errors. At a high level, our results show that the performance of DR mechanisms under reasonable assumptions on the provider's capabilities are significantly lower than those suggested by previous studies, but that the gap reduces when the population's flexibility increases.

7.5. Simulation of HPC Applications

Beside continuous development and contribution to the SimGrid project, the two following contributions have been published this year. Both build on the SMPI interface which allows to efficiently predict the performance of MPI applications.

- Finite-difference methods are commonplace in High Performance Computing applications. Despite their apparent regularity, they often exhibit load imbalance that damages their efficiency. In [9], we characterize the spatial and temporal load imbalance of Ondes3D, a typical finite-differences application dedicated to earthquake modeling. Our analysis reveals imbalance originating from the structure of the input data, and from low-level CPU optimizations. Ondes3D was successfully ported to AMPI/CHARM++ using over-decomposition and MPI process migration techniques to dynamically rebalance the load. However, this approach requires careful selection of the over-decomposition level, the load balancing algorithm, and its activation frequency. These choices are usually tied to application structure and platform characteristics. We have thus proposed a workflow that leverages the capabilities of SimGrid to conduct such study at low experimental cost. We rely on a combination of emulation, simulation, and application modeling that requires minimal code modification and manages to capture both spatial and temporal load imbalance to faithfully predict the performance of dynamic load balancing. We evaluate the quality of our simulation by comparing simulation results with the outcome of real executions and demonstrate how this approach can be used to quickly find the optimal load balancing configuration for a given application/hardware configuration.
- It is typical in High Performance Computing (HPC) courses to give students access to HPC platforms so that they can benefit from hands-on learning opportunities. Using such platforms, however, comes with logistical and pedagogical challenges. For instance, a logistical challenge is that access to representative platforms must be granted to students, which can be difficult for some institutions or course modalities; and a pedagogical challenge is that hands-on learning opportunities are constrained by the configurations of these platforms. A way to address these challenges is to instead simulate program executions on arbitrary HPC platform configurations. In [19] we focus on simulation in the specific context of distributed-memory computing and MPI programming education. While using simulation in this context has been explored in previous works, our approach offers two crucial advantages. First, students write standard MPI programs and can both debug and analyze the performance of their programs in simulation mode. Second, large-scale executions can be simulated in short amounts of time on a single standard laptop computer. This is possible thanks to SMPI, an MPI simulator provided as part of SimGrid. After detailing the challenges involved when using HPC platforms for HPC education and providing background information about SMPI, we present SMPI Courseware. SMPI Courseware is a set of in-simulation assignments that can be incorporated into HPC courses to provide students with hands-on experience for distributed-memory computing and MPI programming learning objectives. We describe some these assignments, highlighting how simulation with SMPI enhances the student learning experience.

7.6. Mean Field and Refined Mean Field Methods

Mean field approximation is a popular means to approximate stochastic models that can be represented as a system of N interacting objects. It is know to be exact as N goes to infinity. In a recent series of paper, [24], [25], [7], we establish theoretical results and numerical methods that allows us to define an approximation that is much more accurate than the classical mean field approximation. This new approximation, that we call the *refined mean field approximation*, is based on the computation of an expansion term of the order $1/N$. By considering a variety of applications, that include coupon collector, load balancing and bin packing problems, we illustrate that the proposed refined mean field approximation is significantly more accurate that the classic mean field approximation for small and moderate values of N: relative errors are often below 1% for systems with $N = 10$.

In [23], [8], we improve this result in two directions. First, we show how to obtain the same result for the transient regime. Second, we provide a further refinement by expanding the term in $1/N^2$ (both for transient and steady-state regime). Our derivations are inspired by moment-closure approximation, a popular technique in theoretical biochemistry. We provide a number of examples that show: (1) that this new approximation is usable in practice for systems with up to a few tens of dimensions, and (2) that it accurately captures the transient and steady state behavior of such systems.

7.7. Optimization of Networks and Communication

This section describes two independent contributions on the analysis and optimization of networks and communication.

- Telecommunication networks are converging to a massively distributed cloud infrastructure interconnected with software defined networks. In the envisioned architecture, services will be deployed flexibly and quickly as network slices. Our paper [26] addresses a major bottleneck in this context, namely the challenge of computing the best resource provisioning for network slices in a robust and efficient manner. With tractability in mind, we propose a novel optimization framework which allows fine-grained resource allocation for slices both in terms of network bandwidth and cloud processing. The slices can be further provisioned and auto-scaled optimally based on a large class of utility functions in real-time. Furthermore, by tuning a slice-specific parameter, system designers can trade off traffic-fairness with computing-fairness to provide a mixed fairness strategy. We also propose an iterative algorithm based on the alternating direction method of multipliers (ADMM) that provably converges to the optimal resource allocation and we demonstrate the method's fast convergence in a wide range of quasi-stationary and dynamic settings.
- Distributed power control schemes in wireless networks have been well-examined, but standard methods rarely consider the effect of potentially random delays, which occur in almost every real-world network. We present in paper [33] Robust Feedback Averaging, a novel power control algorithm that is capable of operating in delay-ridden and noisy environments. We prove optimal convergence of this algorithm in the presence of random, time-varying delays, and present numerical simulations that indicate that Robust Feedback Averaging outperforms the ubiquitous Foschini-Miljanic algorithm in several regimes.

7.8. Privacy, Fairness, and Transparency in Online Social Medias

Bringing transparency to algorithmic decision making systems and guaranteeing that the system satisfies properties of fairness and privacy is crucial in today's world. To start tackling this broad challenge, we focused on the case of online advertising and we had the following contributions.

- *Transparency properties for social media advertising and audit of Facebook's explanations.* In [15], we took a first step towards exploring the transparency mechanisms provided by social media sites, focusing on the two processes for which Facebook provides transparency mechanisms: the process of how Facebook infers data about users, and the process of how advertisers use this data to target users. We call explanations about those two processes *data explanations* and *ad explanations*, respectively.

  We identify a number of *properties* that are key for different types of explanations aimed at bringing transparency to social media advertising. We then evaluate empirically how well Facebook's explanations satisfy these properties and discuss the implications of our findings in view of the possible purposes of explanations. In particular, for *ad explanations*, we define five key properties: *personalization*, *completeness*, *correctness* (and the companion property of *misleadingness*), *consistency*, and *determinism*, and we show that Facebook's ad explanations are often *incomplete* and sometimes *misleading*. In particular, we observe that Facebook reveals only the most prevalent attribute used by the advertisers, which may allow malicious advertisers to easily obfuscate ad explanations from ad campaigns that are discriminatory or that target privacy-sensitive attributes. For *data explanations*, we define four key properties of the explanations: *specificity*, *snapshot completeness*, *temporal completeness*, and *correctness*; and we show that Facebook's explanations are *incomplete* and often *vague*; hence potentially limiting user control.

Overall, our study provides a first step towards better understanding and improving transparency in social media advertising. During this work, we developed the tool AdAnalyst (https://adanalyst.mpi-sws.org/), which was instrumental for the study but also provides a transparency tool on its own for the large public, and is anticipated to be the basis of a number of further research studies in transparency.

- *Potential for discrimination in social media advertising.* Recently, online targeted advertising platforms like Facebook have been criticized for allowing advertisers to discriminate against users belonging to sensitive groups, i.e., to exclude users belonging to a certain race or gender from receiving their ads. Such criticisms have led, for instance, Facebook to disallow the use of attributes such as ethnic affinity from being used by advertisers when targeting ads related to housing or employment or financial services. In our paper [30], we systematically investigate the different targeting methods offered by Facebook (traditional attribute- or interest-based targeting, custom audience and lookalike audience) for their ability to enable discriminatory advertising and showed that a malicious advertiser can create highly discriminatory ads without using sensitive attributes (hence banning those features is inefficient to solve the problem). We argue that discrimination measures should be based on the targeted population and not on the attributes used for targeting and propose a discrimination metric in this direction.
- *Identification and resolution of privacy leakages in the Facebook's advertising platform.* In paper [31] we discovered that the information provided to advertisers through the custom audience feature (where an advertisers can upload PIIs (Personally Identifiable Information) of their customers and Facebook matches those with their users) was very severely leaking personal information. Specifically, it was making it possible for a malicious advertiser knowing the email address of a user to discover its phone number. Perhaps even worse, it was allowing a malicious advertiser to de-anonymize visitors of a website he controls. We discovered that the problem was due to the way Facebook computes estimates of the number of users matching a list of PIIs and proposed a solution based on not de-duplicating records with different PIIs belonging to the same users; and we proved the robustness of our solution theoretically. Our work led to Facebook implementing a solution inspired by the one we proposed.

7.9. Optimization Methods

This section describes four independent contributions on optimization.

- In view of solving convex optimization problems with noisy gradient input, we analyze in the paper [11] the asymptotic behavior of gradient-like flows under stochastic disturbances. Specifically, we focus on the widely studied class of mirror descent schemes for convex programs with compact feasible regions, and we examine the dynamics' convergence and concentration properties in the presence of noise. In the vanishing noise limit, we show that the dynamics converge to the solution set of the underlying problem (a.s.). Otherwise, when the noise is persistent, we show that the dynamics are concentrated around interior solutions in the long run, and they converge to boundary solutions that are sufficiently "sharp". Finally, we show that a suitably rectified variant of the method converges irrespective of the magnitude of the noise (or the structure of the underlying convex program), and we derive an explicit estimate for its rate of convergence.
- We examine in paper [12] a class of stochastic mirror descent dynamics in the context of monotone variational inequalities (including Nash equilibrium and saddle-point problems). The dynamics under study are formulated as a stochastic differential equation driven by a (single-valued) monotone operator and perturbed by a Brownian motion. The system's controllable parameters are two variable weight sequences that respectively pre- and post-multiply the driver of the process. By carefully tuning these parameters, we obtain global convergence in the ergodic sense, and we estimate the average rate of convergence of the process. We also establish a large deviations principle showing that individual trajectories exhibit exponential concentration around this average.

- We develop in [17] a new stochastic algorithm with variance reduction for solving pseudo-monotone stochastic variational inequalities. Our method builds on Tseng's forward-backward-forward algorithm, which is known in the deterministic literature to be a valuable alternative to Korpelevich's extragradient method when solving variational inequalities over a convex and closed set governed with pseudo-monotone and Lipschitz continuous operators. The main computational advantage of Tseng's algorithm is that it relies only on a single projection step, and two independent queries of a stochastic oracle. Our algorithm incorporates a variance reduction mechanism, and leads to a.s. convergence to solutions of a merely pseudo-monotone stochastic variational inequality problem. To the best of our knowledge, this is the first stochastic algorithm achieving this by using only a single projection at each iteration.
- One of the most widely used training methods for large-scale machine learning problems is distributed asynchronous stochastic gradient descent (DASGD). However, a key issue in its implementation is that of delays: when a "worker" node asynchronously contributes a gradient update to the "master", the global model parameter may have changed, rendering this information stale. In massively parallel computing grids, these delays can quickly add up if a node is saturated, so the convergence of DASGD is uncertain under these conditions. Nevertheless, by using a judiciously chosen quasilinear step-size sequence, we show in [35] that it is possible to amortize these delays and achieve global convergence with probability 1, even under polynomially growing delays, reaffirming in this way the successful application of DASGD to large-scale optimization problems.

7.10. Multi-agent Learning and Distributed Best Response

This section describes several independent contributions on multi-agent learning.

- In [5], [22], [21], we study how fast can simple algorithms compute Nash equilibria. We study the case of random potential games for which we have designed and analyzed distributed algorithms to compute a Nash equilibrium. Our algorithms are based on best-response dynamics, with suitable revision sequences (orders of play). We compute the average complexity over all potential games of best response dynamics under a random i.i.d. revision sequence, since it can be implemented in a distributed way using Poisson clocks. We obtain a distributed algorithm whose execution time is within a constant factor of the optimal centralized one. We also showed how to take advantage of the structure of the interactions between players in a network game: non- interacting players can play simultaneously. This improves best response algorithm, both in the centralized and in the distributed case.
- In [10], we study a class of evolutionary game dynamics defined by balancing a gain determined by the game's payoffs against a cost of motion that captures the difficulty with which the population moves between states. Costs of motion are represented by a Riemannian metric, i.e., a state-dependent inner product on the set of population states. The replicator dynamics and the (Euclidean) projection dynamics are the archetypal examples of the class we study. Like these representative dynamics, all Riemannian game dynamics satisfy certain basic desiderata, including positive correlation and global convergence in potential games. Moreover, when the underlying Riemannian metric satisfies a Hessian integrability condition, the resulting dynamics preserve many further properties of the replicator and projection dynamics. We examine the close connections between Hessian game dynamics and reinforcement learning in normal form games, extending and elucidating a well-known link between the replicator dynamics and exponential reinforcement learning.
- The paper [18] examines the long-run behavior of learning with bandit feedback in non-cooperative concave games. The bandit framework accounts for extremely low-information environments where the agents may not even know they are playing a game; as such, the agents' most sensible choice in this setting would be to employ a no-regret learning algorithm. In general, this does not mean that the players' behavior stabilizes in the long run: no-regret learning may lead to cycles, even with perfect gradient information. However, if a standard monotonicity condition is satisfied, our analysis shows that no-regret learning based on mirror descent with bandit feedback converges to Nash equilibrium

with probability 1. We also derive an upper bound for the convergence rate of the process that nearly matches the best attainable rate for single-agent bandit stochastic optimization.

- In [34], we consider a game-theoretical multi-agent learning problem where the feedback information can be lost during the learning process and rewards are given by a broad class of games known as variationally stable games. We propose a simple variant of the classical online gradient descent algorithm, called reweighted online gradient descent (ROGD) and show that in variationally stable games, if each agent adopts ROGD, then almost sure convergence to the set of Nash equilibria is guaranteed, even when the feedback loss is asynchronous and arbitrarily corrrelated among agents. We then extend the framework to deal with unknown feedback loss probabilities by using an estimator (constructed from past data) in its replacement. Finally, we further extend the framework to accomodate both asynchronous loss and stochastic rewards and establish that multi-agent ROGD learning still converges to the set of Nash equilibria in such settings. Together, these results contribute to the broad lanscape of multi-agent online learning by significantly relaxing the feedback information that is required to achieve desirable outcomes.
- Regularized learning is a fundamental technique in online optimization, machine learning and many other fields of computer science. A natural question that arises in these settings is how regularized learning algorithms behave when faced against each other. In the paper [27], we study a natural formulation of this problem by coupling regularized learning dynamics in zero-sum games. We show that the system's behavior is Poincaré recurrent, implying that almost every trajectory revisits any (arbitrarily small) neighborhood of its starting point infinitely often. This cycling behavior is robust to the agents' choice of regularization mechanism (each agent could be using a different regularizer), to positive-affine transformations of the agents' utilities, and it also persists in the case of networked competition, i.e., for zero-sum polymatrix games.

7.11. Blotto games

The Colonel Blotto game is a famous game commonly used to model resource allocation problems in many domains ranging from security to advertising. Two players distribute a fixed budget of resources on multiple battlefields to maximize the aggregate value of battlefields they win, each battlefield being won by the player who allocates more resources to it. The continuous version of the game –where players can choose any fractional allocation– has been extensively studied, albeit only with partial results to date. Recently, the discrete version –where allocations can only be integers– started to gain traction and algorithms were proposed to compute the equilibrium in polynomial time; but these remain computationally impractical for large (or even moderate) numbers of battlefields. In [32], [46], we propose an algorithm to compute very efficiently an approximate equilibrium for the discrete Colonel Blotto game with many battlefields. We provide a theoretical bound on the approximation error as a function of the game's parameters, in particular number of battlefields and resource budgets. We also propose an efficient dynamic programming algorithm to compute the best-response to any strategy that allows computing for each game instance the actual value of the error. We perform numerical experiments that show that the proposed strategy provides a fast and good approximation to the equilibrium even for moderate numbers of battlefields.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- Bilateral contrat with Enedis (Linky-Lab), Post-doctoral position for 18th months (Mouhcine Mendil).
- ULTRON, bilateral contract with Huawei over 18 months, supporting two postdoctoral researchers, Amélie Heliou and Luigi Vigneri.
- Inria/Orange Labs Laboratory. Polaris is involved in this partnership with Orange Labs by supervising two PhD students in the context of this common laboratory.

- Cifre contract with Schneider Electric. The PhD thesis of Benoit Vinot (supervised by Nicolas Gast and Florent Cadoux (G2Elab)) is supported by this collaboration.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. IDEX UGA

Nicolas Gast received a grant from the IDEX UGA that funds a two-years post-doctoral researcher (Takai Kennouche) for two years (2018 and 2019) to work on the smart-grid project that focus on distributed optimization in electrical distribution networks.

Patrick Loiseau and Panayotis Mertikopoulos received a grant from the IDEX UGA that partly funds a PhD student (Benjamin Roussillon) to work on game theoretic models for adversarial classification.

9.2. National Initiatives

9.2.1. Inria Project Labs

Arnaud Legrand is the leader of the HAC SPECIS project. The goal of the HAC SPECIS (High-performance Application and Computers: Studying PErformance and Correctness In Simulation) project is to answer methodological needs of HPC application and runtime developers and to allow to study real HPC systems both from the correctness and performance point of view. To this end, we gather experts from the HPC, formal verification and performance evaluation community. Inria Teams: AVALON, POLARIS, MYRIADS, SUMO, HIEPACS, STORM, MEXICO, VERIDIS.

9.2.2. DGA Grants

Patrick Loiseau and Panayotis Mertikopoulos received a grant from DGA that complements the funding of PhD student (Benjamin Roussillon) to work on game theoretic models for adversarial classification.

9.2.3. PGMO Projects

PGMO projects are supported by the Jacques Hadamard Mathematical Foundation (FMJH). Our project (HEAVY.NET) is focused on congested networks and their asymptotic properties.

9.2.4. PEPS

Panayotis Mertikopoulos est co-PI of a PEPS I3A project: MixedGAN ("Mixed-strategy generative adversarial networks") (PI: R. Laraki, U. Dauphine).

9.2.5. Fondation Blaise Pascal

Project IAM (Informatique à la Main) funded by fondation Blaise Pascal (Jean-Marc Vincent).

9.2.6. ANR

- *ORACLESS (2016–2021)*
  ORACLESS is an ANR starting grant (JCJC) coordinated by Panayotis Mertikopoulos. The goal of the project is to develop highly adaptive resource allocation methods for wireless communication networks that are provably capable of adapting to unpredictable changes in the network. In particular, the project will focus on the application of online optimization and online learning methodologies to multi-antenna systems and cognitive radio networks.
- *CONNECTED (2016–2019)*
  CONNECTED is an ANR Tremplin-ERC (T-ERC) grant coordinated by Patrick Loiseau. The goal of the project is to work on several game-theoretic models involving learning agents and data revealed by strategic agents in response to the learning algorithms, so as to derive better learning algorithms for such special data.

9.3. International Initiatives

9.3.1. Inria International Labs

The POLARIS team is involved in the JLESC (Joint Laboratory for Extreme-Scale Computing) with University of University of Illinois Urbana Champaign, Argonne Nat. Lab and BSC.

9.3.2. Participation in Other International Programs

- *LICIA:* The CNRS, Inria, the Universities of Grenoble, Grenoble INP, and Universidade Federal do Rio Grande do Sul have created the LICIA (*Laboratoire International de Calcul intensif et d'Informatique Ambiante*). LICIA's main research themes are high performance computing, language processing, information representation, interfaces and visualization as well as distributed systems. Jean-Marc Vincent is the director of the laboratory on the French side and visited Porto Alegre for three weeks in November 2018.

  More information can be found at http://www.inf.ufrgs.br/licia/.
- *GENE*: Stochastic dynamics of large games and networks. This is a joint project (2018 - 2019) with Universidad de Buenos Aires, Argentina (Matthieu Jonckheere), Universidad de la Republica Uruguay (Federico La Rocca), CNRS (Balakrishna Prabhu) and Universidad ORT Uruguay (Andrés Ferragut).

  Through the creation and consolidation of strong research and formation exchanges between Argentina, France and Uruguay, the GENE project will contribute to the fields of performance evaluation and control of communication networks, using tools of game theory, probability theory and control theory. Some of the challenges this project will address are: (1) Mean-field games and their application to load balancing and resource allocations, (2) Scaling limits for centralized and decentralized load balancing strategies and implementation of practical policies for web servers farms, (3) Information diffusion and communication protocols in large and distributed wireless networks.
- *LEARN*: Learning algorithms for games and applications (2016-2018). POLARIS is a member of the Franco-Chilean collaboration network LEARN with CONICYT (the Chilean national research agency), formed under the ECOS-Sud framework. The main research themes of this network is the application of continuous optimization and game-theoretic learning methods to traffic routing and congestion control in data networks. Panayotis Mertikopoulos was an invited researcher at the University of Chile in October 2016.

  More information can be found at http://www.conicyt.cl/pci/2016/02/11/programa-ecos-conicyt-adjudica-proyectos-para-el-ano-2016.

9.4. International Research Visitors

9.4.1. Visits to International Teams

9.4.1.1. Research Stays Abroad

Panayotis Mertikopoulos was a visiting scientist at UC Berkeley / Simons Institute for the Theory of Computing (Feb.-March 2018) and a visiting scientist at U Athens / STSM in the framework of the EU COST Action GAMENET (Apr. - May 2018).

Jean-Marc Vincent is the director of Licia (Laboratoire de Calcul Intensif et d'Informatique Ambiante) and stayed 20 days at Porto Alegre to teaching and nurture research collaborations.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Panayotis Mertikopoulos was involved in the following events:

- 2018 French Days on Optimization and Decision Science ("Journées SMAI MODE 2018"): general co-chair
- 2018 Paris Symposium on Game Theory (Paris, June 2018): co-organizer
- GDO '18: the 2018 Workshop on Games, Dynamics, and Optimization (Vienna, March 2018): co-organizer

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- SBAC-PAD 2018 (Arnaud Legrand: chair of the performance evaluation track)

10.1.2.2. Member of the Conference Program Committees

- Performance 2018 (Bruno Gaujal, Nicolas Gast)
- SIGMETRICS 2018 (Nicolas Gast)
- WiOpt 2018 (Bruno Gaujal, Patrick Loiseau)
- NetGCoop 2018 (Bruno Gaujal, Patrick Loiseau)
- NIPS 2018 (Panayotis Mertikopoulos, Patrick Loiseau)
- ICML 2018 (Patrick Loiseau)
- SuperComputing 2018 (Arnaud Legrand)
- RescueHPC 2018 (Arnaud Legrand)
- EPEW 2018 (Jean-Marc Vincent)
- Valuetools 2018 (Jean-Marc Vincent)
- NetEcon 2018 (Patrick Loiseau)

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Patrick Loiseau is Associate Editor of ACM Trans. on Internet Technology and of IEEE Trans. on Big Data.

10.1.4. Invited Talks

- Arnaud Legrand gave a keynote on "Simulation of Large-Scale Distributed Computing Infrastructures" at Orange Labs, Chatillon, October 2018 and on "Reproducible Research"at Inria Rennes in May 2018.
- Bruno Gaujal gave invited presentations at Paris Symposium on Game Theory (Paris), International Symposium on Dynamic Games (Grenoble), New trends in Scheduling (Aussois).
- Panayotis Mertikopoulos gave invited presentations at Trinity College Dublin, Ireland on *Efficient network utility maximization algorithms*, at National Technical University of Athens (Athens Polytechnic) Athens, Greece on *Traffic in congested networks: Equilibrium, efficiency, and dynamics*, at GDO 2018 in Vienna, Austria, and on *Bandit learning in concave N-person games*, at UC Berkeley (Simons Institute for the Theory of Computing), USA on *Online learning in games*.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

The POLARIS members teach regularly. We only mention here lectures at the Master level.

- Master: Bruno Gaujal and Mouhcine Mendil, *"Advanced Performance Evaluation"*, 18h (M2), ENSIMAG
- Master: Bruno Gaujal and Panayotis Mertikopoulos, *"Online decision making"*, M2 ENS Lyon
- Master: Bruno Gaujal and Ana Busic, *"Cours théorie des réseaux"*, M2 MPRI, Paris
- Master M2R : Nicolas Gast *"Optimization Under Uncertainties"*, 18h (M2), Master ORCO, Grenoble.
- Master: Arnaud Legrand and Jean-Marc Vincent, *"Scientific Methodology and Performance Evaluation"*, 18h M2, M2R MOSIG
- Master: Arnaud Legrand, *"Scientific Methodology and Performance Evaluation"*, 18h M2, ENS Rennes
- Master: Panayotis Mertikopoulos, *"Advanced optimization algorithms"*, 16h M2, University of Athens, Athens, Greece
- Master: Guillaume Huard, *"Conception des Systèmes d'Exploitation"* (M1), Université Grenoble-Alpes
- Master: Guillaume Huard, *"Conception des Systèmes d'Exploitation"* (M1), Université Grenoble-Alpes
- Master: Florence Perronnin, *"Simulation"*, M1, Université Versailles – Saint-Quentin
- Master: Arnaud Legrand and Florence Perronnin, *"Probabilités–Simulation"* and *"Performance evaluation"* 72 h (M1), RICM4 Polytech Grenoble
- Master: Jean-Marc Vincent, Mathematics for computer science, 18 h , (M1) Mosig.
- Master/PhD: Jean-Marc Vincent, Litterate Programming and Statistics, UFRGS (Porto Alegre, Brazil)
- Master: Vincent Danjean, Architecture and Software project, engineering school
- Master: Vincent Danjean, Conception of operating systems, concurrent programming and systems project, MOSIG and CS Master, Grenoble

E-learning Arnaud Legrand has designed and organized a MOOC on Reproducible Research with Konrad Hinsen (CNRS/Centre de Biophysique Moléculaire) and Christophe Pouzat (CNRS/ Mathématiques Appliquées à Paris 5) with the support of the Inria MOOC-lab.

This MOOC is hosted on the FUN platform https://www.fun-mooc.fr/courses/course-v1:inria+41016+session01bis/about and the first edition (Oct-Dec 2018) targets graduate students, PhD students, post-doc, engineers and researchers working in any scientific domain relying on computations. In this MOOC, some modern and reliable tools are presented: GitLab for version control and collaborative working, Computational notebooks (Jupyter, RStudio, and Org-Mode) for efficiently combining the computation, presentation, and analysis of data. More than 3,400 people have registered to the first edition.

10.2.2. Supervision

Stephane Durand (PhD UGA defended on Dec. 11, 2018): Distributed Best Response Algorithms in random Potential Games, co-advised by Bruno Gaujal and Federica Garin (funded by Labex Persyval, Grenoble).

Benoit Vinot (PhD UGA defended on April 2018): Design of a distributed information system for the control of flexibilities in a power distribution network: modelling, simulation and implementation, co-advised by Nicolas Gast and Florent Cadoux

Vinicius Garcia Pinto (PhD in co-tutelle with UFRGS defended in 2018): Performance analysis and visualization of dynamic task-based applications: co-advised by Arnaud Legrand, Lucas Schnorr and Nicolas Maillard (funded by the Brazilian government).

Rafael Tesser (PhD in co-tutelle with UFRGS defended in 2018): Simulation and performance evaluation of dynamical load balancing of an over-decomposed Geophysics application, co-advised by Arnaud Legrand, Lucas Schnorr and Philippe Navaux (funded by the Brazilian government).

Alexis Janon (PhD in progress, 2018-...): Tasks Placement on Hierarchical Computational Platforms, co-advised by Guillaume Huard and Arnaud Legrand (funded by the French Ministry)

Stephan Plassart (PhD in progress, 2016-...): Energy Optimization in Embedded Systems, co-advised by Bruno Gaujal and Alain Girault (funded by Labex Persyval, Grenoble).

Baptiste Jonglez (PhD in progress, 2016-...): Leveraging Diversity in Communication Networks, co-advised by Bruno Gaujal and Martin Heusse (funded by Univ Grenoble Alpes).

Vitalii Emelianov (PhD in progress, 2018-...): Fairness and transparency in data-driven decision making, co-advised by Patrick Loiseau and Nicolas Gast (funded by Inria).

Benjamin Roussillon (PhD in progress, 2018-...): Classification in the presence adversarial data: models and solutions, co-advised by Patrick Loiseau and Panayotis Mertikopoulos (funded by IDEX UGA and DGA).

Dong Quan Vu (PhD in progress, 2017-...): Learning in Blotto games and applications to modeling attention in social networks, co-advised by Patrick Loiseau and Alonso Silva (Cifre PhD with Nokia Bell-Labs)

Athanasios Andreou (PhD in progress, 2015-...): Bringing transparency to personalized systems through statistical inference, co-advised by Patrick Loiseau and Oana Goga (funded by Institut Mines Telecom and ANR)

Alexandre Marcastel (PhD in progress, 2015-...): co-advised by E. Veronica Belmega, Panayotis Mertikopoulos and Inbar Fijalkow

Kimon Antonakopoulos (PhD in progress, 2017-...): Variational inequalities and optimization, co-advised by E. Veronica Belmega, Panayotis Mertikopoulos and Bruno Gaujal

Bruno Donassolo (PhD in progress, 2017-...): Decentralized management of applications in Fog computing environments, co-supervised by Panayotis Mertikopoulos, Arnaud Legrand and Ilhem Fajjari (Cifre PhD with Orange)

Pedro Bruel (PhD in progress co-advised with USP 2017-...): Design of experiments and autotuning of HPC computation kernels, co-advised by Arnaud Legrand, Alfredo Goldman and Brice Videau (funded by the Brazilian Government).

Tom Cornebize (PhD in progress 2017-...): Capacity planning and performance evaluation of supercomputers, advised by Arnaud Legrand (funded by the French Ministry for Research).

Christian Heinrich (PhD in progress 2015-...): Modeling of performance and energy consumption of HPC systems, advised by Arnaud Legrand (funded by Inria).

Umar Ozeer (PhD in progress 2017-...): co-advised by Jean-Marc Vincent, Gwen Salaün, François-Gaël Ottogalli and Loic Letondeur (within the Inria-Orange lab).

Amélie Héliou (PostDoc, Sep. 2017-May 2018): co-supervised by Panayotis Mertikopoulos and Bruno Gaujal

Mouhcine Mendil (PostDoc, 2017-...): supervised by Nicolas Gast

Takai Kennouche (PostDoc, 2017-...): supervised by Nicolas Gast

Luigi Vigneri (PostDoc, Sep. 2017-Sep.2018): co-supervised by Panayotis Mertikopoulos and G. Paschos

Olivier Bilenne (PostDoc, 2018-...): co-supervised by Panayotis Mertikopoulos and E. V. Belmega

10.2.3. Juries

- Bruno Gaujal was president of the jury for the CRCN Inria Grenoble Rhone-Alpes competition.
- Jean-Marc Vincent was member of the jury of Capes de mathématiques (option Informatique).
- Bruno Gaujal was reviewer of the PhD of Adil Salim (Telecom Paris Tech) and Panayotis Mertikopoulos was examinator.
- Nicolas Gast was reviewer of the PhD of Fabbio Cecchi (Univ. Eindhoven).
- Arnaud Legrand was president of the jury for the PhD of Louis Poirel (Univ. Bordeaux).
- Arnaud Legrand was president of the jury for the PhD of Nicolas Denoyelle (Univ. Bordeaux).
- Arnaud Legrand was president of the jury for the PhD of Valentin Reis (Univ. Grenoble Alpes).
- Arnaud Legrand was examinator of the jury for the PhD of Marcos Amaris Gonzales (Univ. São Paulo).

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- Jean-Marc Vincent is responsible for the mediation in the Inria Rhône-Alpes center in relation with the Rectorat de l'Académie de Grenoble (organisation of ISN conferences, Class'Code for digital referents)
- Jean-Marc Vincent coordinates the group "Info sans ordi" (computer science without computer)–in which Florence Perronnin participates

10.3.2. Articles and contents

Jean-Marc Vincent has coordinated and participated to the redaction of the 96 pages special issue of *Tangente* on *Informatique Débranchée*: http://www.infinimath.com/librairie/descriptif_livre.php?type=magazines&theme=7&soustheme=26&ref=2568#article.

10.3.3. Education

- Highschool Professors: Vincent Danjean is responsible of the University Dept. for Highschool Professors Training in Computer Science (in relation with the rectorat).
- Highschool Professors: Jean-Marc Vincent is member of the steering committee for the training of Highschool Professors towards the new option NSI (Numérique et Sciences Informatiques) for Baccalauréat and participated in the creation of a inter-university diploma in CS.
- Arnaud Legrand gave a lecture on "Reproducible Research" at CIRM in May 2018 to the computer science teachers of classes préparatoires.
- Bruno Gaujal: Course on game theory at the 7 laux of ENS Lyon students
- Bruno Gaujal: Course on dynamique optimization for high school teachers.
- Florence Perronnin: creation of an option in Licence on "Sciences Informatiques et Médiation" (CS and mediation)

10.3.4. Interventions

- Arnaud Legrand made a podcast with Interstice on reproducible research https://interstices.info/la-recherche-reproductible-pour-une-science-transparente/
- Participation to mediation events of the center (Fête de la science, journées math C2+, journée login)
- Participation to mediation actions Inria/Irem/Maison for Science (Jean-Marc Vincent)

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] V. GARCIA PINTO.*Performance Analysis Strategies for Task-based Applications on Hybrid Platforms*, Universidade Federal do Rio Grande do Sul - UFRGS ; UGA - Université Grenoble Alpes, October 2018, https://tel.archives-ouvertes.fr/tel-01962333

[2] R. KELLER TESSER.*A Simulation Workflow to Evaluate the Performance of Dynamic Load Balancing with Over-decomposition for Iterative Parallel Applications*, Universidade Federal Do Rio Grande Do Sul, April 2018, https://tel.archives-ouvertes.fr/tel-01962082

[3] B. VINOT.*Design of a distributed information system for the control of flexibilities in a power distribution network: modelling, simulation and implementation*, UGA - Université Grenoble Alpes ; MSTII, June 2018, https://tel.archives-ouvertes.fr/tel-01875320

Articles in International Peer-Reviewed Journal

[4] A. BENEGIAMO, P. LOISEAU, G. NEGLIA.*Dissecting demand response mechanisms: The role of consumption forecasts and personalized offers*, in "Sustainable Energy, Grids and Networks", December 2018, vol. 16, p. 156-166 [*DOI :* 10.1016/J.SEGAN.2018.07.005], https://hal.archives-ouvertes.fr/hal-01955356

[5] S. DURAND, F. GARIN, B. GAUJAL.*Distributed best response dynamics with high playing rates in potential games*, in "Performance Evaluation", October 2018, vol. 129, p. 40-59 [*DOI :* 10.1016/J.PEVA.2018.09.007], https://hal.inria.fr/hal-01940128

[6] V. GARCIA PINTO, L. M. SCHNORR, L. STANISIC, A. LEGRAND, S. THIBAULT, V. DANJEAN.*A Visual Performance Analysis Framework for Task-based Parallel Applications running on Hybrid Clusters*, in "Concurrency and Computation: Practice and Experience", April 2018, vol. 30, n^{o} 18, p. 1-31 [*DOI :* 10.1002/CPE.4472], https://hal.inria.fr/hal-01616632

[7] N. GAST, L. BORTOLUSSI, M. TRIBASTONE.*Size Expansions of Mean Field Approximation: Transient and Steady-State Analysis*, in "Performance Evaluation", 2018, p. 1-15, https://hal.inria.fr/hal-01891632

[8] N. GAST, D. LATELLA, M. MASSINK.*A refined mean field approximation of synchronous discrete-time population models*, in "Performance Evaluation", July 2018, p. 1-27, https://arxiv.org/abs/1807.08585 [*DOI :* 10.1016/J.PEVA.2018.05.002], https://hal.inria.fr/hal-01845235

[9] R. KELLER TESSER, L. MELLO SCHNORR, A. LEGRAND, C. HEINRICH, F. DUPROS, P. O. ALEXANDRE NAVAUX.*Performance Modeling of a Geophysics Application to Accelerate the Tuning of Over-decomposition Parameters through Simulation*, in "Concurrency and Computation: Practice and Experience", 2018, p. 1-21 [*DOI :* 10.1002/CPE.5012], https://hal.inria.fr/hal-01891416

[10] P. MERTIKOPOULOS, W. H. SANDHOLM.*Riemannian game dynamics*, in "Journal of Economic Theory", September 2018, vol. 177, p. 315-364 [*DOI :* 10.1016/J.JET.2018.06.002], https://hal.archives-ouvertes.fr/hal-01382281

[11] P. MERTIKOPOULOS, M. STAUDIGL.*On the convergence of gradient-like flows with noisy gradient input*, in "SIAM Journal on Optimization", January 2018, vol. 28, n^{o} 1, p. 163-197, https://hal.inria.fr/hal-01404586

[12] P. MERTIKOPOULOS, M. STAUDIGL.*Stochastic mirror descent dynamics and their convergence in monotone variational inequalities*, in "Journal of Optimization Theory and Applications", 2018, vol. 179, n^{o} 3, p. 838-867 [*DOI :* 10.1007/S10957-018-1346-X], https://hal.archives-ouvertes.fr/hal-01643343

Invited Conferences

[13] B. VINOT, F. CADOUX, N. GAST.*Congestion Avoidance in Low-Voltage Networks by using the Advanced Metering Infrastructure*, in "ePerf 2018 - IFIP WG PERFORMANCE - 36th International Symposium on

Computer Performance, Modeling, Measurements and Evalution", Toulouse, France, December 2018, p. 1-3, https://hal.inria.fr/hal-01953386

International Conferences with Proceedings

[14] A. ANDREOU, M. SILVA, F. BENEVENUTO, O. GOGA, P. LOISEAU, A. MISLOVE.*Measuring the Facebook Advertising Ecosystem*, in "NDSS 2019 - Proceedings of the Network and Distributed System Security Symposium", San Diego, United States, February 2019 [*DOI :* 10.14722/NDSS.2019.23280], https://hal.archives-ouvertes.fr/hal-01959145

[15] A. ANDREOU, G. VENKATADRI, O. GOGA, K. P. GUMMADI, P. LOISEAU, A. MISLOVE.*Investigating Ad Transparency Mechanisms in Social Media: A Case Study of Facebook's Explanations*, in "NDSS 2018 - Network and Distributed System Security Symposium", San Diego, United States, February 2018, p. 1-15 [*DOI :* 10.14722/NDSS.2018.23204], https://hal.archives-ouvertes.fr/hal-01955309

[16] R. BORRELL, J. C. CAJAS, L. M. SCHNORR, A. LEGRAND, G. HOUZEAUX.*SFC based multi-partitioning for accurate load balancing of CFD simulations*, in "Tenth International Conference on Computational Fluid Dynamics (ICCFD10)", Barcelona, Spain, July 2018, https://hal.inria.fr/hal-01896947

[17] R. BOT, P. MERTIKOPOULOS, M. STAUDIGL, P. T. VUONG.*On the convergence of stochastic forward-backward-forward algorithms with variance reduction in pseudo-monotone variational inequalities*, in "NIPS 2018 - Workshop on Smooth Games, Optimization and Machine Learning", Montréal, Canada, December 2018, p. 1-5, https://hal.inria.fr/hal-01949361

[18] M. BRAVO, D. S. LESLIE, P. MERTIKOPOULOS.*Bandit learning in concave N-person games*, in "NIPS 2018 - Thirty-second Conference on Neural Information Processing Systems", Montréal, Canada, December 2018, p. 1-24, https://hal.inria.fr/hal-01891523

[19] H. CASANOVA, A. LEGRAND, M. QUINSON, F. SUTER.*SMPI Courseware: Teaching Distributed-Memory Computing with MPI in Simulation*, in "EduHPC-18 - Workshop on Education for High-Performance Computing", Dallas, United States, November 2018, p. 1-10, https://hal.inria.fr/hal-01891513

[20] B. DONASSOLO, I. FAJJARI, A. LEGRAND, P. MERTIKOPOULOS.*Fog Based Framework for IoT Service Provisioning*, in "IEEE Consumer Communications & Networking Conference", Las Vegas, United States, January 2019, https://hal.inria.fr/hal-01859695

[21] S. DURAND, F. GARIN, B. GAUJAL.*Distributed Best Response Algorithms for Potential Games*, in "ECC 2018 - 16th European Control Conference", Limassol, Cyprus, June 2018, p. 1-6, https://hal.inria.fr/hal-01726836

[22] S. DURAND, F. GARIN, B. GAUJAL.*Efficiency of Best Response Dynamics with High Playing Rates in Potential Games*, in "IFIP WG 7.3 Performance 2018 - 36th International Symposium on Computer Performance, Modeling, Measurements and Evaluation", Toulouse, France, December 2018, p. 1-2, https://hal.inria.fr/hal-01940150

[23] N. GAST, L. BORTOLUSSI, M. TRIBASTONE.*Size Expansions of Mean Field Approximation: Transient and Steady-State Analysis*, in "2018 - 36th International Symposium on Computer Performance, Modeling, Measurements and Evaluation", Toulouse, France, December 2018, p. 1-2, https://hal.inria.fr/hal-01891636

[24] *Best Paper*
N. GAST, B. V. HOUDT.*A Refined Mean Field Approximation*, in "ACM SIGMETRICS 2018", Irvine, France, June 2018, 1, https://hal.inria.fr/hal-01891642.

[25] N. GAST, D. LATELLA, M. MASSINK.*A Refined Mean Field Approximation for Synchronous Population Processes*, in "MAMA 2018Workshop on MAthematical performance Modeling and Analysis", Irvine, United States, June 2018, p. 1-3, https://hal.inria.fr/hal-01891629

[26] M. LECONTE, G. PASCHOS, P. MERTIKOPOULOS, U. KOZAT.*A resource allocation framework for network slicing*, in "INFOCOM 2018 - IEEE International Conference on Computer Communications", Honolulu, United States, April 2018, p. 1-9, https://hal.archives-ouvertes.fr/hal-01643349

[27] P. MERTIKOPOULOS, C. H. PAPADIMITRIOU, G. PILIOURAS.*Cycles in adversarial regularized learning*, in "SODA '18 - Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms", New Orleans, United States, January 2018, p. 2703-2717, https://hal.archives-ouvertes.fr/hal-01643338

[28] M. MINAEI, M. MONDAL, P. LOISEAU, K. P. GUMMADI, A. KATE.*Forgetting the Forgotten with Lethe: Conceal Content Deletion from Persistent Observers*, in "PETS 2019 - 19th Privacy Enhancing Technologies Symposium", Stockholm, Sweden, July 2019, p. 1-21, https://hal.archives-ouvertes.fr/hal-01959119

[29] U. OZEER, X. ETCHEVERS, L. LETONDEUR, F.-G. OTTOGALLI, G. SALAÜN, J.-M. VINCENT.*Resilience of Stateful IoT Applications in a Dynamic Fog Environment*, in "EAI International Conference on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous '18)", New York, United States, November 2018, p. 1-10 [*DOI :* 10.1145/3286978.3287007], https://hal.archives-ouvertes.fr/hal-01927286

[30] *Best Paper*
T. SPEICHER, M. ALI, G. VENKATADRI, F. RIBEIRO, G. ARVANITAKIS, F. BENEVENUTO, K. P. GUMMADI, P. LOISEAU, A. MISLOVE.*Potential for Discrimination in Online Targeted Advertising*, in "FAT 2018 - Conference on Fairness, Accountability, and Transparency", New-York, United States, February 2018, vol. 81, p. 1-15, https://hal.archives-ouvertes.fr/hal-01955343.

[31] G. VENKATADRI, A. ANDREOU, Y. LIU, A. MISLOVE, K. P. GUMMADI, P. LOISEAU, O. GOGA.*Potential for Discrimination in Online Targeted Advertising*, in "FAT 2018 - Conference on Fairness, Accountability, and Transparency", New-York, United States, February 2018, vol. 81, p. 1-15, https://hal.archives-ouvertes.fr/hal-01955327

[32] D. Q. VU, P. LOISEAU, A. SILVA.*Efficient Computation of Approximate Equilibria in Discrete Colonel Blotto Games*, in "IJCAI-ECAI 2018 - 27th International Joint Conference on Artificial Intelligence and the 23rd European Conference on Artificial Intelligence", Stockholm, Sweden, July 2018, p. 1-8 [*DOI :* 10.24963/IJCAI.2018/72], https://hal.archives-ouvertes.fr/hal-01955448

[33] A. WARD, Z. ZHOU, P. MERTIKOPOULOS, N. BAMBOS.*Power Control with Random Delays: Robust Feedback Averaging*, in "CDC 2018 - 57th IEEE Conference on Decision and Control", Miami Beach, United States, December 2018, p. 1-6, https://hal.inria.fr/hal-01891528

[34] Z. ZHOU, P. MERTIKOPOULOS, S. ATHEY, N. BAMBOS, P. GLYNN, Y. YE.*Learning in Games with Lossy Feedback*, in "NIPS 2018 - Thirty-second Conference on Neural Information Processing Systems", Montreal, Canada, December 2018, p. 1-11, https://hal.inria.fr/hal-01904461

[35] Z. ZHOU, P. MERTIKOPOULOS, N. BAMBOS, P. W. GLYNN, Y. YE, L.-J. LI, F.-F. LI.*Distributed Asynchronous Optimization with Unbounded Delays: How Slow Can You Go?*, in "ICML 2018 - 35th International Conference on Machine Learning", Stockholm, Sweden, July 2018, p. 1-10, https://hal.inria.fr/hal-01891449

National Conferences with Proceeding

[36] V. G. PINTO, L. MELLO SCHNORR, A. LEGRAND, S. THIBAULT, L. STANISIC, V. DANJEAN.*Detecção de Anomalias de Desempenho em Aplicações de Alto Desempenho baseadas em Tarefas em Clusters Híbridos*, in "WPerformance 2018 - 17º Workshop em Desempenho de Sistemas Computacionais e de Comunicação", Natal, Brazil, July 2018, p. 1-14, https://hal.inria.fr/hal-01842038

Conferences without Proceedings

[37] U. OZEER, L. LETONDEUR, F.-G. OTTOGALLI, G. SALAÜN, J.-M. VINCENT.*Designing and Implementing Resilient IoT Applications in the Fog: A Smart Home Use Case*, in "22nd Conference on Innovation in Clouds, Internet and Networks", Paris, France, February 2019, https://hal.archives-ouvertes.fr/hal-01979686

Research Reports

[38] A. LEGRAND, D. TRYSTRAM, S. ZRIGUI.*Adapting Batch Scheduling to Workload Characteristics: What can we expect From Online Learning?*, Grenoble 1 UGA - Université Grenoble Alpe, October 2018, n^{o} RR-9212, p. 1-23, https://hal.inria.fr/hal-01896121

Other Publications

[39] I. BOMZE, P. MERTIKOPOULOS, W. SCHACHINGER, M. STAUDIGL.*Hessian barrier algorithms for linearly constrained optimization problems*, October 2018, https://arxiv.org/abs/1809.09449 - 27 pages, 6 figures, https://hal.inria.fr/hal-01891539

[40] P. BRUEL, S. QUINITO MASNADA, B. VIDEAU, A. LEGRAND, J.-M. VINCENT, A. GOLDMAN.*Autotuning under Tight Budget Constraints: A Transparent Design of Experiments Approach*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01953287

[41] R. COLINI-BALDESCHI, R. COMINETTI, P. MERTIKOPOULOS, M. SCARSINI.*When is selfish routing bad? The price of anarchy in light and heavy traffic*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01891558

[42] B. DONASSOLO, I. FAJJARI, A. LEGRAND, P. MERTIKOPOULOS.*FogIoT Orchestrator: an Orchestration System for IoT Applications in Fog Environment*, April 2018, p. 1-3, 1st Grid'5000-FIT school, https://hal.inria.fr/hal-01859689

[43] B. DUVOCELLE, P. MERTIKOPOULOS, M. STAUDIGL, D. VERMEULEN.*Learning in time-varying games*, October 2018, https://arxiv.org/abs/1809.03066 - 38 pages, https://hal.inria.fr/hal-01891545

[44] P. MERTIKOPOULOS, B. LECOUAT, H. ZENATI, C.-S. FOO, V. CHANDRASEKHAR, G. PILIOURAS.*Optimistic mirror descent in saddle-point problems: Going the extra (gradient) mile*, October 2018, https://arxiv.org/abs/1807.02629 - 26 pages, 14 figures, https://hal.inria.fr/hal-01891551

[45] B. VINOT, F. CADOUX, N. GAST, R. HELIOT, V. GOUIN.*Congestion Avoidance in Low-Voltage Networks by using the Advanced Metering Infrastructure*, May 2018, working paper or preprint [*DOI :* 10.1145/NNNNNNN.NNNNNNN], https://hal.inria.fr/hal-01784386

[46] D. Q. VU, P. LOISEAU, A. SILVA.*Efficient computation of approximate equilibria in discrete Colonel Blotto games*, May 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01787505

References in notes

[47] R. M. BADIA, J. LABARTA, J. GIMÉNEZ, F. ESCALÉ.*Dimemas: Predicting MPI Applications Behaviour in Grid Environments*, in "Proc. of the Workshop on Grid Applications and Programming Tools", June 2003

[48] C. BAIER, B. HAVERKORT, H. HERMANNS, J.-P. KATOEN.*Model-checking algorithms for continuous-time Markov chains*, in "Software Engineering, IEEE Transactions on", 2003, vol. 29, n^o 6, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1205180

[49] A. BASU, S. FLEMING, J. STANIER, S. NAICKEN, I. WAKEMAN, V. K. GURBANI.*The State of Peer-to-peer Network Simulators*, in "ACM Computing Survey.", August 2013, vol. 45, n^o 4

[50] D. BECKER, F. WOLF, W. FRINGS, M. GEIMER, B. WYLIE, B. MOHR.*Automatic Trace-Based Performance Analysis of Metacomputing Applications*, in "Parallel and Distributed Processing Symposium, 2007. IPDPS 2007. IEEE International", March 2007, http://dx.doi.org/10.1109/IPDPS.2007.370238

[51] P. BEDARIDE, A. DEGOMME, S. GENAUD, A. LEGRAND, G. MARKOMANOLIS, M. QUINSON, M. L. STILLWELL, F. SUTER, B. VIDEAU.*Toward Better Simulation of MPI Applications on Ethernet/TCP Networks*, in "PMBS13 - 4th International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computer Systems", Denver, United States, November 2013, https://hal.inria.fr/hal-00919507

[52] G. BIANCHI.*Performance analysis of the IEEE 802.11 distributed coordination function*, in "Selected Areas in Communications, IEEE Journal on", 2000, vol. 18, n^o 3

[53] L. BOBELIN, A. LEGRAND, M. A. G. DAVID, P. NAVARRO, M. QUINSON, F. SUTER, C. THIERY.*Scalable Multi-Purpose Network Representation for Large Scale Distributed System Simulation*, in "CCGrid 2012 – The 12th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing", Ottawa, Canada, May 2012, 19, https://hal.inria.fr/hal-00650233

[54] L. BORTOLUSSI, J. HILLSTON.*Model checking single agent behaviours by fluid approximation*, in "Information and Computation", 2015, vol. 242, http://dx.doi.org/10.1016/j.ic.2015.03.002

[55] L. BORTOLUSSI, R. LANCIANI.*Model Checking Markov Population Models by Central Limit Approximation*, in "Quantitative Evaluation of Systems", Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2013, n^o 8054

[56] L. BORTOLUSSI, R. LANCIANI.*Fluid Model Checking of Timed Properties*, in "Formal Modeling and Analysis of Timed Systems", Springer International Publishing, 2015

[57] H. BRUNST, D. HACKENBERG, G. JUCKELAND, H. ROHLING.*Comprehensive Performance Tracking with Vampir 7*, in "Tools for High Performance Computing 2009", M. S. MÜLLER, M. M. RESCH, A. SCHULZ, W. E. NAGEL (editors), Springer Berlin Heidelberg, 2010, http://dx.doi.org/10.1007/978-3-642-11261-4_2

[58] A. BUSIC, B. GAUJAL, G. GORGO, J.-M. VINCENT.*PSI2 : Envelope Perfect Sampling of Non Monotone Systems*, in "QEST 2010 - International Conference on Quantitative Evaluation of Systems", Williamsburg, VA, United States, IEEE, September 2010, p. 83-84, https://hal.inria.fr/hal-00788884

[59] A. BUSIC, B. GAUJAL, F. PERRONNIN.*Perfect Sampling of Networks with Finite and Infinite Capacity Queues*, in "19th International Conference on Analytical and Stochastic Modelling Techniques and Applications (ASMTA) 2012", Grenoble, France, K. AL-BEGAIN, D. FIEMS, J.-M. VINCENT (editors), Lecture Notes in Computer Science, Springer, 2012, vol. 7314, p. 136-149 [*DOI :* 10.1007/978-3-642-30782-9_10], https://hal.inria.fr/hal-00788003

[60] S. BÖHM, C. ENGELMANN.*xSim: The Extreme-Scale Simulator*, in "Proceedings of the International Conference on High Performance Computing and Simulation (HPCS) 2011", Istanbul, Turkey, IEEE Computer Society, Los Alamitos, CA, USA, July 2011

[61] H. CASANOVA, A. GIERSCH, A. LEGRAND, M. QUINSON, F. SUTER.*Versatile, Scalable, and Accurate Simulation of Distributed Applications and Platforms*, in "Journal of Parallel and Distributed Computing", June 2014, vol. 74, n^{o} 10, p. 2899-2917 [*DOI :* 10.1016/J.JPDC.2014.06.008], https://hal.inria.fr/hal-01017319

[62] A. CHAINTREAU, J.-Y. LE BOUDEC, N. RISTANOVIC.*The Age of Gossip: Spatial Mean Field Regime*, in "SIGMETRICS Perform. Eval. Rev.", June 2009, vol. 37, n^{o} 1, http://doi.acm.org/10.1145/2492101.1555363

[63] K. COULOMB, M. FAVERGE, J. JAZEIX, O. LAGRASSE, J. MARCOUEILLE, P. NOISETTE, A. REDONDY, C. VUCHENER.*Visual trace explorer (ViTE)*, October, 2009

[64] J. DONCEL, N. GAST, B. GAUJAL.*Mean-Field Games with Explicit Interactions*, February 2016, https://hal.inria.fr/hal-01277098

[65] S. DURAND, B. GAUJAL, F. PERRONNIN, J.-M. VINCENT.*A perfect sampling algorithm of random walks with forbidden arcs*, in "QEST 2014 - 11th International Conference on Quantitative Evaluation of Systems", Florence, Italy, Springer, September 2014, vol. 8657, p. 178-193 [*DOI :* 10.1007/978-3-319-10696-0_15], https://hal.inria.fr/hal-01069975

[66] C. FRICKER, N. GAST.*Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity*, in "EURO Journal on Transportation and Logistics", June 2014, 31 [*DOI :* 10.1007/S13676-014-0053-5], https://hal.inria.fr/hal-01086009

[67] C. FRICKER, N. GAST, H. MOHAMED.*Mean field analysis for inhomogeneous bike sharing systems*, in "AofA", Montreal, Canada, July 2012, https://hal.inria.fr/hal-01086055

[68] D. FUDENBERG, D. K. LEVINE.*The Theory of Learning in Games*, Economic learning and social evolution, MIT Press, Cambridge, MA, 1998, vol. 2

[69] R. M. FUJIMOTO.*Parallel Discrete Event Simulation*, in "Commun. ACM", October 1990, vol. 33, n^{o} 10, http://doi.acm.org/10.1145/84537.84545

[70] N. GAST, B. GAUJAL.*Markov chains with discontinuous drifts have differential inclusion limits*, in "Performance Evaluation", 2012, vol. 69, n^{o} 12, p. 623-642 [*DOI :* 10.1016/J.PEVA.2012.07.003], https://hal.inria.fr/hal-00787999

[71] N. GAST, B. GAUJAL, J.-Y. LE BOUDEC.*Mean field for Markov Decision Processes: from Discrete to Continuous Optimization*, in "IEEE Transactions on Automatic Control", 2012, vol. 57, n^{o} 9, p. 2266 - 2280 [*DOI :* 10.1109/TAC.2012.2186176], https://hal.inria.fr/hal-00787996

[72] N. GAST, J.-Y. LE BOUDEC, D.-C. TOMOZEI.*Impact of Demand-Response on the Efficiency and Prices in Real-Time Electricity Markets*, in "ACM e-Energy 2014", Cambridge, United Kingdom, June 2014 [*DOI :* 10.1145/2602044.2602052], https://hal.inria.fr/hal-01086036

[73] N. GAST, B. VAN HOUDT.*Transient and Steady-state Regime of a Family of List-based Cache Replacement Algorithms*, in "ACM SIGMETRICS 2015", Portland, United States, June 2015 [*DOI :* 10.1145/2745844.2745850], https://hal.inria.fr/hal-01143838

[74] J. GONZÁLEZ, J. GIMÉNEZ, J. LABARTA.*Automatic detection of parallel applications computation phases*, in "Parallel and Distributed Processing Symposium, International", 2009, vol. 0, http://doi.ieeecomputersociety.org/10.1109/IPDPS.2009.5161027

[75] M. HEATH, J. ETHERIDGE.*Visualizing the performance of parallel programs*, in "IEEE software", 1991, vol. 8, n^{o} 5

[76] T. HOEFLER, T. SCHNEIDER, A. LUMSDAINE.*LogGOPSim - Simulating Large-Scale Applications in the LogGOPS Model*, in "Proc. of the ACM Workshop on Large-Scale System and Application Performance", June 2010

[77] L. HU, J.-Y. LE BOUDEC, M. VOJNOVIĆ.*Optimal channel choice for collaborative ad-hoc dissemination*, in "INFOCOM, 2010 Proceedings IEEE", IEEE, 2010

[78] L. V. KALÉ, G. ZHENG, C. W. LEE, S. KUMAR.*Scaling applications to massively parallel machines using Projections performance analysis tool*, in "Future Generation Comp. Syst.", 2006, vol. 22, n^{o} 3

[79] T. G. KURTZ.*Approximation of population processes*, SIAM, 1981, vol. 36

[80] Y.-B. LIN, E. D. LAZOWSKA.*A Time-division Algorithm for Parallel Simulation*, in "ACM Trans. Model. Comput. Simul.", January 1991, vol. 1, n^{o} 1, http://doi.acm.org/10.1145/102810.214307

[81] G. LLORT, J. GONZÁLEZ, H. SERVAT, J. GIMÉNEZ, J. LABARTA.*On-line Detection of Large-scale Parallel Application's Structure*, in "24th IEEE International Parallel and Distributed Processing Symposium (IPDPS'2010)", 2010

[82] L. MELLO SCHNORR, A. LEGRAND.*Visualizing More Performance Data Than What Fits on Your Screen*, in "Tools for High Performance Computing 2012", A. CHEPTSOV, S. BRINKMANN, J. GRACIA, M. M. RESCH,

W. E. NAGEL (editors), Springer Berlin Heidelberg, 2013, p. 149-162 [*DOI :* 10.1007/978-3-642-37349-7_10], https://hal.inria.fr/hal-00842761

[83] S. MEYN, P. BAROOAH, A. BUSIC, J. EHREN.*Ancillary service to the grid from deferrable loads: the case for intelligent pool pumps in Florida*, in "Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on", IEEE, 2013

[84] M. MITZENMACHER.*The power of two choices in randomized load balancing*, in "Parallel and Distributed Systems, IEEE Transactions on", 2001, vol. 12, n^o 10

[85] K. MOHROR, K. KARAVANIC, A. SNAVELY.*Scalable Event Trace Visualization*, in "Euro-Par 2009 – Parallel Processing Workshops", H.-X. LIN, M. ALEXANDER, M. FORSELL, A. KNÜPFER, R. PRODAN, L. SOUSA, A. STREIT (editors), Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2010, vol. 6043, http://dx.doi.org/10.1007/978-3-642-14122-5_27

[86] W. NAGEL, A. ARNOLD, M. WEBER, H. HOPPE, K. SOLCHENBACH.*VAMPIR: Visualization and Analysis of MPI Resources*, in "Supercomputer", 1996, vol. 12, n^o 1

[87] V. PILLET, J. LABARTA, T. CORTES, S. GIRONA.*PARAVER: A tool to visualise and analyze parallel code*, in "Proceedings of Transputer and occam Developments, WOTUG-18.", Transputer and Occam Engineering, IOS Press, 1995, vol. 44

[88] J. PROPP, D. WILSON.*Coupling from the past: a user's guide*, in "DIMACS Series on Discrete Mathematics and Theoretical Computer Science", 1998, vol. 41, Microsurveys in discrete probability

[89] M. L. PUTERMAN.*Markov decision processes: discrete stochastic dynamic programming*, John Wiley & Sons, 2014

[90] D. REED, P. ROTH, R. AYDT, K. SHIELDS, L. TAVERA, R. NOE, B. SCHWARTZ.*Scalable performance analysis: the Pablo performance analysis environment*, in "Scalable Parallel Libraries Conference, 1993., Proceedings of the", 1993

[91] W. H. SANDHOLM.*Population Games and Evolutionary Dynamics*, Economic learning and social evolution, MIT Press, Cambridge, MA, 2010

[92] W. H. SANDHOLM, M. STAUDIGL.*A Sample Path Large Deviation Principle for a Class of Population Processes*, in "arXiv preprint arXiv:1511.07897", 2015

[93] H. SERVAT, G. LLORT, J. GIMÉNEZ, K. HUCK, J. LABARTA.*Folding: detailed analysis with coarse sampling*, in "Tools for High Performance Computing 2011", Springer Berlin Heidelberg, 2012

[94] H. SERVAT, G. LLORT, J. GONZALEZ, J. GIMENEZ, J. LABARTA.*Identifying code phases using piece-wise linear regressions*, in "Parallel and Distributed Processing Symposium, 2014 IEEE 28th International", IEEE, 2014

[95] B. SHNEIDERMAN.*The eyes have it: A task by data type taxonomy for information visualizations*, in "Visual Languages, 1996. Proceedings., IEEE Symposium on", IEEE, 1996

[96] M. TIKIR, M. LAURENZANO, L. CARRINGTON, A. SNAVELY.*PSINS: An Open Source Event Tracer and Execution Simulator for MPI Applications*, in "Proc. of the 15th International Euro-Par Conference on Parallel Processing", LNCS, Springer, August 2009, n^{o} 5704

[97] B. VAN HOUDT.*A Mean Field Model for a Class of Garbage Collection Algorithms in Flash-based Solid State Drives*, in "Proceedings of the ACM SIGMETRICS", New York, NY, USA, SIGMETRICS '13, ACM, 2013, http://doi.acm.org/10.1145/2465529.2465543

[98] P. VELHO, L. SCHNORR, H. CASANOVA, A. LEGRAND.*On the Validity of Flow-level TCP Network Models for Grid and Cloud Simulations*, in "ACM Transactions on Modeling and Computer Simulation", October 2013, vol. 23, n^{o} 4, https://hal.inria.fr/hal-00872476

[99] J. J. WILKE, K. SARGSYAN, J. P. KENNY, B. DEBUSSCHERE, H. N. NAJM, G. HENDRY.*Validation and Uncertainty Assessment of Extreme-Scale HPC Simulation through Bayesian Inference*, in "Euro-Par 2013 Parallel Processing: 19th International Conference, Aachen, Germany, August 26-30, 2013. Proceedings", Springer Berlin Heidelberg, Berlin, Heidelberg, 2013

[100] F. WOLF, B. MOHR.*Automatic performance analysis of hybrid MPI/OpenMP applications*, in "Journal of Systems Architecture", 2003, vol. 49, n^{o} 10-11

[101] T. YANG, P. G. MEHTA, S. P. MEYN.*A mean-field control-oriented approach to particle filtering*, in "American Control Conference (ACC), 2011", IEEE, 2011

[102] L. YING.*On the Rate of Convergence of Mean-Field Models: Stein's Method Meets the Perturbation Theory*, in "arXiv preprint arXiv:1510.00761", 2015

[103] O. ZAKI, E. LUSK, W. GROPP, D. SWIDER.*Toward Scalable Performance Visualization with Jumpshot*, in "International Journal of High Performance Computing Applications", 1999, vol. 13, n^{o} 3, http://dx.doi.org/10.1177/109434209901300310

[104] G. ZHENG, G. KAKULAPATI, L. KALÉ.*BigSim: A Parallel Simulator for Performance Prediction of Extremely Large Parallel Machines*, in "Proc. of the 18th International Parallel and Distributed Processing Symposium (IPDPS)", April 2004

[105] J. C. DE KERGOMMEAUX, B. STEIN, P. BERNARD.*Paje, an interactive visualization tool for tuning multi-threaded parallel applications*, in "Parallel Computing", 2000, vol. 10, n^{o} 26, p. 1253–1274

Project-Team PRIVATICS

Privacy Models, Architectures and Tools for the Information Society

IN COLLABORATION WITH: Centre of Innovation in Telecommunications and Integration of services

IN PARTNERSHIP WITH:
Institut national des sciences appliquées de Lyon

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Security and Confidentiality

Table of contents

Project-Team PRIVATICS

Creation of the Team: 2013 January 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A1. - Architectures, systems and networks
A3. - Data and knowledge
A4. - Security and privacy
A9. - Artificial intelligence

Other Research Topics and Application Domains:

B2. - Health
B6. - IT and telecom
B8. - Smart Cities and Territories
B9. - Society and Knowledge

1. Team, Visitors, External Collaborators

Research Scientists

Claude Castelluccia [Team leader, Inria, Senior Researcher, HDR]
Cédric Lauradoux [Inria, Researcher]
Daniel Le Metayer [Inria, Senior Researcher, HDR]
Vincent Roca [Inria, Researcher, HDR]

Faculty Members

Antoine Boutet [INSA Lyon, Associate Professor]
Mathieu Cunche [INSA Lyon, Associate Professor]

Technical Staff

Gabor Gulyas [Inria, until Jun 2018]
Belkacem Teibi [Inria, until Sep 2018]

PhD Students

Supriya Sreekant Adhatarao [Inria, from Mar 2018]
Coline Boniface [Univ Grenoble Alpes, from Sep 2018]
Guillaume Celosia [INSA Lyon]
Theo Jourdan [INSERM, from Sep 2018]
Raouf Kerkouche [Univ Grenoble Alpes]
Victor Morel [Inria]
Mathieu Thiery [Inria]
Clement Henin [MTES-ENPC, from Sep 2018]

Post-Doctoral Fellow

Raul Pardo Jimenez [Inria, from Feb 2018]

Visiting Scientist

Gergely Acs [Budapest University of Technology and Economics, from May 2018 until Jul 2018]

Administrative Assistant

Helen Pouchot-Rouge-Blanc [Inria]

2. Overall Objectives

2.1. Context

The promises of new technologies: Many advances in new technologies are very beneficial to the society and provide services that can drastically improve life's quality. A good example is the emergence of reality mining. Reality mining is a new discipline that infers human relationships and behaviors from information collected by cell-phones. Collected information include data collected by the sensors, such as location or physical activities, as well as data recorded by the phones themselves, such as call duration and dialed numbers. Reality mining could be used by individuals to get information about themselves, their state or performances ("quantified self"). More importantly, it could help monitoring health. For example, the motions of a mobile phone might reveal changes in gait, which could be an early indicator of ailments or depression. The emergence of location-based or mobile/wireless services is also often very beneficial. These systems provide very useful and appreciated services, and become almost essential and inevitable nowadays. For example, RFID cards allow users to open doors or pay their metro tickets. GPS systems help users to navigate and find their ways. Some services tell users where their friends are or provide services personalized to their current location (such as indicating the closest restaurant or hotel). Some wireless parking meters send users a text message when their time is running out. The development of smart grids, smart houses, or more generally smart spaces/environments, can also positively contribute to the well-being of the society. Smart-grids and smart houses attempt to minimize energy consumption by monitoring users' energy consumptions and applying adequate actions. These technologies can help reducing pollution and managing energy resources.

Privacy threats of new technologies: While the potential benefits provided by these systems are numerous, they also pose considerable privacy threats that can potentially turn new technologies into a nightmare. Most of these systems leave digital traces that can potentially be used to profile or monitor users. Content on the Internet (documents, emails, chats, images, videos etc) is often disseminated and replicated on different peers or servers. As a result, users lose the control of their content as soon as they release it. Furthermore most users are unaware of the information that is collected about them beyond requested data. It was shown that consumption data provided by smart meters to electricity providers is so accurate that it can be used to infer physical activities (e.g. when the house occupant took a shower or switched-on TV). Also, a picture taken by a user may reveal additional contextual information inferred from the background or the style of any associated text. For example, photos and videos taken with smart phones or cameras contain geo-location information. This may be considered as a potential source of information leakage and may lead to a privacy breach if used for location tracking or in conjunction with data retrieved from OSN (Online Social Networks). The risk becomes higher as the border between OSN and LBS (Location Based Services) becomes fuzzier. For instance, OSN such as FourSquare and Gowalla are designed to encourage users to share their geolocated data. Information posted on social applications such as Twitter can be used to infer whether or not an individual is at home. Other applications, such as Google Latitude, allow users to track the movements of their friends' cellphones and display their position on a map. In addition to social applications, there are other public sources of information that can be exploited by potential adversaries, such as the free geographic data provided by Google Maps, Yahoo! Maps and Google Earth. The danger is to move into a surveillance society where all our online and physical activities are recorded and correlated. Some companies already offer various services that gather different types of information from users. The combination and concentration of all these information provide a powerful tool to accurately profile users. For example, Google is one of the main third-party aggregators and tracks users across most web sites. In addition, it also runs the most popular search engine and, as such, stores web histories of most users (i.e. their search requests), their map searches (i.e. their requests to the Google Map service), their images and so on [8]. Web searches have been shown to often be sensitive. Furthermore, Google is also going into the mobile and energy business, which will potentially allow it to correlate online profile with physical profiles.

The "Internet of the future" should solve these privacy problems. However, privacy is not something that occurs naturally online, it must be deliberately designed. This architecture of Privacy must be updated and reconsidered as the concept of privacy evolves and new technologies appear.

Even if our main goal is to develop general techniques with a potentially broad impact, Privatics will consider different and various concrete case studies to ensure the relevance and significance of its results. We plan to work on several case studies related to the Internet, online social networks (OSN), mobile services and smart spaces/environments (such as smart grids, smart houses,..), which correspond to challenging application domains with great impact on society.

3. Application Domains

3.1. Domain 1: Privacy in smart environments

Privacy in smart environments. One illustrative example is our latest work on privacy-preserving smart-metering [2]. Several countries throughout the world are planning to deploy smart meters in house-holds in the very near future. Traditional electrical meters only measure total consumption on a given period of time (i.e., one month or one year). As such, they do not provide accurate information of when the energy was consumed. Smart meters, instead, monitor and report consumption in intervals of few minutes. They allow the utility provider to monitor, almost in real-time, consumption and possibly adjust generation and prices according to the demand. Billing customers by how much is consumed and at what time of day will probably change consumption habits to help matching energy consumption with production. In the longer term, with the advent of smart appliances, it is expected that the smart grid will remotely control selected appliances to reduce demand. Although smart metering might help improving energy management, it creates many new privacy problems. Smart-meters provide very accurate consumption data to electricity providers. As the interval of data collected by smart meters decreases, the ability to disaggregate low-resolution data increases. Analysing high-resolution consumption data, Non-intrusive Appliance Load Monitoring (NALM) can be used to identify a remarkable number of electric appliances (e.g., water heaters, well pumps, furnace blowers, refrigerators, and air conditioners) employing exhaustive appliance signature libraries. We developed DREAM, DiffeRentially privatE smArt Metering, a scheme that is private under the differential privacy model and therefore provides strong and provable guarantees. With our scheme, an (electricity) supplier can periodically collect data from smart-meters and derive aggregated statistics while learning only limited information about the activities of individual households. For example, a supplier cannot tell from a user's trace when he watched TV or turned on heating.

3.2. Domain 2: Big Data and Privacy

We believe that another important problem will be related to privacy issues in big data. Public datasets are used in a variety of applications spanning from genome and web usage analysis to location-based and recommendation systems. Publishing such datasets is important since they can help us analyzing and understanding interesting patterns. For example, mobility trajectories have become widely collected in recent years and have opened the possibility to improve our understanding of large-scale social networks by investigating how people exchange information, interact, and develop social interactions. With billion of handsets in use worldwide, the quantity of mobility data is gigantic. When aggregated, they can help understand complex processes, such as the spread of viruses, and build better transportation systems. While the benefits provided by these datasets are indisputable, they unfortunately pose a considerable threat to individual privacy. In fact, mobility trajectories might be used by a malicious attacker to discover potential sensitive information about a user, such as his habits, religion or relationships. Because privacy is so important to people, companies and researchers are reluctant to publish datasets by fear of being held responsible for potential privacy breaches. As a result, only very few of them are actually released and available. This limits our ability to analyze such data to derive information that could benefit the general public. It is now an urgent need to develop Privacy-Preserving Data Analytics (PPDA) systems that collect and transform raw data into a version that is immunized against privacy attacks but that still preserves useful information for data analysis. This is one of the objectives of Privatics. There exists two classes of PPDA according to whether the entity that is collecting and anonymizing the data is trusted or not. In the trusted model, that we refer to as Privacy-Preserving Data Publishing (PPDP), individuals trust the publisher to which they disclose their data. In the untrusted model, that we refer to as Privacy-Preserving Data Collection (PPDC), individuals do not trust the data publisher. They may add some noise to their data to protect sensitive information from the data publisher.

Privacy-Preserving Data Publishing: In the trusted model, individuals trust the data publisher and disclose all their data to it. For example, in a medical scenario, patients give their true information to hospitals to receive proper treatment. It is then the responsibility of the data publisher to protect privacy of the individuals' personal data. To prevent potential data leakage, datasets must be sanitized before possible release. Several proposals have been recently proposed to release private data under the Differential Privacy model [25, 56, 26, 57, 50]. However most of these schemes release a "snapshot" of the datasets at a given period of time. This release often consists of histograms. They can, for example, show the distributions of some pathologies (such as cancer, flu, HIV, hepatitis, etc.) in a given population. For many analytics applications, "snapshots" of data are not enough, and sequential data are required. Furthermore, current work focusses on rather simple data structures, such as numerical data. Release of more complex data, such as graphs, are often also very useful. For example, recommendation systems need the sequences of visited websites or bought items. They also need to analyse people connection graphs to identify the best products to recommend. Network trace analytics also rely on sequences of events to detect anomalies or intrusions. Similarly, traffic analytics applications typically need sequences of visited places of each user. In fact, it is often essential for these applications to know that user A moved from position 1 to position 2, or at least to learn the probability of a move from position 1 to position 2. Histograms would typically represent the number of users in position 1 and position 2, but would not provide the number of users that moved from position 1 to position 2. Due to the inherent sequentiality and high-dimensionality of sequential data, one major challenge of applying current data sanitization solutions on sequential data comes from the uniqueness of sequences (e.g., very few sequences are identical). This fact makes existing techniques result in poor utility. Schemes to privately release data with complex data structures, such as sequential, relational and graph data, are required. This is one the goals of Privatics. In our current work, we address this challenge by employing a variable-length n-gram model, which extracts the essential information of a sequential database in terms of a set of variable-length n - grams [15]. We then intend to extend this approach to more complex data structures.

Privacy-Preserving Data Collection: In the untrusted model, individuals do not trust their data publisher. For example, websites commonly use third party web analytics services, such as Google Analytics to obtain aggregate traffic statistics such as most visited pages, visitors' countries, etc. Similarly, other applications, such as smart metering or targeted advertising applications, are also tracking users in order to derive aggregated information about a particular class of users. Unfortunately, to obtain this aggregate information, services need to track users, resulting in a violation of user privacy. One of our goals is to develop Privacy-Preserving Data Collection solutions. We propose to study whether it is possible to provide efficient collection/aggregation solutions without tracking users, i.e. without getting or learning individual contributions.

4. Highlights of the Year

4.1. Highlights of the Year

Cédric Lauradoux, Vincent Roca with the participation of Claude Castelluccia have created a MOOC on Privacy which has been followed this year by more than 20000 persons.

5. New Software and Platforms

5.1. FECFRAME

FEC Framework following RFC 6363 specifications (https://datatracker.ietf.org/doc/rfc6363/)
KEYWORDS: Error Correction Code - Content delivery protocol - Robust transmission

FUNCTIONAL DESCRIPTION: This sofware implements the FECFRAME IETF standard (RFC 6363) co-authored by V. Roca, and is compliant with 3GPP specifications for mobile terminals. It enables the simultaneous transmission of multimedia flows to one or several destinations, while being robust to packet erasures that happen on wireless networks (e.g., 4G or Wifi). This software relies on the OpenFEC library (the open-source http://openfec.org version or the commercial version) that provides the erasure correction codes (or FEC) and thereby offer robustness in front of packet erasures.

- Participant: Vincent Roca
- Contact: Vincent Roca

5.2. Mobilitcs

FUNCTIONAL DESCRIPTION: Mobilitics is a joint project, started in 2012 between Inria and CNIL, which targets privacy issues on smartphones. The goal is to analyze the behavior of smartphones applications and their operating system regarding users private data, that is, the time they are accessed or sent to third party companies usually neither with user's awareness nor consent.

In the presence of a wide range of different smartphones available in terms of operating systems and hardware architecture, Mobilitics project focuses actually its study on the two mostly used mobile platforms, IOS (Iphone) and Android. Both versions of the Mobilitics software: (1) capture any access to private data, any modification (e.g., ciphering or hashing of private data), or transmission of data to remote locations on the Internet, (2) store these events in a local database on the phone for offline analysis, and (3) provide the ability to perform an in depth database analysis in order to identify personnal information leakage.

- Authors: Jagdish Achara, James-Douglass Lefruit, Claude Castelluccia, Franck Baudot, Geoffrey Delcroix, Gwendal Le Grand, Stéphane Petitcolas and Vincent Roca
- Contact: Claude Castelluccia

5.3. MyTrackingChoices

KEYWORDS: Privacy - User control
FUNCTIONAL DESCRIPTION: This extension lets you control how you are being tracked on the Internet. It allows you to choose the categories (e.g., health, adult) of the websites where you don't want to be tracked on. When you browse the web, your visited webpages will be categorized on the fly and, depending on your choices, the extension will block the trackers (webpage by webpage) or not.

Existing anti-tracking (Ghostery, Disconnect etc.) and ad-blocking (AdBlock Plus etc.) tools block almost ALL trackers and as a result, ads. This has a negative impact on the Internet economy because free services/content on the Internet are fuelled by ads. As a result, websites are starting to block access to their content if they detect use of Ad-blockers or they ask users to move to a subscription-based model (where users have to pay to get access to the website).

This extension is testing another approach: It is trying to find a trade-off between privacy and economy, that would allow users to protect their privacy while still accessing to free content.

It is based on the assumption that most people are not against advertisements, but want to keep control over their data. We believe that some sites are more sensitive than others. In fact, most people don't want to be tracked on "sensitive" websites (for example related to religion, health,...), but don't see any problem to be tracked on less sensitive ones (such as news, sport,...). This extension allows you to take control and specify which on which categories of sites you don't want to be tracked on! Furthermore, the extension also gives you the option to block the trackers on specific websites.

- Contact: Claude Castelluccia
- URL: https://addons.mozilla.org/FR/firefox/addon/mytrackingchoices/

5.4. OMEN+

FUNCTIONAL DESCRIPTION: Omen+ is a password cracker following our previous work. It is used to guess possible passwords based on specific information about the target. It can also be used to check the strength of user password by effectively looking at the similarity of that password with both usual structures and information relative to the user, such as his name, birth date...

It is based on a Markov analysis of known passwords to build guesses. The previous work Omen needs to be cleaned in order to be scaled to real problems and to be distributed or transfered to the security community (maintainability): eventually it will become an open source software. The main challenge of Omen+ is to optimize the memory consumption.

- Participants: Claude Castelluccia and Pierre Rouveyrol
- Contact: Claude Castelluccia

5.5. OPENFEC

KEYWORD: Error Correction Code
FUNCTIONAL DESCRIPTION: OpenFEC is a C-language implementation of several Application-Level Forward Erasure Correction (AL-FEC) codecs, namely: Reed-Solomon (RFC 5510), LDPC-Staircase (RFC 5170) codes, and RLC (https://datatracker.ietf.org/doc/draft-ietf-tsvwg-rlc-fec-scheme/). Two versions are available: an open-source, unsupported version (http://openfec.org), and an advanced version commercialized by the Expway SME.

RELEASE FUNCTIONAL DESCRIPTION: Added support of RLC codes (Random Linear Codes), based on a sliding encoding window.

- Participants: Christophe Neumann, Belkacem Teibi, Jérome Lacan, Jonathan Detchart, Julien Laboure, Kevin Chaumont, Mathieu Cunche and Vincent Roca
- Partner: Expway
- Contact: Vincent Roca
- URL: http://openfec.org/

6. New Results

6.1. Fine-Grained Control over Tracking to Support the Ad-Based Web Economy

Participant: Claude Castelluccia.

The intrusiveness of Web tracking and the increasing invasiveness of digital advertising have raised serious concerns regarding user privacy and Web usability, leading a substantial chunk of the populace to adopt ad-blocking technologies in recent years. The problem with these technologies, however, is that they are extremely limited and radical in their approach, and they completely disregard the underlying economic model of the Web, in which users get content free in return for allowing advertisers to show them ads. Nowadays, with around 200 million people regularly using such tools, said economic model is in danger. In this article, we investigate an Internet technology that targets users who are not, in general, against advertising, accept the trade-off that comes with the “free” content, but—for privacy concerns—they wish to exert fine-grained control over tracking. Our working assumption is that some categories of web pages (e.g., related to health or religion) are more privacy-sensitive to users than others (e.g., about education or science). Capitalizing on this, we propose a technology that allows users to specify the categories of web pages that are privacy-sensitive to them and block the trackers present on such web pages only. As tracking is prevented by blocking network connections of third-party domains, we avoid not only tracking but also third-party ads. Since users continue receiving ads on those web pages that belong to non-sensitive categories, our approach may provide a better point of operation within the trade-off between user privacy and the Web economy. To test the appropriateness and feasibility of our solution, we implemented it as a Web-browser plug-in, which is currently available for Google Chrome and Mozilla Firefox. Experimental results from the collected data of 746 users during one year show that only 16.25% of ads are blocked by our tool, which seems to indicate that the economic impact of the ad-blocking exerted by privacy-sensitive users could be significantly reduced.

6.2. Differentially Private Mixture of Generative Neural Networks

Participant: Claude Castelluccia.

Generative models are used in a wide range of applications building on large amounts of contextually rich information. Due to possible privacy violations of the individuals whose data is used to train these models, however, publishing or sharing generative models is not always viable. In this paper, we present a novel technique for privately releasing generative models and entire high-dimensional datasets produced by these models. We model the generator distribution of the training data with a mixture of k generative neural networks. These are trained together and collectively learn the generator distribution of a dataset. Data is divided into k clusters, using a novel differentially private kernel k-means, then each cluster is given to separate generative neural networks, such as Restricted Boltzmann Machines or Variational Autoencoders, which are trained only on their own cluster using differentially private gradient descent. We evaluate our approach using the MNIST dataset, as well as call detail records and transit datasets, showing that it produces realistic synthetic samples, which can also be used to accurately compute arbitrary number of counting queries.

6.3. On the Cost-Effectiveness of Mass Surveillance

Participant: Claude Castelluccia.

In recent times, we have witnessed an increasing concern by governments and intelligence agencies to deploy mass-surveillance systems that help them fight terrorism. Although a government may be perfectly legitimate to do so, it is questionable whether a preventive-surveillance state is rational and cost-effective. In this paper, we conduct a theoretical analysis of the cost of such surveillance systems. Our analysis starts with a fairly well-known result in statistics, namely, the false-positive paradox. We propose a quantitative measure of the total cost of a monitoring program, and study a detection system that is designed to minimize it, subject to a constraint in the percentage of terrorists the agency wishes to capture. Our formulation is first illustrated by means of several simple albeit insightful examples of terrorist and innocent profiles. Then, we conduct an extensive experimental study from real-world socio-demographic data of jihadist terrorism in the U.K. and Spain, and provide insight into the rationality and cost-effectiveness of two countries with two of the biggest defense budgets in the world.

6.4. To Extend or not to Extend: on the Uniqueness of Browser Extensions and Web Logins

Participants: Claude Castelluccia, Gabor Gulyas.

Recent works showed that websites can detect browser extensions that users install and websites they are logged into. This poses significant privacy risks, since extensions and Web logins that re ect user's behavior, can be used to uniquely identify users on the Web. This paper reports on the rst large-scale behavioral uniqueness study based on 16,393 users who visited our website. We test and detect the presence of 16,743 Chrome extensions, covering 28% of all free Chrome extensions. We also detect whether the user is connected to 60 different websites. We analyze how unique users are based on their behavior, and nd out that 54.86% of users that have installed at least one detectable extension are unique; 19.53% of users are unique among those who have logged into one or more detectable websites; and 89.23% are unique among users with at least one extension and one login. We use an advanced ngerprinting algorithm and show that it is possible to identify a user in less than 625 milliseconds by selecting the most unique combinations of extensions. Because privacy extensions contribute to the uniqueness of users, we study the trade-o between the amount of trackers blocked by such extensions and how unique the users of these extensions are. We have found that privacy extensions should be considered more useful than harmful. The paper concludes with possible counter- measures.

6.5. Privacy-Preserving Release of Spatio-Temporal Density

Participants: Claude Castelluccia, Gergely Acs.

In today's digital society, increasing amounts of contextually rich spatio-temporal information are collected and used, e.g., for knowledge-based decision making, research purposes, optimizing operational phases of city management, planning infrastructure networks, or developing timetables for public transportation with an increasingly autonomous vehicle fleet. At the same time, however, publishing or sharing spatio-temporal data, even in aggregated form, is not always viable owing to the danger of violating individuals' privacy, along with the related legal and ethical repercussions. In this chapter, we review some fundamental approaches for anonymizing and releasing spatio-temporal density, i.e., the number of individuals visiting a given set of locations as a function of time. These approaches follow different privacy models providing different privacy guarantees as well as accuracy of the released anonymized data. We demonstrate some sanitization (anonymization) techniques with provable privacy guarantees by releasing the spatio-temporal density of Paris, in France. We conclude that, in order to achieve meaningful accuracy, the sanitization process has to be carefully customized to the application and public characteristics of the spatio-temporal data.

6.6. Algorithmic Decision Systems in the Health and Justice Sectors: Certification and Explanations for Algorithms in European and French Law

Participant: Daniel Le Metayer.

Algorithmic decision systems are already used in many everyday tools and services on the Internet, and they also play an increasing role in many situations in which people's lives and rights are strongly affected, such as job and loans applications, but also medical diagnosis and therapeutic choices, or legal advice and court decisions. This evolution gives rise to a whole range of questions. In this paper, we argue that certification and explanation are two complementary means of strengthening the European legal framework and enhancing trust in algorithmic decision systems. The former can be seen as the delegation of the task of checking certain criteria to an authority, while the latter allows the stakeholders themselves (for example, developers, users and decision-subjects) to understand the results or the logic of the system. We explore potential legal requirements of accountability in this sense and their effective implementation. These two aspects are tackled from the perspective of the European and French legal frameworks. We focus on two particularly sensitive application domains, namely the medical and legal sectors.

6.7. Capacity: an Abstract Model of Control over Personal Data

Participant: Daniel Le Metayer.

While the control of individuals over their personal data is increasingly seen as an essential component of their privacy, the word "control" is usually used in a very vague way, both by lawyers and by computer scientists. This lack of precision may lead to misunderstandings and makes it difficult to check compliance. To address this issue, we propose a formal framework based on capacities to specify the notion of control over personal data and to reason about control properties. We illustrate our framework with social network systems and show that it makes it possible to characterize the types of control over personal data that they provide to their users and to compare them in a rigorous way.

6.8. Biometric Systems Private by Design: Reasoning about privacy properties of biometric system architectures

Participant: Daniel Le Metayer.

In is to show the applicability of the privacy by design approach to biometric systems and the benefit of using formal methods to this end. We build on a general framework for the definition and verification of privacy architectures introduced at STM 2014 and show how it can be adapted to biometrics. The choice of particular techniques and the role of the components (central server, secure module, biometric terminal, smart card, etc.) in the architecture have a strong impact on the privacy guarantees provided by a biometric system. Some architectures have already been analysed but on a case by case basis, which makes it difficult to draw comparisons and to provide a rationale for the choice of specific options. In this paper, we describe the application of a general privacy architecture framework to specify different design options for biometric systems and to reason about them in a formal way.

6.9. Privacy Risk Analysis to Enable Informed Privacy Settings

Participant: Daniel Le Metayer.

is a contribution to enhancing individual control over personal data which is promoted, inter alia, by the new EU General Data Protection Regulation. We propose a method to enable better informed choices of privacy preferences or privacy settings. The method relies on a privacy risk analysis framework parameterized with privacy settings. The user can express his choices, visualize their impact on the privacy risks through a user-friendly interface, and decide to revise them as necessary to reduce risks to an acceptable level.

6.10. Enhancing Transparency and Consent in the IoT

Participants: Daniel Le Metayer, Claude Castelluccia, Mathieu Cunche, Victor Morel.

The development of the IoT raises specific questions in terms of privacy, especially with respect to information to users and consent. We argue that (1) all necessary information about collected data and the collecting devices should be communicated electronically to all data subjects in their range and (2) data subjects should be able to reply also electronically and express their own privacy choices. In this position paper, we take some examples of technologies and initiatives to illustrate our position (including direct and registry-based communications) and discuss them in the light of the GDPR and the WP29 recommendations.

6.11. Toward privacy in IoT mobile devices for activity recognition

Participant: Antoine Boutet.

Recent advances in wireless sensors for personal healthcare allow to recognise human real-time activities with mobile devices. While the analysis of those datastream can have many benefits from a health point of view, it can also lead to privacy threats by exposing highly sensitive information. In this work, we propose a privacy-preserving framework for activity recognition. This framework relies on a machine learning technique to efficiently recognise the user activity pattern, useful for personal healthcare monitoring, while limiting the risk of re-identification of users from biometric patterns that characterizes each individual. To achieve that, we first deeply analysed different features extraction schemes in both temporal and frequency domain. We show that features in temporal domain are useful to discriminate user activity while features in frequency domain lead to distinguish the user identity. On the basis of this observation, we second design a novel protection mechanism that processes the raw signal on the user's smartphone and transfers to the application server only the relevant features unlinked to the identity of users. In addition, a generalisation-based approach is also applied on features in frequency domain before to be transmitted to the server in order to limit the risk of re-identification. We extensively evaluate our framework with a reference dataset: results show an accurate activity recognition (87%) while limiting the re-identification rate (33%). This represents a slightly decrease of utility (9%) against a large privacy improvement (53%) compared to state-of-the-art baselines, while reducing the computational cost on the application server.

6.12. The Long Road to Computational Location Privacy: A Survey

Participant: Antoine Boutet.

The widespread adoption of continuously connected smartphones and tablets developed the usage of mobile applications, among which many use location to provide geolocated services. These services provide new prospects for users: getting directions to work in the morning, leaving a check-in at a restaurant at noon and checking next day's weather in the evening are possible right from any mobile device embedding a GPS chip. In these location-based applications, the user's location is sent to a server, which uses them to provide contextual and personalised answers. However, nothing prevents the latter from gathering, analysing and possibly sharing the collected information, which opens the door to many privacy threats. Indeed, mobility data can reveal sensitive information about users, among which one's home, work place or even religious and political preferences. For this reason, many privacy-preserving mechanisms have been proposed these last years to enhance location privacy while using geolocated services. This work surveys and organises contributions in this area from classical building blocks to the most recent developments of privacy threats and location privacy-preserving mechanisms. We divide the protection mechanisms between online and offline use cases, and organise them into six categories depending on the nature of their algorithm. Moreover, this work surveys the evaluation metrics used to assess protection mechanisms in terms of privacy, utility and performance. Finally, open challenges and new directions to address the problem of computational location privacy are pointed out and discussed.

6.13. CYCLOSA: Decentralizing Private Web Search Through SGX-Based Browser Extensions

Participant: Antoine Boutet.

By regularly querying Web search engines, users (unconsciously) disclose large amounts of their personal data as part of their search queries, among which some might reveal sensitive information (e.g. health issues, sexual, political or religious preferences). Several solutions exist to allow users querying search engines while improving privacy protection. However, these solutions suffer from a number of limitations: some are subject to user re-identification attacks, while others lack scalability or are unable to provide accurate results. This contribution presents CYCLOSA, a secure, scalable and accurate private Web search solution. CYCLOSA improves security by relying on trusted execution environments (TEEs) as provided by Intel SGX. Further, CYCLOSA proposes a novel adaptive privacy protection solution that reduces the risk of user re-identification. CYCLOSA sends fake queries to the search engine and dynamically adapts their count according to the sensitivity of the user query. In addition, CYCLOSA meets scalability as it is fully decentralized, spreading the load for distributing fake queries among other nodes. Finally, CYCLOSA achieves accuracy of Web search as it handles the real query and the fake queries separately, in contrast to other existing solutions that mix fake and real query results.

6.14. ACCIO: How to Make Location Privacy Experimentation Open and Easy

Participant: Antoine Boutet.

The advent of mobile applications collecting and exploiting the location of users opens a number of privacy threats. To mitigate these privacy issues, several protection mechanisms have been proposed this last decade to protect users' location privacy. However, these protection mechanisms are usually implemented and evaluated in monolithic way, with heterogeneous tools and languages. Moreover, they are evaluated using different methodologies, metrics and datasets. This lack of standard makes the task of evaluating and comparing protection mechanisms particularly hard. In this work, we present ACCIO, a unified framework to ease the design and evaluation of protection mechanisms. Thanks to its Domain Specific Language, ACCIO allows researchers and practitioners to define and deploy experiments in an intuitive way, as well as to easily collect and analyse the results. ACCIO already comes with several state-of-the-art protection mechanisms and a toolbox to manipulate mobility data. Finally, ACCIO is open and easily extensible with new evaluation metrics and protection mechanisms. This openness, combined with a description of experiments through a user-friendly DSL, makes ACCIO an appealing tool to reproduce and disseminate research results easier. In this work, we present ACCIO's motivation and architecture, and demonstrate its capabilities through several use cases involving multiples metrics, state-of-the-art protection mechanisms, and two real-life mobility datasets collected in Beijing and in the San Francisco area.

6.15. Collaborative Filtering Under a Sybil Attack: Similarity Metrics do Matter!

Participant: Antoine Boutet.

Recommendation systems help users identify interesting content, but they also open new privacy threats. In this contribution, we deeply analyze the effect of a Sybil attack that tries to infer information on users from a user-based collaborative-filtering recommendation systems. We discuss the impact of different similarity metrics used to identity users with similar tastes in the trade-off between recommendation quality and privacy. Finally, we propose and evaluate a novel similarity metric that combines the best of both worlds: a high recommendation quality with a low prediction accuracy for the attacker. Our results, on a state-of-the-art recommendation framework and on real datasets show that existing similarity metrics exhibit a wide range of behaviors in the presence of Sybil attacks, while our new similarity metric consistently achieves the best trade-off while outperforming state-of-the-art solutions.

6.16. Automatic Privacy and Utility Preservation of Mobility Data: A Nonlinear Model-Based Approach

Participant: Antoine Boutet.

The widespread use of mobile devices and location-based services has generated massive amounts of mobility databases. While processing these data is highly valuable, privacy issues can occur if personal information is revealed. The prior art has investigated ways to protect mobility data by providing a large range of Location Privacy Protection Mechanisms (LPPMs). However, the privacy level of the protected data significantly varies depending on the protection mechanism used, its configuration and on the characteristics of the mobility data. Meanwhile, the protected data still needs to enable some useful processing. To tackle these issues, in this work we present PULP, a framework that finds the suitable protection mechanism and automatically configures it for each user in order to achieve user-defined objectives in terms of both privacy and utility. PULP uses nonlinear models to capture the impact of each LPPM on data privacy and utility levels. Evaluation of our framework is carried out with two protection mechanisms of the literature and four real-world mobility datasets. Results show the efficiency of PULP, its robustness and adaptability. Comparisons between LPPMs' configurator and the state of the art further illustrate that PULP better realizes users' objectives and its computations time is in orders of magnitude faster.

6.17. Privacy Preserving Analytics

Participant: Mathieu Cunche.

As communications-enabled devices are becoming more ubiquitous, it becomes easier to track the movements of individuals through the radio signals broadcasted by their devices. Thus, while there is a strong interest for physical analytics platforms to leverage this information for many purposes, this tracking also threatens the privacy of individuals. To solve this issue, we propose a privacy-preserving solution for collecting aggregate mobility patterns while satisfying the strong guarantee of ε-differential privacy. More precisely, we introduce a sanitization mechanism for efficient, privacy-preserving and non-interactive approximate distinct counting for physical analytics based on perturbed Bloom filters called Pan-Private BLIP. We also extend and generalize previous approaches for estimating distinct count of events and joint events (i.e., intersection and more generally t-out-of-n cardinalities). Finally, we evaluate expirementally our approach and compare it to previous ones on real datasets.

6.18. Detecting smartphone state changes through a Bluetooth based timing attack

Participants: Mathieu Cunche, Guillaume Celosia.

Bluetooth is a popular wireless communication technology that is available on most mobile devices. Although Bluetooth includes security and privacy preserving mechanisms, we show that a Bluetooth harmless inherent request-response mechanism can taint users privacy. More specifically, we introduce a timing attack that can be triggered by a remote attacker in order to infer information about a Bluetooth device state. By observing the L2CAP layer ping mechanism timing variations, it is possible to detect device state changes, for instance when the device goes in or out of the locked state. Our experimental results show that change point detection analysis of the timing allows to detect device state changes with a high accuracy. Finally, we discuss applications and countermeasures.

6.19. Analyzing Ultrasound-based Physical Tracking Systems

Participant: Mathieu Cunche.

A trending application of ultrasound communication is the implementation of ultrasound beacons to track owners of mobile phones in stores and shopping centers. We present the analysis of an Ultrasound-based tracking application. By analyzing several mobile applications along with the network communication and sample of the original audio signal, we were able to reverse engineer the ultrasonic communications and some other elements of the system. Based on those finding we show how arbitrary ultrasonic signal can be generated and how to perform jamming. Finally we analyze a real world deployment and discuss privacy implications.

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. AMNECYS

- Title: AMNECYS
- Duration: 2015 - .
- Coordinator: CESICE, UPMF.
- Others partners: Inria/Privatics and LIG/Moais, Gipsa-lab, LJK, Institut Fourier, TIMA, Vérimag, LISTIC (Pole MSTIC) .
- Abstract: Privatics participates to the creation of an Alpine Multidisciplinary NEtwork on CYbersecurity Studies (AMNECYS). The academic teams and laboratories participating in this project have already developed great expertise on encryption technologies, vulnerabilities analysis, software engineering, protection of privacy and personal data, international & European aspects of cybersecurity. The first project proposal (ALPEPIC ALPs-Embedded security: Protecting Iot & Critical infrastructure) focuses on the protection of the Internet of Things (IoT) and Critical Infrastructure (CI).

7.1.2. Data Institute

- Title: Data Institute UGA
- Duration: 2017 - .
- Coordinator: TIMC-IMAG.
- Others partners: AGEIS, BIG, CESICE, GIN, GIPSA-lab, IAB, IGE, IPAG, LAPP, LARHRA, LIDILEM, LIG, LISTIC, LITT&ArTS, LJK, LUHCIE, LECA, OSUG, PACTE, TIMC-IMAG

- Abstract: Privatics is leading the WP5 (Data Governance, Data Protection and Privacy). This action (WP5) aims to analyze, in a multi-disciplinary perspective, why and how specific forms of data governance emerge as well as the consequences on the interaction between the state, the market and society. The focus will be on the challenges raised by the collection and use of data for privacy, on the data subjects' rights and on the obligations of data controllers and processors. A Privacy Impact/Risk assessments methodology and software will be proposed. A case study will focus on medical and health data and make recommendations on how they should be collected and processed.

7.1.3. CyberAlps

- Title: CyberAlps
- Duration: 2018 - .
- Coordinator: IF.
- Others partners: CEA LETI, CERAG, CESICE, CREg, G2E lab, GIPSA-lab, GSCOP, IF, LCIS, LIG, LISTIC, LJK, PACTE, TIMC-IMAG, VERIMAG.
- Abstract: The Grenoble Alpes Cybersecurity Institute aims at undertaking ground-breaking interdisciplinary research in order to address cybersecurity and privacy challenges. Our main technical focus is on low-cost secure elements, critical infrastructures, vulnerability analysis and validation of large systems, including practical resilience across the industry and the society. Our approach to cybersecurity is holistic, encompassing technical, legal, law-enforcement, economic, social, diplomatic, military and intelligence-related aspects with strong partnerships with the private sector and robust national and international cooperation with leading institutions in France and abroad.

7.1.4. Antidot

- Title: Antidot
- Type: Fédération Informatique de Lyon (inter laboratories project)
- Duration: September 2018 - 2020.
- Coordinator: Inria.
- Others partners: LIRIS.
- Abstract: The ANTIDOT project is interested in the privacy issues raised by the increasingly ubiquitous collection of mobility data and their exploitation by third-party applications. The objective of this project is to propose solutions and tools to increase the user awareness about the risks of violation of their privacy in the context of the mobile Internet. In order to achieve this objective, ANTIDOT will jointly address the study of information gathering mechanisms, the study of mobility data vulnerabilities and the protection of this personal data.

7.2. National Initiatives

7.2.1. FUI

Title: ADAGE (Anonymous Mobile Traffic Data Generation).

Type: FUI.

Duration: July 2016 - September 2018.

Coordinator: Orange.

Others partners: Inria, CNRS LAAS.

Abstract: The project ADAGE aims at developping solutions for the anonymization of mobility traces produced by mobile operators.

7.2.2. ANR

7.2.2.1. CISC

Title: Certification of IoT Secure Compilation.

Type: ANR.

Duration: April 2018 - March 2022.

Coordinator: Inria INDES project-team (France)

Others partners: Inria CELTIC project-team (France), College de France (France) (France).

See also: http://cisc.gforge.inria.fr.

Abstract: The objective of the ANR CISC project is to investigate multitier languages and compilers to build secure IoT applications with private communication. A first goal is to extend multitier platforms by a new orchestration language that we call Hiphop.js to synchronize internal and external activities of IoT applications as a whole. CISC will define the language, semantics, attacker models, and policies for the IoT and investigate automatic implementation of privacy and security policies by multitier compilation of IoT applications. To guarantee such applications are correct, and in particular that the required security and privacy properties are achieved, the project will certify them using the Coq proof assistant.

7.2.2.2. SIDES 3.0

Title: Application of privacy by design to biometric access control.

Type: ANR.

Duration: August 2017 - August 2020.

Coordinator: Uness (France).

Others partners: Inria, UGA, ENS, Theia, Viseo.

Abstract: Since 2013, faculties of medicine have used a shared national platform that enables them to carry out all of their validating exams on tablets with automatic correction. This web platform entitled SIDES allowed the preparation of the medical students to the Computerized National Classing Events (ECN) which were successfully launched in June 2016 (8000 candidates simultaneously throughout France). SIDES 3.0 proposes to upgrade the existing platform. Privatics goals in this project is to ensure that privacy is respected and correctly assessed .

7.2.2.3. DAPCODS/IOTics

Title: DAPCODS/IOTics.

Type: ANR 2016.

Duration: May 2017 - Dec. 2020.

Coordinator: Inria PRIVATICS.

Others partners: Inria DIANA, EURECOM, Univ. Paris Sud, CNIL.

Abstract:

Thanks to the exponential growth of Internet, citizens have become more and more exposed to personal information leakage in their digital lives. This trend began with web tracking when surfing the Internet with our computers. The advent of smartphones, our personal assistants always connected and equipped with many sensors, further reinforced this tendency. And today the craze for "quantified self" wearable devices, for smart home appliances or for other connected devices enable the collection of potentially highly sensitive personal information in domains that were so far out of reach. However, little is known about the actual practices in terms of security, confidentiality, or data exchanges. The enduser is therefore prisoner of a highly asymmetric system. This has important consequences in terms of regulation, sovereignty, and leads to the hegemony of the GAFAs (Google, Amazon, Facebook and Apple). Security, transparency and user control are three key properties that should be followed by all the stakeholders of the smartphone and connected devices ecosystem. Recent scandals show that the reality is sometimes at the opposite.

The DAPCODS project gathers four renowned research teams, experts in security, privacy and digital economy. They are seconded by CNIL, the French data protection agency. The project aims at contributing along several axes:

- by analyzing the inner working of a significant set of connected devices in terms of personal information leaks. This will be made possible by analyzing their data flows (and associated smartphone application if applicable) from outside (smartphone and/or Wifi network) or inside, through ondevice static and dynamic analyses. New analysis methods and tools will be needed, some of them leveraging on previous works when applicable;
- by studying the device manufacturers' privacy policies along several criteria (e.g., accessibility, precision, focus, privacy risks). In a second step, their claims will be compared to the actual device behavior, as observed during the test campaigns. This will enable an accurate and unique ranking of connected devices;
- by understanding the underlying ecosystem, from the economical viewpoint. Data collected will make it possible to define the blurred boundaries of personal information market, a key aspect to set up an efficient regulation;
- and finally, by proposing a public website that will rank those connected devices and will inform citizens. We will then test the impact of this information on the potential change of behavior of stakeholders.

By giving transparent information of hidden behaviors, by highlighting good and bad practices, this project will contribute to reduce the information asymmetry of the system, to give back some control to the endusers, and hopefully to encourage certain stakeholders to change practices.

7.2.3. Inria Innovation Laboratory

Title: LEELCO (Low End-to-End Latency COmmunications).

Duration: 3 years (2015 - 2018).

Coordinator: Inria PRIVATICS.

Others partners: Expway.

Abstract:

This Inria Innovation Lab aims at strengthening Expway (http://www.expway.com/) commercial offer with technologies suited to real-time data transmissions, typically audio/video flows. In this context, the end-to-end latency must be reduced to a minimum in order to enable a high quality interaction between users, while keeping the ability to recover from packet losses that are unavoidable with wireless communications in harsh environments. In this collaboration we focus on new types of Forward Erasure Correction (FEC) codes based on a sliding encoding windows, and on the associated communication protocols, in particular an extension to FECFRAME (RFC6363) to such FEC codes. The outcomes of this work are proposed to both IETF and 3GPP standardisation organisations, in particular in the context of 3GPP mission critical communication services activity. The idea of this 3GPP activity is to leverage on the 3GPP Evolved Multimedia Broadcast Multicast Services (eMBMS) and on the existing Long Term Evolution (LTE) infrastructure for critical communications and such services as group voice transmissions, live high-definition video streams and large data transmissions. In this context, the advanced FEC codes studied in LEELCO offer a significant improvement both from the reduced latency and increased loss recovery viewpoints compared to the Raptor codes included in the existing standard (https://hal.inria.fr/hal-01571609v1/en/).

7.2.4. Inria CNIL project

Privatics is in charged of the Cnil-Inria collaboration. This collaboration was at the origin of the Mobilitics project and it is now at the source of many discussions and collaborations on data anoymisation, risk analysis, consent or IoT Privacy. Privatics and Cnil are both actively involved on the IoTics project, that is the follow-up of the Mobilitics projects. The goal of the Mobilitics project was to study information leakage in mobile phones. The goal of IoTics is to extend this work to IoT and connected devices.

Privatics is also in charged of the organization of the Cnil-Inria prize that is awarded every year to an outstanding publication in the field of data privacy.

7.3. European Initiatives

7.3.1. Collaborations in European Programs, Except FP7 & H2020

7.3.1.1. COPES

Title: COnsumer-centric Privacy in smart Energy gridS

Programm: CHISTERA

Duration: December 2015 - december 2018

Coordinator: KTH Royal Institute of Technology

Inria contact: Cédric Lauradoux

Smart meters have the capability to measure and record consumption data at a high time resolution and communicate such data to the energy provider. This provides the opportunity to better monitor and control the power grid and to enable demand response at the residential level. This not only improves the reliability of grid operations but also constitutes a key enabler to integrate variable renewable generation, such as wind or solar. However, the communication of high resolution consumption data also poses privacy risks as such data allows the utility, or a third party, to derive detailed information about consumer behavior. Hence, the main research objective of COPES is to develop new technologies to protect consumer privacy, while not sacrificing the ”smartness”, i.e., advanced control and monitoring functionalities. The core idea is to overlay the original consumption pattern with additional physical consumption or generation, thereby hiding the consumer privacy sensitive consumption. The means to achieve this include the usage of storage, small scale distributed generation and/or elastic energy consumptions. Hence, COPES proposes and develops a radically new approach to alter the physical energy flow, instead of purely relying on encryption of meter readings, which provides protection against third party intruders but does not prevent the use of this data by the energy provider.

7.3.1.2. UPRISE-IoT

Title: User-centric PRIvacy & Security in IoT

Programm: CHISTERA

Duration: December 2016 - december 2019

Coordinator: SUPSI (Suisse)

Inria contact: Claude Castelluccia

The call states that “Traditional protection techniques are insufficient to guarantee users’ security and privacy within the future unlimited interconnection”: UPRISE-IoT will firstly identify the threats and model the behaviours in IoT world, and further will build new privacy mechanisms centred around the user. Further, as identified by the call “all aspects of security and privacy of the user data must be under the control of their original owner by means of as simple and efficient technical solutions as possible”, UPRISE-IoT will rise the awareness of data privacy to the users. Finally, it will deeply develop transparency mechanisms to “guarantee both technically and regulatory the neutrality of the future internet.” as requested by the call. The U-HIDE solution developed inn UPRISE-IoT will “empower them to understand and make their own decisions regarding their data, which is essential in gaining informed consent and in ensuring the take-up of IoT technologies”, using a methodology that includes “co-design with users to address the key, fundamental, but inter-related and interdisciplinary aspects of privacy, security and trust.”

7.4. International Initiatives

7.4.1. DATA

Title: Data and Algorithmic Transparency and Accountability

International Partner (Institution - Laboratory - Researcher):

Université du Québec à Montréal (UQAM) (Canada) - Département d'informatique - Sébastien Gambs

Start year: 2018

See also: http://planete.inrialpes.fr/data-associated-team/

The accelerated growth of the Internet has outpaced our abilities as individuals to maintain control of our personal data. The recent advent of personalized services has lead to the massive collection of personal data and the construction of detailed profiles about users. However, users have no information about the data which constitute its profile and how they are exploited by the different entities (Internet companies, telecom operators, ...). This lack of transparency gives rise to ethical issues such as discrimination or unfair processing.

In this associate team, we propose to strengthen the complementary nature and the current collaborations between the Inria Privatics group and UQAM to advance research and understanding on data and the algorithmic transparency and accountability.

7.5. International Research Visitors

7.5.1. Visits of International Scientists

- Sébastien Gambs visited the team in Lyon in April 2018 for a week to initiate the DATA collaboration. We also organized a workshop in data and algorithmic transparency during this week.
- Gergely Acs, assistant professor at Budapest University (Hungary), visited our team for 2 months, from mi-May to mid-July. He worked together with Claude Castelluccia on machine learning (in)security. In particular, he studied how adversarial examples can be used to evade monitoring, and consequently improve privacy.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. General Chair, Scientific Chair

Antoine Boutet: Workshop on data transparency, 23/04/2018, Lyon, France.

Claude Castelluccia: APVP 2018, 3-6/06/2018, Porquerolles, France).

Claude Castelluccia: *Intelligence Oversight : Is Human Rights-Preserving Surveillance Possible?*, Grenoble Data institute, 25/01/2018, Grenoble, France.

8.1.1.2. Member of the Organizing Committees

Antoine Boutet: Winter School on Distributed Systems and Networks 2018, 4-8/02/2018, Sept Laux, France.

Antoine Boutet: SRDS 2018, 02-05/10/2018, Salvador, Brasil.

Mathieu Cunche: French GNURadio Days, 03/07/2018, INSA Lyon, France.

Daniel Le Metayer: Panel *Physical tracking: nowhere to hide*, CPDP 2018, 24/01/2018, Brussels, Belgium.

Vincent Roca: IEEE WiMob 2018.

8.1.2. Scientific Events Selection

8.1.2.1. Member of the Conference Program Committees

Antoine Boutet: Compas 2018, APVP 2018, Middleware 2018.

Mathieu Cunche: ACM WiSec 2018, Mobiquitous 2018, WCNC 2019, ICISSP 2018.

Claude Castelluccia: APF 2018.

Daniel Le Metayer: IWPE 2018, CPDP 2018, APF 2018.

Vincent Roca: SPACOMM 2018, SSCC 2018.

8.1.3. Invited Talks

Antoine Boutet: *Feedback on the Shonan Meeting on Anonymization methods and inference attacks*, 4th Franco-Japanese Cybersecurity workshop, 16/05/2018, Annecy, France.

Claude Castelluccia: *Cognitive security*, 4th Franco-Japanese Cybersecurity workshop, 16/05/2018, Annecy, France.

Claude Castelluccia: *Brain Hacking*, College des Bernardins, 11/12/2018, Paris, France.

Claude Castelluccia: *Plateforme en ligne et transparence*, AFDIT, 09/11/2018, Paris, France.

Claude Castelluccia: *Internet Surveillance*, Intelligence Oversight workshop, 25/01/2018, Grenoble, France.

Cédric Lauradoux: *Cybersécurité et cybermenaces*, Club Democraties, 09/03/2018, Paris, France.

Cédric Lauradoux: *Cybersécurité et cybermenaces*, Rectorat Académie de Grenoble, 10/12/2018, Grenoble, France.

Daniel Le Metayer, *Intelligility and transparency in machine learning and AI*, Société Française de Statistique, 18/05/2018, Paris, France.

Daniel Le Metayer, *Transparency and opacity in IT systems*, INSA Lyon CITI, 23/04/2018, Lyon, France.

Vincent Roca, *Privacy and Connected Objects*, Eclipse IoT Days Grenoble, 18/01/2018, Grenoble, France.

Vincent Roca, *Archéologie de la fuite de nos données personnelles par le biais de nos téléphones*, Atelier Internet – ENSIIB, 06/04/2018, Lyon, France.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master : Antoine Boutet, *Privacy*, 12h, INSA-Lyon, France.

Master : Antoine Boutet, *Security*, 12h, INSA-Lyon, France.

Undergraduate course : Mathieu Cunche, *Introduction to computer science*, 120h, L1, INSA-Lyon, France.

Master : Mathieu Cunche, *Wireless Security*, 6h, M2, INSA-Lyon, France.

Undergraduate course : Mathieu Cunche, *On Wireless Network Security*, 10h, L1, IUT-2 (UPMF - Grenoble University) , France.

Undergraduate course : Mathieu Cunche, *Security & Privacy*, 21h, L3, INSA-Lyon, France.

Master : Mathieu Cunche, *Privacy and Data protection*, 14h, M2, INSA-Lyon, France.

Master : Mathieu Cunche, *Cryptography and Communication Security*, 18h, M1, INSA-Lyon, France.

Master : Cédric Lauradoux, *Advanced Topics in Security*, 20h, M2, Ensimag/INPG, France.

Master : Cédric Lauradoux, *Systems and Network Security*, 30h, M1, Ensimag, France.

Master : Cédric Lauradoux, *Internet Security*, 12h, M2, University of Grenoble Alpes, France.

Master : Cédric Lauradoux, *Cyber Security*, 3h, M2, Laws School of University of Grenoble Alpes, France.

Master : Claude Castelluccia, *Advanced Topics in Security*, 15h, M2, Ensimag/INPG, France.

Master : Claude Castelluccia, *Cyber Security*, 6h, M2, Laws School of University of Grenoble Alpes, France.

Master : Claude Castelluccia, *Data Privacy*, 6h, M2, Laws School of University of Grenoble Alpes, France.

Master : Daniel Le Metayer, *Privacy*, 12h, M2 MASH, Université Paris Dauphine, France.

Master : Daniel Le Metayer, *Privacy*, 12h, M2, Insa Lyon, France.

Master : Vincent Roca, *On Wireless Communications*, 12h, M1, Polytech' Grenoble, France.

Undergraduate course : Vincent Roca, *On Network Communications*, 44h, L1, IUT-2 (University of Grenoble Alpes), France.

Undergraduate course : Vincent Roca, *On Security and Privacy in smartphones*, 6h, L-Pro, University of Grenoble Alpes, France.

Master : Vincent Roca, *On Security and Privacy in smartphones*, 3h, M2, France.

8.2.2. E-learning

E-learning

Mooc: Cédric Lauradoux and Vincent Roca, *, 2 month session, FUN-MOOC, Inria, public ciblé, 23000 (13000 first session and 10000 second session).*

8.2.3. Supervision

- PhD in progress : Victor Morel, *IoT privacy*, September 2016, Daniel Le Métayer and Claude Castelluccia.
- PhD in progress : Mathieu Thiery, *IoT privacy* , September 2016, Vincent Roca.
- PhD in progress : Guillaume Celosia, *Wireless Privacy in the Internet of Things* , November 2017, Mathieu Cunche and Daniel Le Métayer.
- PhD in progress : Supryia Adhatarao, *Privacy of E-learning systems* , March 2018, Cédric Lauradoux.
- PhD in progress : Coline Boniface, *Cyberweapons: from bug bounties to zero days* , March 2018, Cédric Lauradoux.
- PhD in progress : Raoul Kerkouche, *Privacy-Preserving Processing of Medical Data* , January 2018, Claude Castelluccia.
- PhD in progress : Clement Henin, *Explainable AI* , September 2018, Claude Castelluccia et Daniel Le Metayer.
- PhD in progress: Théo Jourdan, *Privacy-preserving machine learning in medical domain, October 2018, Antoine Boutet.*
- *Intern (M2): Louis Beziaud,* Privacy of national identity systems*, M2 ENS Rennes, Claude Castelluccia et Daniel Le Metayer.*
- *Intern (L3): Alexandre van Beurden, Inspect what location history reveals about an individual, Antoine Boutet.*
- *Intern (L3): Romain Fournier, Development of a cybersecurity platform, Antoine Boutet.*
- *Intern (L3): Bastien Durand, Analysis of the correlation between the mobility and the personality of an individual, Antoine Boutet.*

8.2.4. Juries

PhD: David Gerault, *Security Analysis of Contactless Communication Protocols*, Université Clermont Auvergne , 27/11/2018, Cédric Lauradoux.

PhD: Jonathan Detchart, *Optimisation de codes correcteurs d'effacements par application de transformées polynomiales*, Université de Toulouse, 05/12/2018, Vincent Roca.

PhD: Elise Tourne, *Le phénome de circulation des données à caractère personnel dans le cloud: étude de droit matériel dans le contexte de l'Union Européenne*, Université Lyon 3, 11/06/2018, Daniel Le Métayer.

8.3. Popularization

8.3.1. Hearings

- Claude Castelluccia: *Understanding Algorithmic Decision-Making Systems*, European Parliament, 10/2018, Strasbourg, France.
- Daniel Le Metayer: *Understanding Algorithmic Decision-Making Systems*, European Parliament, 10/2018, Strasbourg, France.
- Daniel Le Metayer: *Transparence et explicabilité des algorithmes d'aide à la décision*, CCNE, 02/2018, Paris, France.

8.3.2. Internal or external Inria responsibilities

- Claude Castelluccia is co-leader of the Worpackage 5 (data governance and privacy) of the Grenoble Data Institute.
- Claude Castelluccia is co-leader of Grenoble CyberAlps (cybersecurity institute of Grenoble).
- Daniel Le Metayer chairs the CNIL-Inria privacy award.
- Vincent Roca is co-editor of the white book *Cybersecurity: current challenges and Inria's research directions*.

8.3.3. Articles and contents

- Claude Castelluccia: *Manipulation informationnelle et psychologique*, Le blog binaire du Monde, 05/2018.
- Claude Castelluccia: *Data surveillance and manipulation*,Interview for Magazine Capital, 12/2018.
- Mathieu Cunche: *Cybersécurité et menace informatique*, Sommet des start-up sciencesetavenir.fr, 11/2018, Lyon, France.
- Mathieu Cunche: *Attaque par déni de service dans le Wi-Fi*, GNULinux Magazine HS 99, 11/2018.
- Mathieu Cunche: *Comprendre les attaques Krack*, GNULinux Magazine HS 99, 11/2018.
- Daniel Le Metayer: *Weighting the impact of the GDPR*, Communications of the ACM, 11/2018.
- Daniel Le Metayer: *Qui gouverne les algorithmes ?*, Revue THIIRD, 11/2018.
- Vincent Roca: Inria White Paper in Cyber-Security.

8.3.4. Education

- Cédric Lauradoux: *Action nombres et cryptographie*, Maison pour la science, Inria, 06/02/2018, Grenoble, France.
- Cédric Lauradoux: *Action nombres et cryptographie*, Maison pour la science, 18/12/2018, Annecy, France.
- Cédric Lauradoux: *Animation du forum du MOOC Protection de la vie privée dans le monde numérique*, 02-03/2018.
- Cédric Lauradoux: : *Animation du forum du MOOC Protection de la vie privée dans le monde numérique*, 11-12/2018.

8.3.5. Interventions

- Cédric Lauradoux: *Atelier cryptograpie*, Fête de la Science, 11-12/10/2018, Grenoble, France.
- Cédric Lauradoux: *Cryptologie et Vie privée*, Semaine des mathématiques, Lycée Gabriel-Faure, 15/03/2018, Tournon, France.
- Cédric Lauradoux: *Cryptologie et Vie privée*, Semaine des mathématiques, Lycée Boissy d'Anglas, 16/03/2018, Annonay, France.
- Cédric Lauradoux: *Challenge de cryptologie*, MathC2+ internship, 26/06/18, Grenoble, France.
- Cédric Lauradoux: *Challenge de cryptologie*, Cité scolaire Jean PREVOST, 02/06/2018, Villard de Lans, France.
- Cédric Lauradoux: *Challenge de cryptologie*, Collège Barnave, 18/01/2018, Saint-Égrève, France.

8.3.6. Internal action

- Cédric Lauradoux: *Réglementation sur les données*, Inria, Grenoble, 19/06/2018.
- Cédric Lauradoux: *Réglementation sur les données*, IMAG, Grenoble, 09/07/2018.

9. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] J. P. ACHARA, J. PARRA-ARNAU, C. CASTELLUCCIA.*Fine-Grained Control over Tracking to Support the Ad-Based Web Economy*, in "ACM Transactions on Internet Technology", October 2018, vol. 18, n^{o} 4, p. 1-25 [*DOI :* 10.1145/3158372], https://hal.inria.fr/hal-01921868

[2] G. ACS, L. MELIS, C. CASTELLUCCIA, E. DE CRISTOFARO.*Differentially Private Mixture of Generative Neural Networks (ext)*, in "IEEE Transactions on Knowledge and Data Engineering", July 2018, p. 1-13 [*DOI :* 10.1109/TKDE.2018.2855136], https://hal.inria.fr/hal-01921923

[3] M. ALAGGAN, M. CUNCHE, S. GAMBS.*Privacy-preserving Wi-Fi Analytics*, in "Proceedings on Privacy Enhancing Technologies", April 2018, vol. 2018, n^{o} 2, p. 4-26 [*DOI :* 10.1515/POPETS-2018-0010], https://hal.inria.fr/hal-01719211

[4] J.-S. BERGÉ, D. LE MÉTAYER.*Phénomènes de masse et droit des données*, in "Communication Commerce Électronique", 2018, vol. n° 12, https://hal.archives-ouvertes.fr/hal-01907540

[5] J. BRINGER, H. CHABANNE, D. LE MÉTAYER, R. LESCUYER.*Biometric Systems Private by Design: Reasoning about privacy properties of biometric system architectures*, in "Transactions on Data Privacy", 2018, vol. 11, n^{o} 2, p. 111-137, https://hal.archives-ouvertes.fr/hal-01939841

[6] S. CERF, S. BOUCHENAK, B. ROBU, N. MARCHAND, V. PRIMAULT, S. BEN MOKHTAR, A. BOUTET, L. Y. CHEN.*Automatic Privacy and Utility Preservation of Mobility Data: A Nonlinear Model-Based Approach*, in "IEEE Transactions on Dependable and Secure Computing", 2018, https://hal.archives-ouvertes.fr/hal-01910687

[7] L. DEMIR, A. KUMAR, M. CUNCHE, C. LAURADOUX.*The Pitfalls of Hashing for Privacy*, in "Communications Surveys and Tutorials, IEEE Communications Society", 2018, vol. 20, n^{o} 1, p. 551-565 [*DOI :* 10.1109/COMST.2017.2747598], https://hal.inria.fr/hal-01589210

[8] D. LE MÉTAYER, S. DESMOULIN-CANSELIER.*Algorithmic Decision Systems in the Health and Justice Sectors: Certification and Explanations for Algorithms in European and French Law*, in "European Journal of Law and Technology", December 2018, https://hal.archives-ouvertes.fr/hal-01972259

[9] J. PARRA-ARNAU, C. CASTELLUCCIA.*On the Cost-Effectiveness of Mass Surveillance*, in "IEEE Access", 2018, vol. 6, p. 46538 - 46557 [*DOI :* 10.1109/ACCESS.2018.2866310], https://hal.inria.fr/hal-01921899

[10] V. PRIMAULT, A. BOUTET, S. BEN MOKHTAR, L. BRUNIE.*The Long Road to Computational Location Privacy: A Survey*, in "Communications Surveys and Tutorials, IEEE Communications Society", 2018, 1 [*DOI :* 10.1109/COMST.2018.2873950], https://hal.archives-ouvertes.fr/hal-01890014

International Conferences with Proceedings

[11] A. BOUTET, F. DE MOOR, D. FREY, R. GUERRAOUI, A.-M. KERMARREC, A. RAULT.*Collaborative Filtering Under a Sybil Attack: Similarity Metrics do Matter!*, in "DSN 2018 - the 48th International Conference on Dependable Systems and Networks", Luxembourg, Luxembourg, IEEE, June 2018, p. 466-477 [*DOI :* 10.1109/DSN.2018.00055], https://hal.inria.fr/hal-01787060

[12] G. CELOSIA, M. CUNCHE.*Detecting smartphone state changes through a Bluetooth based timing attack*, in "WiSec '18 - 11th ACM Conference on Security & Privacy in Wireless and Mobile Networks", Stockholm, Sweden, June 2018, p. 154-159 [*DOI :* 10.1145/3212480.3212494], https://hal.inria.fr/hal-01870011

[13] S. J. DE, D. LE MÉTAYER.*Privacy Risk Analysis to Enable Informed Privacy Settings*, in "IWPE 2018 – 4th IEEE International Workshop on Privacy Engineering", London, United Kingdom, Proceedings of the 4th IEEE International Workshop on Privacy Engineering (IWPE 2018), April 2018, p. 1-8, https://hal.archives-ouvertes.fr/hal-01939845

[14] G. G. GULYÁS, D. F. SOMÉ, N. BIELOVA, C. CASTELLUCCIA.*To Extend or not to Extend: On the Uniqueness of Browser Extensions and Web Logins*, in "WPES'18 - Workshop on Privacy in the Electronic Society", Toronto, Canada, ACM Press, October 2018, p. 14-27 [*DOI :* 10.1145/3267323.3268959], https://hal.inria.fr/hal-01921863

[15] T. JOURDAN, A. BOUTET, C. FRINDEL.*Toward privacy in IoT mobile devices for activity recognition*, in "Privacy Preserving Machine Learning NeurIPS 2018 Workshop", Montréal, Canada, December 2018, p. 1-6, https://hal.inria.fr/hal-01941453

[16] T. JOURDAN, A. BOUTET, C. FRINDEL.*Toward privacy in IoT mobile devices for activity recognition*, in "MobiQuitous 2018 - 15th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services", New York city, United States, November 2018, p. 1-10, https://hal.inria.fr/hal-01882330

[17] R. KERKOUCHE, R. ALAMI, R. FÉRAUD, N. VARSIER, P. MAILLÉ.*Node-based optimization of LoRa transmissions with Multi-Armed Bandit algorithms*, in "ICT 2018 - 25th International Conference on Telecommunications", Saint Malo, France, June 2018, p. 1-6, https://hal-imt-atlantique.archives-ouvertes.fr/hal-01946456

[18] D. LE MÉTAYER, P. RAUZY.*Capacity: an Abstract Model of Control over Personal Data*, in "CODASPY 2018 - ACM Conference on Data and Application Security and Privacy", Tempe, United States, Proceedings of the ACM Conference on Data and Application Security and Privacy (CODASPY 2018), ACM, March 2018, p. 1-12, https://hal.archives-ouvertes.fr/hal-01939847

[19] V. MOREL, D. LE MÉTAYER, M. CUNCHE, C. CASTELLUCCIA.*Enhancing Transparency and Consent in the IoT*, in "IWPE 2018 - International Workshop on Privacy Engineering", London, United Kingdom, Proceedings of the International Workshop on Privacy Engineering (IWPE 2018), IEEE, April 2018, p. 116-119 [*DOI :* 10.1109/EUROSPW.2018.00023], https://hal.archives-ouvertes.fr/hal-01709255

[20] R. PIRES, D. GOLTZSCHE, S. BEN MOKHTAR, S. BOUCHENAK, A. BOUTET, P. FELBER, R. KAPITZA, M. PASIN, V. SCHIAVONI.*CYCLOSA: Decentralizing Private Web Search Through SGX-Based Browser Extensions*, in "ICDCS 2018 - 38th IEEE International Conference on Distributed Computing Systems", Vienne, Austria, IEEE, July 2018, p. 467-477 [*DOI :* 10.1109/ICDCS.2018.00053], https://hal.inria.fr/hal-01882430

[21] V. PRIMAULT, M. MAOUCHE, A. BOUTET, S. BEN MOKHTAR, S. BOUCHENAK, L. BRUNIE.*ACCIO: How to Make Location Privacy Experimentation Open and Easy*, in "ICDCS 2018 - 38th IEEE International Conference on Distributed Computing Systems", Vienna, Austria, Proceedings of the 38th IEEE International Conference on Distributed Computing Systems, IEEE, July 2018, p. 896-906 [*DOI :* 10.1109/ICDCS.2018.00091], https://hal.archives-ouvertes.fr/hal-01784557

Conferences without Proceedings

[22] M. CUNCHE, L. SAMPAIO CARDOSO.*Analyzing Ultrasound-based Physical Tracking Systems*, in "GreHack 2018", Grenoble, France, November 2018, https://hal.inria.fr/hal-01927513

Scientific Books (or Scientific Book chapters)

[23] G. ACS, G. BICZÓK, C. CASTELLUCCIA.*Privacy-Preserving Release of Spatio-Temporal Density*, in "Handbook of Mobile Data Privacy", Springer, October 2018, p. 307-335 [*DOI :* 10.1007/978-3-319-98161-1_12], https://hal.inria.fr/hal-01921891

[24] M. CUNCHE.*Syria*, in "The SAGE Encyclopedia of Surveillance, Security, andPrivacy", B. A. ARRIGO (editor), SAGE, March 2018, https://hal.inria.fr/hal-01798933

Research Reports

[25] A. BOUTET, F. DE MOOR, D. FREY, R. GUERRAOUI, A.-M. KERMARREC, A. RAULT.*Collaborative Filtering Under a Sybil Attack: Similarity Metrics do Matter!*, Inria Rennes - Bretagne Atlantique, April 2018, p. 1-12, https://hal.inria.fr/hal-01767059

[26] A. KASSEM, G. ACS, C. CASTELLUCCIA.*Differential Inference Testing A Practical Approach to Evaluate Anonymized Data*, Inria, January 2018, https://hal.inria.fr/hal-01681014

[27] D. LE MÉTAYER, V. MOREL, M. CUNCHE.*A Generic Information and Consent Framework for the IoT*, Inria, December 2018, n^o RR-9234, p. 1-18, https://arxiv.org/abs/1812.06773 , https://hal.inria.fr/hal-01953052

[28] C. MATTE, M. CUNCHE.*Overview of real-world deployment of physical analytics systems*, Inria Grenoble Rhône-Alpes, January 2018, n^o RR-9143, 18, https://hal.inria.fr/hal-01682373

[29] C. MATTE, M. CUNCHE.*Spread of MAC address randomization studied using locally administered MAC addresses use historic*, Inria Grenoble Rhône-Alpes, January 2018, n^o RR-9142, https://hal.inria.fr/hal-01682363

Scientific Popularization

[30] M. CUNCHE.*Attaque par déni de service dans le Wi-Fi*, in "Linux Magazine France", November 2018, vol. HS 99, https://hal.inria.fr/hal-01927520

[31] M. CUNCHE.*Comprendre les attaques Krack*, in "GNU/Linux Magazine France", November 2018, vol. HS 99, https://hal.inria.fr/hal-01927522

Other Publications

[32] B. ADAMSON, C. ADJIH, J. BILBAO, V. FIROIU, F. FITZEK, G. A. M. SAMAH, E. LOCHIN, A. MASUCCI, M.-J. MONTPETIT, M. V. PEDERSEN, G. PERALTA, V. ROCA, P. SAXENA, S. SIVAKUMAR.*Taxonomy of Coding Techniques for Efficient Network Communications*, June 2018, Internet Research Task Force, Request For Comments (RFC) 8406, https://datatracker.ietf.org/doc/rfc8406/, https://hal.inria.fr/hal-00998506

[33] A. BOUTET, M. CUNCHE.*A Privacy-Preserving Mechanism for Requesting Location Data Provider with Wi-Fi Access Points*, December 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01949419

[34] C. BURDINAT, C. THIENOT, T. TRAN, V. ROCA, B. TEIBI.*Pseudo-CR Performance evaluation of AL-FEC and MCS dimensionning*, February 2018, 3GPP TSG-SA WG4 Meeting #97, Fukuoka, 5th – 6th February 2018, https://hal.inria.fr/hal-01731622

[35] M. CUNCHE, L. S. CARDOSO.*Analysis of an Ultrasound-Based Physical Tracking System*, May 2018, working paper or preprint, https://hal.inria.fr/hal-01798091

[36] J. PARRA-ARNAU, C. CASTELLUCCIA.*Dataveillance and the False-Positive Paradox*, April 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01157921

[37] V. ROCA, A. BEGEN.*Forward Error Correction (FEC) Framework Extension to Sliding Window Codes*, September 2018, Working document of the TSVWG (Transport Area Working Group) group of IETF (Internet Engineering Task Force), draft-ietf-tsvwg-fecframe-ext-06 (work in progress), https://datatracker.ietf.org/doc/draft-ietf-tsvwg-fecframe-ext/, https://hal.inria.fr/hal-01345125

[38] V. ROCA, J. DETCHART, C. ADJIH, M. V. PEDERSEN.*Generic Application Programming Interface (API) for Sliding Window FEC Codes*, November 2018, p. 1-23, Internet Research Task Force - Working document of the Network Coding Research Group (NWCRG), draft-roca-nwcrg-generic-fec-api-04 (work in progress), https://datatracker.ietf.org/doc/draft-roca-nwcrg-generic-fec-api/, https://hal.inria.fr/hal-01630138

[39] V. ROCA, I. SWETT, M.-J. MONTPETIT.*Sliding Window Random Linear Code (RLC) Forward Erasure Correction (FEC) Schemes for QUIC*, June 2018, Internet Research Task Force, Coding for Efficient Network Communications Research Group (NWCRG) document, work in progress, https://datatracker.ietf.org/doc/draft-roca-nwcrg-rlc-fec-scheme-for-quic/, https://hal.inria.fr/hal-01919876

[40] V. ROCA, B. TEIBI.*Sliding Window Random Linear Code (RLC) Forward Erasure Correction (FEC) Schemes for FECFRAME*, September 2018, p. 1-25, Working document of the TSVWG (Transport Area Working Group) group of IETF (Internet Engineering Task Force), draft-ietf-tsvwg-rlc-fec-scheme-09 (work in progress), https://datatracker.ietf.org/doc/draft-ietf-tsvwg-rlc-fec-scheme/, https://hal.inria.fr/hal-01630089

[41] I. SWETT, M.-J. MONTPETIT, V. ROCA.*Coding for QUIC*, June 2018, Internet Research Task Force, Coding for Efficient Network Communications Research Group (NWCRG) document, work in progress, https://datatracker.ietf.org/doc/draft-swett-nwcrg-coding-for-quic/, https://hal.inria.fr/hal-01919858

Project-Team ROMA

Optimisation des ressources : modèles, algorithmes et ordonnancement

IN COLLABORATION WITH: Laboratoire de l'Informatique du Parallélisme (LIP)

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Lyon

Université Claude Bernard (Lyon 1)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Distributed and High Performance Computing

Table of contents

Project-Team ROMA

Creation of the Team: 2012 February 01, updated into Project-Team: 2015 January 01

Keywords:

Computer Science and Digital Science:
A1.1.1. - Multicore, Manycore
A1.1.2. - Hardware accelerators (GPGPU, FPGA, etc.)
A1.1.3. - Memory models
A1.1.4. - High performance computing
A1.1.5. - Exascale
A1.1.9. - Fault tolerant systems
A1.6. - Green Computing
A6.1. - Methods in mathematical modeling
A6.2.3. - Probabilistic methods
A6.2.5. - Numerical Linear Algebra
A6.2.6. - Optimization
A6.2.7. - High performance computing
A6.3. - Computation-data interaction
A7.1. - Algorithms
A8.1. - Discrete mathematics, combinatorics
A8.2. - Optimization
A8.7. - Graph theory
A8.9. - Performance evaluation

Other Research Topics and Application Domains:
B3.2. - Climate and meteorology
B3.3. - Geosciences
B4. - Energy
B4.1. - Fossile energy production (oil, gas)
B4.5.1. - Green computing
B5.2.3. - Aviation
B5.5. - Materials

1. Team, Visitors, External Collaborators

Research Scientists

Frédéric Vivien [Team leader, Inria, Senior Researcher, HDR]
Jean-Yves L'Excellent [Inria, Researcher, HDR]
Loris Marchal [CNRS, Researcher, HDR]
Bora Uçar [CNRS, Researcher]

Faculty Members

Anne Benoit [Ecole Normale Supérieure Lyon, Associate Professor, HDR]
Louis-Claude Canon [Univ. de Franche-Comté, Associate Professor, until Aug 2018]
Yves Robert [Ecole Normale Supérieure Lyon, Professor, HDR]

Samuel Thibault [Univ. Bordeaux 1, Associate Professor, until July 2018]

External Collaborators

Patrick Amestoy [INP Toulouse, external collaborator, HDR]
Alfredo Buttari [CNRS, external collaborator]

Technical Staff

Marie Durand [Inria, from Sep 2018]
Guillaume Joslin [Inria]
Chiara Puglisi [Inria]

PhD Students

Yiqin Gao [Univ de Lyon, from Oct 2018]
Changjiang Gou [China Scholarship Council]
Li Han [China Scholarship Council]
Aurélie Kong Win Chang [Ecole Normale Supérieure Lyon]
Valentin Le Fèvre [Ecole Normale Supérieure Lyon]
Gilles Moreau [Inria]
Ioannis Panagiotas [Inria]
Filip Pawlowski [CIFRE Huawei]
Loïc Pottier [Ecole Normale Supérieure Lyon, until Sep 2018]
Bertrand Simon [Ecole Normale Supérieure Lyon, until Aug 2018]
Issam Raïs [Inria]

Post-Doctoral Fellow

Adrien Rémy [Univ de Lyon, until Jun 2018]

Administrative Assistants

Solene Audoux [Inria, from Nov 2018]
Evelyne Blesle [Inria, from Mar 2018 until July 2018]

2. Overall Objectives

2.1. Overall Objectives

The ROMA project aims at designing models, algorithms, and scheduling strategies to optimize the execution of scientific applications.

Scientists now have access to tremendous computing power. For instance, the four most powerful computing platforms in the TOP 500 list [60] each includes more than 500,000 cores and deliver a sustained performance of more than 10 Peta FLOPS. The volunteer computing platform BOINC [56] is another example with more than 440,000 enlisted computers and, on average, an aggregate performance of more than 9 Peta FLOPS. Furthermore, it had never been so easy for scientists to have access to parallel computing resources, either through the multitude of local clusters or through distant cloud computing platforms.

Because parallel computing resources are ubiquitous, and because the available computing power is so huge, one could believe that scientists no longer need to worry about finding computing resources, even less to optimize their usage. Nothing is farther from the truth. Institutions and government agencies keep building larger and more powerful computing platforms with a clear goal. These platforms must allow to solve problems in reasonable timescales, which were so far out of reach. They must also allow to solve problems more precisely where the existing solutions are not deemed to be sufficiently accurate. For those platforms to fulfill their purposes, their computing power must therefore be carefully exploited and not be wasted. This often requires an efficient management of all types of platform resources: computation, communication, memory, storage, energy, etc. This is often hard to achieve because of the characteristics of new and emerging platforms. Moreover, because of technological evolutions, new problems arise, and fully tried and tested solutions need to be thoroughly overhauled or simply discarded and replaced. Here are some of the difficulties that have, or will have, to be overcome:

- computing platforms are hierarchical: a processor includes several cores, a node includes several processors, and the nodes themselves are gathered into clusters. Algorithms must take this hierarchical structure into account, in order to fully harness the available computing power;
- the probability for a platform to suffer from a hardware fault automatically increases with the number of its components. Fault-tolerance techniques become unavoidable for large-scale platforms;
- the ever increasing gap between the computing power of nodes and the bandwidths of memories and networks, in conjunction with the organization of memories in deep hierarchies, requires to take more and more care of the way algorithms use memory;
- energy considerations are unavoidable nowadays. Design specifications for new computing platforms always include a maximal energy consumption. The energy bill of a supercomputer may represent a significant share of its cost over its lifespan. These issues must be taken into account at the algorithm-design level.

We are convinced that dramatic breakthroughs in algorithms and scheduling strategies are required for the scientific computing community to overcome all the challenges posed by new and emerging computing platforms. This is required for applications to be successfully deployed at very large scale, and hence for enabling the scientific computing community to push the frontiers of knowledge as far as possible. The ROMA project-team aims at providing fundamental algorithms, scheduling strategies, protocols, and software packages to fulfill the needs encountered by a wide class of scientific computing applications, including domains as diverse as geophysics, structural mechanics, chemistry, electromagnetism, numerical optimization, or computational fluid dynamics, to quote a few. To fulfill this goal, the ROMA project-team takes a special interest in dense and sparse linear algebra.

The work in the ROMA team is organized along three research themes.

1. **Algorithms for probabilistic environments.** In this theme, we consider problems where some of the platform characteristics, or some of the application characteristics, are described by probability distributions. This is in particular the case when considering the resilience of applications in failure-prone environments: the possibility of faults is modeled by probability distributions.
2. **Platform-aware scheduling strategies.** In this theme, we focus on the design of scheduling strategies that finely take into account some platform characteristics beyond the most classical ones, namely the computing speed of processors and accelerators, and the communication bandwidth of network links. In the scope of this theme, when designing scheduling strategies, we focus either on the energy consumption or on the memory behavior. All optimization problems under study are multi-criteria.
3. **High-performance computing and linear algebra.** We work on algorithms and tools for both sparse and dense linear algebra. In sparse linear algebra, we work on most aspects of direct multifrontal solvers for linear systems. In dense linear algebra, we focus on the adaptation of factorization kernels to emerging and future platforms. In addition, we also work on combinatorial scientific computing, that is, on the design of combinatorial algorithms and tools to solve combinatorial problems, such as those encountered, for instance, in the preprocessing phases of solvers of sparse linear systems.

3. Research Program

3.1. Algorithms for probabilistic environments

There are two main research directions under this research theme. In the first one, we consider the problem of the efficient execution of applications in a failure-prone environment. Here, probability distributions are used to describe the potential behavior of computing platforms, namely when hardware components are subject to faults. In the second research direction, probability distributions are used to describe the characteristics and behavior of applications.

3.1.1. Application resilience

An application is resilient if it can successfully produce a correct result in spite of potential faults in the underlying system. Application resilience can involve a broad range of techniques, including fault prediction, error detection, error containment, error correction, checkpointing, replication, migration, recovery, etc. Faults are quite frequent in the most powerful existing supercomputers. The Jaguar platform, which ranked third in the TOP 500 list in November 2011 [59], had an average of 2.33 faults per day during the period from August 2008 to February 2010 [84]. The mean-time between faults of a platform is inversely proportional to its number of components. Progresses will certainly be made in the coming years with respect to the reliability of individual components. However, designing and building high-reliability hardware components is far more expensive than using lower reliability top-of-the-shelf components. Furthermore, low-power components may not be available with high-reliability. Therefore, it is feared that the progresses in reliability will far from compensate the steady projected increase of the number of components in the largest supercomputers. Already, application failures have a huge computational cost. In 2008, the DARPA white paper on “System resilience at extreme scale” [58] stated that high-end systems wasted 20% of their computing capacity on application failure and recovery.

In such a context, any application using a significant fraction of a supercomputer and running for a significant amount of time will have to use some fault-tolerance solution. It would indeed be unacceptable for an application failure to destroy centuries of CPU-time (some of the simulations run on the Blue Waters platform consumed more than 2,700 years of core computing time [54] and lasted over 60 hours; the most time-consuming simulations of the US Department of Energy (DoE) run for weeks to months on the most powerful existing platforms [57]).

Our research on resilience follows two different directions. On the one hand we design new resilience solutions, either generic fault-tolerance solutions or algorithm-based solutions. On the other hand we model and theoretically analyze the performance of existing and future solutions, in order to tune their usage and help determine which solution to use in which context.

3.1.2. Scheduling strategies for applications with a probabilistic behavior

Static scheduling algorithms are algorithms where all decisions are taken before the start of the application execution. On the contrary, in non-static algorithms, decisions may depend on events that happen during the execution. Static scheduling algorithms are known to be superior to dynamic and system-oriented approaches in stable frameworks [65], [72], [73], [83], that is, when all characteristics of platforms and applications are perfectly known, known a priori, and do not evolve during the application execution. In practice, the prediction of application characteristics may be approximative or completely infeasible. For instance, the amount of computations and of communications required to solve a given problem in parallel may strongly depend on some input data that are hard to analyze (this is for instance the case when solving linear systems using full pivoting).

We plan to consider applications whose characteristics change dynamically and are subject to uncertainties. In order to benefit nonetheless from the power of static approaches, we plan to model application uncertainties and variations through probabilistic models, and to design for these applications scheduling strategies that are either static, or partially static and partially dynamic.

3.2. Platform-aware scheduling strategies

In this theme, we study and design scheduling strategies, focusing either on energy consumption or on memory behavior. In other words, when designing and evaluating these strategies, we do not limit our view to the most classical platform characteristics, that is, the computing speed of cores and accelerators, and the bandwidth of communication links.

In most existing studies, a single optimization objective is considered, and the target is some sort of absolute performance. For instance, most optimization problems aim at the minimization of the overall execution time of the application considered. Such an approach can lead to a very significant waste of resources, because it does not take into account any notion of efficiency nor of yield. For instance, it may not be meaningful to use twice as many resources just to decrease by 10% the execution time. In all our work, we plan to look only for algorithmic solutions that make a "clever" usage of resources. However, looking for the solution that optimizes a metric such as the efficiency, the energy consumption, or the memory-peak minimization, is doomed for the type of applications we consider. Indeed, in most cases, any optimal solution for such a metric is a sequential solution, and sequential solutions have prohibitive execution times. Therefore, it becomes mandatory to consider multi-criteria approaches where one looks for trade-offs between some user-oriented metrics that are typically related to notions of Quality of Service—execution time, response time, stretch, throughput, latency, reliability, etc.—and some system-oriented metrics that guarantee that resources are not wasted. In general, we will not look for the Pareto curve, that is, the set of all dominating solutions for the considered metrics. Instead, we will rather look for solutions that minimize some given objective while satisfying some bounds, or "budgets", on all the other objectives.

3.2.1. Energy-aware algorithms

Energy-aware scheduling has proven an important issue in the past decade, both for economical and environmental reasons. Energy issues are obvious for battery-powered systems. They are now also important for traditional computer systems. Indeed, the design specifications of any new computing platform now always include an upper bound on energy consumption. Furthermore, the energy bill of a supercomputer may represent a significant share of its cost over its lifespan.

Technically, a processor running at speed s dissipates s^α watts per unit of time with $2 \le \alpha \le 3$ [63], [64], [70]; hence, it consumes $s^\alpha \times d$ joules when operated during d units of time. Therefore, energy consumption can be reduced by using speed scaling techniques. However it was shown in [85] that reducing the speed of a processor increases the rate of transient faults in the system. The probability of faults increases exponentially, and this probability cannot be neglected in large-scale computing [81]. In order to make up for the loss in *reliability* due to the energy efficiency, different models have been proposed for fault tolerance: (i) *re-execution* consists in re-executing a task that does not meet the reliability constraint [85]; (ii) *replication* consists in executing the same task on several processors simultaneously, in order to meet the reliability constraints [62]; and (iii) *checkpointing* consists in "saving" the work done at some certain instants, hence reducing the amount of work lost when a failure occurs [80].

Energy issues must be taken into account at all levels, including the algorithm-design level. We plan to both evaluate the energy consumption of existing algorithms and to design new algorithms that minimize energy consumption using tools such as resource selection, dynamic frequency and voltage scaling, or powering-down of hardware components.

3.2.2. Memory-aware algorithms

For many years, the bandwidth between memories and processors has increased more slowly than the computing power of processors, and the latency of memory accesses has been improved at an even slower pace. Therefore, in the time needed for a processor to perform a floating point operation, the amount of data transferred between the memory and the processor has been decreasing with each passing year. The risk is for an application to reach a point where the time needed to solve a problem is no longer dictated by the processor computing power but by the memory characteristics, comparable to the *memory wall* that limits CPU performance. In such a case, processors would be greatly under-utilized, and a large part of the computing power of the platform would be wasted. Moreover, with the advent of multicore processors, the amount of memory per core has started to stagnate, if not to decrease. This is especially harmful to memory intensive applications. The problems related to the sizes and the bandwidths of memories are further exacerbated on modern computing platforms because of their deep and highly heterogeneous hierarchies. Such a hierarchy can extend from core private caches to shared memory within a CPU, to disk storage and even tape-based storage systems, like in the Blue Waters supercomputer [55]. It may also be the case that heterogeneous cores are used (such as hybrid CPU and GPU computing), and that each of them has a limited memory.

Because of these trends, it is becoming more and more important to precisely take memory constraints into account when designing algorithms. One must not only take care of the amount of memory required to run an algorithm, but also of the way this memory is accessed. Indeed, in some cases, rather than to minimize the amount of memory required to solve the given problem, one will have to maximize data reuse and, especially, to minimize the amount of data transferred between the different levels of the memory hierarchy (minimization of the volume of memory inputs-outputs). This is, for instance, the case when a problem cannot be solved by just using the in-core memory and that any solution must be out-of-core, that is, must use disks as storage for temporary data.

It is worth noting that the cost of moving data has lead to the development of so called "communication-avoiding algorithms" [76]. Our approach is orthogonal to these efforts: in communication-avoiding algorithms, the application is modified, in particular some redundant work is done, in order to get rid of some communication operations, whereas in our approach, we do not modify the application, which is provided as a task graph, but we minimize the needed memory peak only by carefully scheduling tasks.

3.3. High-performance computing and linear algebra

Our work on high-performance computing and linear algebra is organized along three research directions. The first direction is devoted to direct solvers of sparse linear systems. The second direction is devoted to combinatorial scientific computing, that is, the design of combinatorial algorithms and tools that solve problems encountered in some of the other research themes, like the problems faced in the preprocessing phases of sparse direct solvers. The last direction deals with the adaptation of classical dense linear algebra kernels to the architecture of future computing platforms.

3.3.1. Direct solvers for sparse linear systems

The solution of sparse systems of linear equations (symmetric or unsymmetric, often with an irregular structure, from a few hundred thousand to a few hundred million equations) is at the heart of many scientific applications arising in domains such as geophysics, structural mechanics, chemistry, electromagnetism, numerical optimization, or computational fluid dynamics, to cite a few. The importance and diversity of applications are a main motivation to pursue research on sparse linear solvers. Because of this wide range of applications, any significant progress on solvers will have a significant impact in the world of simulation. Research on sparse direct solvers in general is very active for the following main reasons:

- many applications fields require large-scale simulations that are still too big or too complicated with respect to today's solution methods;
- the current evolution of architectures with massive, hierarchical, multicore parallelism imposes to overhaul all existing solutions, which represents a major challenge for algorithm and software development;
- the evolution of numerical needs and types of simulations increase the importance, frequency, and size of certain classes of matrices, which may benefit from a specialized processing (rather than resort to a generic one).

Our research in the field is strongly related to the software package MUMPS, which is both an experimental platform for academics in the field of sparse linear algebra, and a software package that is widely used in both academia and industry. The software package MUMPS enables us to (i) confront our research to the real world, (ii) develop contacts and collaborations, and (iii) receive continuous feedback from real-life applications, which is extremely critical to validate our research work. The feedback from a large user community also enables us to direct our long-term objectives towards meaningful directions.

In this context, we aim at designing parallel sparse direct methods that will scale to large modern platforms, and that are able to answer new challenges arising from applications, both efficiently—from a resource consumption point of view—and accurately—from a numerical point of view. For that, and even with increasing parallelism, we do not want to sacrifice in any manner numerical stability, based on threshold partial pivoting, one of the main originalities of our approach (our "trademark") in the context of direct solvers for distributed-memory computers; although this makes the parallelization more complicated, applying the same pivoting strategy as in the serial case ensures numerical robustness of our approach, which we generally measure in terms of sparse backward error. In order to solve the hard problems resulting from the always-increasing demands in simulations, special attention must also necessarily be paid to memory usage (and not only execution time). This requires specific algorithmic choices and scheduling techniques. From a complementary point of view, it is also necessary to be aware of the functionality requirements from the applications and from the users, so that robust solutions can be proposed for a wide range of applications.

Among direct methods, we rely on the multifrontal method [74], [75], [79]. This method usually exhibits a good data locality and hence is efficient in cache-based systems. The task graph associated with the multifrontal method is in the form of a tree whose characteristics should be exploited in a parallel implementation.

Our work is organized along two main research directions. In the first one we aim at efficiently addressing new architectures that include massive, hierarchical parallelism. In the second one, we aim at reducing the running time complexity and the memory requirements of direct solvers, while controlling accuracy.

3.3.2. Combinatorial scientific computing

Combinatorial scientific computing (CSC) is a recently coined term (circa 2002) for interdisciplinary research at the intersection of discrete mathematics, computer science, and scientific computing. In particular, it refers to the development, application, and analysis of combinatorial algorithms to enable scientific computing applications. CSC's deepest roots are in the realm of direct methods for solving sparse linear systems of equations where graph theoretical models have been central to the exploitation of sparsity, since the 1960s. The general approach is to identify performance issues in a scientific computing problem, such as memory use, parallel speed up, and/or the rate of convergence of a method, and to develop combinatorial algorithms and models to tackle those issues.

Our target scientific computing applications are (i) the preprocessing phases of direct methods (in particular MUMPS), iterative methods, and hybrid methods for solving linear systems of equations, and general sparse matrix and tensor computations; and (ii) the mapping of tasks (mostly the sub-tasks of the mentioned solvers) onto modern computing platforms. We focus on the development and the use of graph and hypergraph models, and related tools such as hypergraph partitioning algorithms, to solve problems of load balancing and task mapping. We also focus on bipartite graph matching and vertex ordering methods for reducing the memory overhead and computational requirements of solvers. Although we direct our attention on these models and algorithms through the lens of linear system solvers, our solutions are general enough to be applied to some other resource optimization problems.

3.3.3. Dense linear algebra on post-petascale multicore platforms

The quest for efficient, yet portable, implementations of dense linear algebra kernels (QR, LU, Cholesky) has never stopped, fueled in part by each new technological evolution. First, the LAPACK library [67] relied on BLAS level 3 kernels (Basic Linear Algebra Subroutines) that enable to fully harness the computing power of a single CPU. Then the SCALAPACK library [66] built upon LAPACK to provide a coarse-grain parallel version, where processors operate on large block-column panels. Inter-processor communications occur through highly tuned MPI send and receive primitives. The advent of multi-core processors has led to a major modification in these algorithms [69], [82], [77]. Each processor runs several threads in parallel to keep all cores within that processor busy. Tiled versions of the algorithms have thus been designed: dividing large block-column panels into several tiles allows for a decrease in the granularity down to a level where many smaller-size tasks are spawned. In the current panel, the diagonal tile is used to eliminate all the lower tiles in the panel. Because the factorization of the whole panel is now broken into the elimination of several tiles, the update operations can also be partitioned at the tile level, which generates many tasks to feed all cores.

The number of cores per processor will keep increasing in the following years. It is projected that high-end processors will include at least a few hundreds of cores. This evolution will require to design new versions of libraries. Indeed, existing libraries rely on a static distribution of the work: before the beginning of the execution of a kernel, the location and time of the execution of all of its component is decided. In theory, static solutions enable to precisely optimize executions, by taking parameters like data locality into account. At run time, these solutions proceed at the pace of the slowest of the cores, and they thus require a perfect load-balancing. With a few hundreds, if not a thousand, cores per processor, some tiny differences between the computing times on the different cores ("jitter") are unavoidable and irremediably condemn purely static solutions. Moreover, the increase in the number of cores per processor once again mandates to increase the number of tasks that can be executed in parallel.

We study solutions that are part-static part-dynamic, because such solutions have been shown to outperform purely dynamic ones [71]. On the one hand, the distribution of work among the different nodes will still be statically defined. On the other hand, the mapping and the scheduling of tasks inside a processor will be dynamically defined. The main difficulty when building such a solution will be to design lightweight dynamic schedulers that are able to guarantee both an excellent load-balancing and a very efficient use of data locality.

4. Application Domains

4.1. Applications of sparse direct solvers

Sparse direct (e.g., multifrontal solvers that we develop) solvers have a wide range of applications as they are used at the heart of many numerical methods in computational science: whether a model uses finite elements or finite differences, or requires the optimization of a complex linear or nonlinear function, one often ends up solving a system of linear equations involving sparse matrices. There are therefore a number of application fields, among which some of the ones cited by the users of our sparse direct solver MUMPS are: structural mechanics, seismic modeling, biomechanics, medical image processing, tomography, geophysics, electromagnetism, fluid dynamics, econometric models, oil reservoir simulation, magneto-hydro-dynamics, chemistry, acoustics, glaciology, astrophysics, circuit simulation, and work on hybrid direct-iterative methods.

5. Highlights of the Year

5.1. Highlights of the Year

- Anne Benoit was the program chair of 32nd IEEE IPDPS conference (IEEE International Parallel & Distributed Processing Symposium), held in Vancouver, Canada, May 21–25, 2018.
- Bora Uçar was the general chair of 32nd IEEE IPDPS conference (IEEE International Parallel & Distributed Processing Symposium), held in Vancouver, Canada, May 21–25, 2018.

5.1.1. Awards

BEST PAPERS AWARDS :
[29]
T. HÉRAULT, Y. ROBERT, A. BOUTEILLER, D. ARNOLD, K. B. FERREIRA, G. BOSILCA, J. DONGARRA.*Optimal Cooperative Checkpointing for Shared High-Performance Computing Platforms*, in "APDCM", Vancouver, Canada, 2018, https://hal.inria.fr/hal-01968441

6. New Software and Platforms

6.1. MUMPS

A MUltifrontal Massively Parallel Solver
KEYWORDS: High-Performance Computing - Direct solvers - Finite element modelling

FUNCTIONAL DESCRIPTION: MUMPS is a software library to solve large sparse linear systems (AX=B) on sequential and parallel distributed memory computers. It implements a sparse direct method called the multifrontal method. It is used worldwide in academic and industrial codes, in the context numerical modeling of physical phenomena with finite elements. Its main characteristics are its numerical stability, its large number of features, its high performance and its constant evolution through research and feedback from its community of users. Examples of application fields include structural mechanics, electromagnetism, geophysics, acoustics, computational fluid dynamics. MUMPS is developed by INPT(ENSEEIHT)-IRIT, Inria, CERFACS, University of Bordeaux, CNRS and ENS Lyon. In 2014, a consortium of industrial users has been created (http://mumps-consortium.org).

RELEASE FUNCTIONAL DESCRIPTION: MUMPS versions 5.1.0, 5.1.1 and 5.1.2, all released in 2017 include many new features and improvements. The two main new features are Block Low-Rank compression, decreasing the complexity of sparse direct solvers for various types of applications, and selective 64-bit integers, allowing to process matrices with more than 2 billion entries. Several new features have been developed in 2017 and 2018 that are included in some MUMPS versions provided to partners for experimentation (e.g. in the context of industrial contracts). These features will appear in the future public versions, starting with MUMPS 5.2.0.

- Participants: Gilles Moreau, Abdou Guermouche, Alfredo Buttari, Aurélia Fevre, Bora Uçar, Chiara Puglisi, Clément Weisbecker, Emmanuel Agullo, François-Henry Rouet, Guillaume Joslin, Jacko Koster, Jean-Yves L'excellent, Marie Durand, Maurice Bremond, Mohamed Sid-Lakhdar, Patrick Amestoy, Philippe Combes, Stéphane Pralet, Theo Mary and Tzvetomila Slavova
- Partners: Université de Bordeaux - CNRS - CERFACS - ENS Lyon - INPT - IRIT - Université de Lyon - Université de Toulouse - LIP
- Contact: Jean-Yves L'excellent
- URL: http://mumps-solver.org/

7. New Results

7.1. Birkhoff–von Neumann decomposition

The well-known Birkhoff-von Neumann (BvN) decomposition expresses a doubly stochastic matrix as a convex combination of a number of permutation matrices. For a given doubly stochastic matrix, there are many BvN decompositions, and finding the one with the minimum number of permutation matrices is NP-hard. There are heuristics to obtain BvN decompositions for a given doubly stochastic matrix. A family of heuristics are based on the original proof of Birkhoff and proceed step by step by subtracting a scalar multiple of a permutation matrix at each step from the current matrix, starting from the given matrix. At every step, the subtracted matrix contains nonzeros at the positions of some nonzero entries of the current matrix and annihilates at least one entry, while keeping the current matrix nonnegative. Our first result, which supports a claim of Brualdi [68], shows that this family of heuristics can miss optimal decompositions. We also investigate the performance of two heuristics from this family theoretically. The findings are published in a journal [10].

7.2. Parallel sparse matrix-vector multiply

There are three common parallel sparse matrix-vector multiply algorithms: 1D 3 row-parallel, 1D column-parallel and 2D row-column-parallel. The 1D parallel algorithms offer the advantage of having only one communication phase. On the other hand, the 2D parallel algorithm is more scalable but it suffers from two communication phases. In this work, we introduce a novel concept of heterogeneous messages where a heterogeneous message may contain both input-vector entries and partially computed output-vector entries. This concept not only leads to a decreased number of messages, but also enables fusing the input-and output-communication phases into a single phase. These findings are exploited to propose a 1.5D parallel sparse matrix-vector multiply algorithm which is called local row-column-parallel. This proposed algorithm requires a constrained fine-grain partitioning in which each fine-grain task is assigned to the processor that contains either its input-vector entry, or its output-vector entry, or both. We propose two methods to carry out the constrained fine-grain partitioning. We conduct our experiments on a large set of test matrices to evaluate the partitioning qualities and partitioning times of these proposed 1.5D methods. The findings are published in a journal [14].

7.3. Scheduling series-parallel task graphs to minimize peak memory

We consider a variant of the well-known, NP-complete problem of minimum cut linear arrangement for directed acyclic graphs. In this variant, we are given a directed acyclic graph and we are asked to find a topological ordering such that the maximum number of cut edges at any point in this ordering is minimum. In our variant, the vertices and edges have weights, and the aim is to minimize the maximum weight of cut edges in addition to the weight of the last vertex before the cut. There is a known, polynomial time algorithm [78] for the cases where the input graph is a rooted tree. We focus on the instances where the input graph is a directed series-parallel graph, and propose a polynomial time algorithm, thus expanding the class of graphs for which a polynomial time algorithm is known. Directed acyclic graphs are used to model scientific applications where the vertices correspond to the tasks of a given application and the edges represent the dependencies between the tasks. In such models, the problem we address reads as minimizing the peak memory requirement in an execution of the application. Our work, combined with Liu's work on rooted trees addresses this practical problem in two important classes of applications. The findings are published in a journal [15].

7.4. Parallel Candecomp/Parafac decomposition of sparse tensors using dimension trees

Tensor factorization has been increasingly used to address various problems in many fields such as signal processing, data compression, computer vision, and computational data analysis. CANDECOMP/PARAFAC (CP) decomposition of sparse tensors has successfully been applied to many well-known problems in web search, graph analytics, recommender systems, health care data analytics, and many other domains. In these applications, computing the CP decomposition of sparse tensors efficiently is essential in order to be able to process and analyze data of massive scale. For this purpose, we investigate an efficient computation and parallelization of the CP decomposition for sparse tensors. We provide a novel computational scheme for reducing the cost of a core operation in computing the CP decomposition with the traditional alternating least squares (CP-ALS) based algorithm. We then effectively parallelize this computational scheme in the context of CP-ALS in shared and distributed memory environments, and propose data and task distribution models for better scalability. We implement parallel CP-ALS algorithms and compare our implementations with an efficient tensor factorization library, using tensors formed from real-world and synthetic datasets. With our algorithmic contributions and implementations, we report up to 3.95x, 3.47x, and 3.9x speedups in sequential, shared memory parallel, and distributed memory parallel executions over the state of the art, and up to 1466x overall speedup over the sequential execution using 4096 cores on an IBM BlueGene/Q supercomputer. The findings are published in a journal [13].

7.5. Approximation algorithms for maximum matchings in undirected graphs

We propose heuristics for approximating the maximum cardinality matching on undirected graphs. Our heuristics are based on the theoretical body of a certain type of random graphs, and are made practical for real-life ones. The idea is based on judiciously selecting a subgraph of a given graph and obtaining a maximum cardinality matching on this subgraph. We show that the heuristics have an approximation guarantee of around $0.866 - \log(n)/n$ for a graph with n vertices. Experiments for verifying the theoretical results in practice are provided. The findings are published in a conference proceedings [25].

7.6. SINA: A Scalable iterative network aligner

Given two graphs, network alignment asks for a potentially partial mapping between the vertices of the two graphs. This arises in many applications where data from different sources need to be integrated. Recent graph aligners use the global structure of input graphs and additional information given for the edges and vertices. We present SINA, an efficient, shared memory parallel implementation of such an aligner. Our experimental evaluations on a 32-core shared memory machine showed that SINA scales well for aligning large real-world graphs: SINA can achieve up to $28.5\times$ speedup, and can reduce the total execution time of a graph alignment problem with 2M vertices and 100M edges from 4.5 hours to under 10 minutes. To the best of our knowledge, SINA is the first parallel aligner that uses global structure and vertex and edge attributes to handle large graphs. The findings are published in a conference proceedings [34].

7.7. Acyclic partitioning of large directed acyclic graphs

We investigate the problem of partitioning the vertices of a directed acyclic graph into a given number of parts. The objective function is to minimize the number or the total weight of the edges having end points in different parts, which is also known as edge cut. The standard load balancing constraint of having an equitable partition of the vertices among the parts should be met. Furthermore, the partition is required to be acyclic, i.e., the inter-part edges between the vertices from different parts should preserve an acyclic dependency structure among the parts. In this work, we adopt the multilevel approach with coarsening, initial partitioning, and refinement phases for acyclic partitioning of directed acyclic graphs. We focus on two-way partitioning (sometimes called bisection), as this scheme can be used in a recursive way for multi-way partitioning. To ensure the acyclicity of the partition at all times, we propose novel and efficient coarsening and refinement heuristics. The quality of the computed acyclic partitions is assessed by computing the edge cut. We also propose effective ways to use the standard undirected graph partitioning methods in our multilevel scheme. We perform a large set of experiments on a dataset consisting of (i) graphs coming from an application and (ii) some others corresponding to matrices from a public collection. We report improvements, on average, around 59% compared to the current state of the art. The findings are published in a research report [50].

7.8. Effective heuristics for matchings in hypergraphs

The problem of finding a maximum cardinality matching in a d-partite d-uniform hypergraph is an important problem in combinatorial optimization and has been theoretically analyzed by several researchers. In this work, we first devise heuristics for this problem by generalizing the existing cheap graph matching heuristics. Then, we propose a novel heuristic based on tensor scaling to extend the matching via judicious hyperedge selections. Experiments on random, synthetic and real-life hypergraphs show that this new heuristic is highly practical and superior to the others on finding a matching with large cardinality. The findings are published in a research report [46].

7.9. Scaling matrices and counting the perfect matchings in graphs

We investigate efficient randomized methods for approximating the number of perfect matchings in bipartite graphs and general graphs. Our approach is based on assigning probabilities to edges. The findings are published in a research report [47].

7.10. A scalable clustering-based task scheduler for homogeneous processors using DAG partitioning

When scheduling a directed acyclic graph (DAG) of tasks on computational platforms, a good trade-off between load balance and data locality is necessary. List-based scheduling techniques are commonly used greedy approaches for this problem. The downside of list-scheduling heuristics is that they are incapable of making short-term sacrifices for the global efficiency of the schedule. In this work, we describe new list-based scheduling heuristics based on clustering for homogeneous platforms. Our approach uses an acyclic partitioner for DAGs for clustering. The clustering enhances the data locality of the scheduler with a global view of the graph. Furthermore, since the partition is acyclic, we can schedule each part completely once its input tasks are ready to be executed. We present an extensive experimental evaluation showing the trade-offs between the granularity of clustering and the parallelism, and how this affects the scheduling. Furthermore, we compare our heuristics to the best state-of-the-art list-scheduling and clustering heuristics, and obtain better performance in cases with many communications. The findings are published in a research report [53].

7.11. Data-Locality Aware Dynamic Schedulers for Independent Tasks with Replicated Inputs

In this work we concentrate on a crucial parameter for efficiency in Big Data and HPC applications: data locality. We focus on the scheduling of a set of independant tasks, each depending on an input file. We assume that each of these input files has been replicated several times and placed in local storage of different nodes of a cluster, similarly of what we can find on HDFS system for example. We consider two optimization problems, related to the two natural metrics: makespan optimization (under the constraint that only local tasks are allowed) and communication optimization (under the constraint of never letting a processor idle in order to optimize makespan). For both problems we investigate the performance of dynamic schedulers, in particular the basic greedy algorithm we can for example find in the default MapReduce scheduler. First we theoretically study its performance, with probabilistic models, and provide a lower bound for communication metric and asymptotic behaviour for both metrics. Second we propose simulations based on traces from a Hadoop cluster to compare the different dynamic schedulers and assess the expected behaviour obtained with the theoretical study.

These findings have been presented at the CEBDA workshop [19].

7.12. Parallel scheduling of DAGs under memory constraints.

Scientific workflows are frequently modeled as Directed Acyclic Graphs (DAG) of tasks, which represent computational modules and their dependencies, in the form of data produced by a task and used by another one. This formulation allows the use of runtime systems which dynamically allocate tasks onto the resources of increasingly complex and heterogeneous computing platforms. However, for some workflows, such a dynamic schedule may run out of memory by exposing too much parallelism. This work focuses on the problem of transforming such a DAG to prevent memory shortage, and concentrates on shared memory platforms. We first propose a simple model of DAG which is expressive enough to emulate complex memory behaviors. We then exhibit a polynomial-time algorithm that computes the maximum peak memory of a DAG, that is, the maximum memory needed by any parallel schedule. We consider the problem of reducing this maximum peak memory to make it smaller than a given bound by adding new fictitious edges, while trying to minimize the critical path of the graph. After proving this problem NP-complete, we provide an ILP solution as well as several heuristic strategies that are thoroughly compared by simulation on synthetic DAGs modeling actual computational workflows. We show that on most instances, we are able to decrease the maximum peak memory at the cost of a small increase in the critical path, thus with little impact on quality of the final parallel schedule.

This work has been presented at the IPDPS 2018 conference [31] and an extended version has been submitted to the Elsevier JPDC journal [52].

7.13. Online Scheduling of Task Graphs on Hybrid Platforms.

Modern computing platforms commonly include accelerators. We target the problem of scheduling applications modeled as task graphs on hybrid platforms made of two types of resources, such as CPUs and GPUs. We consider that task graphs are uncovered dynamically, and that the scheduler has information only on the available tasks, i.e., tasks whose predecessors have all been completed. Each task can be processed by either a CPU or a GPU, and the corresponding processing times are known. Our study extends a previous $4\sqrt{m/k}$-competitive online algorithm [61], where m is the number of CPUs and k the number of GPUs ($m \geq k$). We prove that no online algorithm can have a competitive ratio smaller than $\sqrt{m/k}$. We also study how adding flexibility on task processing, such as task migration or spoliation, or increasing the knowledge of the scheduler by providing it with information on the task graph, influences the lower bound. We provide a $(2\sqrt{m/k}+1)$-competitive algorithm as well as a tunable combination of a system-oriented heuristic and a competitive algorithm; this combination performs well in practice and has a competitive ratio in $\Theta(\sqrt{m/k})$. Finally, simulations on different sets of task graphs illustrate how the instance properties impact the performance of the studied algorithms and show that our proposed tunable algorithm performs the best among the online algorithms in almost all cases and has even performance close to an offline algorithm.

This work has been presented at the EuroPar 2018 conference [24].

7.14. Memory-aware tree partitioning on homogeneous platforms

Scientific applications are commonly modeled as the processing of directed acyclic graphs of tasks, and for some of them, the graph takes the special form of a rooted tree. This tree expresses both the computational dependencies between tasks and their storage requirements. The problem of scheduling/traversing such a tree on a single processor to minimize its memory footprint has already been widely studied. Hence, we move to parallel processing and study how to partition the tree for a homogeneous multiprocessor platform, where each processor is equipped with its own memory. We formally state the problem of partitioning the tree into subtrees such that each subtree can be processed on a single processor and the total resulting processing time is minimized. We prove that the problem is NP-complete, and we design polynomial-time heuristics to address it. An extensive set of simulations demonstrates the usefulness of these heuristics.

This work has been presented as a short paper in the PDP 2018 conference [27].

7.15. Reliability-aware energy optimization for throughput-constrained applications on MPSoC.

Multi-Processor System-on-Chip (MPSoC) has emerged as a promising platform to meet the increasing performance demand of embedded applications. However, due to limited energy budget, it is hard to guarantee that applications on MPSoC can be accomplished on time with a required throughput. The situation becomes even worse for applications with high reliability requirements, since extra energy will be inevitably consumed by task re-executions or duplicated tasks. Based on Dynamic Voltage and Frequency Scaling (DVFS) and task duplication techniques, this paper presents a novel energy-efficient scheduling model, which aims at minimizing the overall energy consumption of MPSoC applications under both throughput and reliability constraints. The problem is shown to be NP-complete, and several polynomial-time heuristics are proposed to tackle this problem. Comprehensive simulations on both synthetic and real application graphs show that our proposed heuristics can meet all the given constraints, while reducing the energy consumption.

This findings have been presented at the ICPADS 2018 conference [26].

7.16. Malleable task-graph scheduling with a practical speed-up model

Scientific workloads are often described by Directed Acyclic task Graphs. Indeed, DAGs represent both a theoretical model and the structure employed by dynamic runtime schedulers to handle HPC applications. A natural problem is then to compute a makespan-minimizing schedule of a given graph. In this paper, we are motivated by task graphs arising from multifrontal factorizations of sparse matrices and therefore work under the following practical model. Tasks are malleable (i.e., a single task can be allotted a time-varying number of processors) and their speedup behaves perfectly up to a first threshold, then speedup increases linearly, but not perfectly, up to a second threshold where the speedup levels off and remains constant.

After proving the NP-hardness of minimizing the makespan of DAGs under this model, we study several heuristics. We propose model-optimized variants for PROPSCHEDULING, widely used in linear algebra application scheduling, and FLOWFLEX. GREEDYFILLING is proposed, a novel heuristic designed for our speedup model, and we demonstrate that PROPSCHEDULING and GREEDYFILLING are 2-approximation algorithms. In the evaluation, employing synthetic data sets and task graphs arising from multifrontal factorization, the proposed optimized variants and GREEDYFILLING significantly outperform the traditional algorithms, whereby GREEDYFILLING demonstrates a particular strength for balanced graphs.

These findings have been published in the IEEE TPDS journal [16].

7.17. Performance and scalability of the block low-rank multifrontal factorization on multicore architectures

Matrices coming from elliptic Partial Differential Equations have been shown to have a low-rank property which can be efficiently exploited in multifrontal solvers to provide a substantial reduction of their complexity. Among the possible low-rank formats, the Block Low-Rank format (BLR) is reasonably easy to use in a general purpose multifrontal solver and its potential compared to standard (full-rank) solvers has been demonstrated. Recently, new variants have been introduced and it was proved that they can further reduce the complexity but their performance remained to be analyzed. We develop a multithreaded BLR factorization, and analyze its efficiency and scalability in shared-memory multicore environments. We identify the challenges posed by the use of BLR approximations in multifrontal solvers and put forward several algorithmic variants of the BLR factorization that overcome these challenges by improving its efficiency and scalability. We illustrate the performance analysis of the BLR multifrontal factorization with numerical experiments on a large set of problems coming from a variety of real-life applications.

This work has been accepted for publication in the ACM Transactions on Mathematical Software [5].

7.18. On exploiting sparsity of multiple right-hand sides in sparse direct solvers

The cost of the solution phase in sparse direct methods is sometimes critical. It can be larger than that of the factorization in applications where systems of linear equations with thousands of right-hand sides (RHS) must be solved. In this work, we focus on the case of multiple sparse RHS with different nonzero structures in each column. In this setting, vertical sparsity reduces the number of operations by avoiding computations on rows that are entirely zero, and horizontal sparsity goes further by performing each elementary solve operation only on a subset of the RHS columns. To maximize the exploitation of horizontal sparsity, we propose a new algorithm to build a permutation of the RHS columns. We then propose an original approach to split the RHS columns into a minimal number of blocks, while reducing the number of operations down to a given threshold. Both algorithms are motivated by geometric intuitions and designed using an algebraic approach, so that they can be applied to general systems. We demonstrate the effectiveness of our algorithms on systems coming from real applications and compare them to other standard approaches. We also give some perspectives and possible applications.

This work has been accepted for publication in the SIAM Journal on Scientific Computing [6].

7.19. Efficient use of sparsity by direct solvers applied to 3D controlled-source EM problems

Controlled-source electromagnetic (CSEM) surveying becomes a widespread method for oil and gas exploration, which requires fast and efficient software for inverting large-scale EM datasets. In this context, one often needs to solve sparse systems of linear equations with a *large* number of *sparse* right-hand sides, each corresponding to a given transmitter position. Sparse direct solvers are very attractive for these problems, especially when combined with low-rank approximations which significantly reduce the complexity and the cost of the factorization. In the case of thousands of right-hand sides, the time spent in the sparse triangular solve tends to dominate the total simulation time and here we propose several approaches to reduce it. A significant reduction is demonstrated for marine CSEM application by utilizing the sparsity of the right-hand sides (RHS) and of the solutions that results from the geometry of the problem. Large gains are achieved by restricting computations at the forward substitution stage to exploit the fact that the RHS matrix might have empty rows (*vertical sparsity*) and/or empty blocks of columns within a non-empty row (*horizontal sparsity*). We also adapt the parallel algorithms that were designed for the factorization to solve-oriented algorithms and describe performance optimizations particularly relevant for the very large numbers of right-hand sides of the

CSEM application. We show that both the operation count and the elapsed time for the solution phase can be significantly reduced. The total time of CSEM simulation can be divided by approximately a factor of 3 on all the matrices from our set (from 3 to 30 million unknowns, and from 4 to 12 thousands RHSs).

These findings are described in a technical report [37] and will be submitted for publication.

7.20. A Generic Approach to Scheduling and Checkpointing Workflows

We dealt with scheduling and checkpointing strategies to execute scientific workflows on failure-prone large-scale platforms. To the best of our knowledge, this work was the first to target fail-stop errors for arbitrary workflows. Most previous work addresses soft errors, which corrupt the task being executed by a processor but do not cause the entire memory of that processor to be lost, contrarily to fail-stop errors. We revisited classical mapping heuristics such as HEFT and MINMIN and complement them with several checkpointing strategies. The objective was to derive an efficient trade-off between checkpointing every task (CKPTALL), which is an overkill when failures are rare events, and checkpointing no task (CKPTNONE), which induces dramatic re-execution overhead even when only a few failures strike during execution. Contrarily to previous work, our approach applies to arbitrary workflows, not just special classes of dependence graphs such as M-SPGs (Minimal Series-Parallel Graphs). Extensive experiments report significant gain over both CKPTALL and CKPTNONE, for a wide variety of workflows.

This findings have been presented at the ICPP 2018 conference [28].

7.21. Scheduling independent stochastic tasks under deadline and budget constraints

We studied scheduling strategies for the problem of maximizing the expected number of tasks that can be executed on a cloud platform within a given budget and under a deadline constraint. The execution times of tasks follow IID probability laws. The main questions are how many processors to enroll and whether and when to interrupt tasks that have been executing for some time. We provide complexity results and an asymptotically optimal strategy for the problem instance with discrete probability distributions and without deadline. We extend the latter strategy for the general case with continuous distributions and a deadline and we design an efficient heuristic which is shown to outperform standard approaches when running simulations for a variety of useful distribution laws.

This findings have been presented at the SBAC-PAD 2018 conference [23].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- In 2018, in the context of the MUMPS consortium (http://mumps-consortium.org), we worked in close collaboration with Toulouse INP to:
  - sign or renew membership contracts with AIRBUS, FFT-MSC, and SHELL, on top of the ongoing contracts with EDF, ALTAIR, Michelin, LSTC, Siemens, ESI Group, Total, SAFRAN, LBNL,
  - organize point-to-point meetings with several members,
  - provide technical support and scientific advice to members,
  - provide experimental releases to members in advance,
  - organize the fourth consortium committee meeting, at SAFRAN (Saclay).

Three engineers have been funded by the membership fees in 2018, for software engineering and software development, performance study and tuning on modern architectures, business development, management of the consortium, and organization of the future of the consortium. Half a year of a PhD student was also funded by the membership fees (see Section 9.1). On top of their membership, an additional contract was finalized with Michelin to study a new functionality and understand how to best exploit MUMPS recent features in their computing environment.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. PhD grant laboratoire d'excellence MILYON-Mumps consortium

The doctoral program from Labex MILYON dedicated to applied research in collaboration with industrial partners funded 50% of a 3-year PhD grant (the other 50% being funded by the MUMPS consortium) to work on improvements of the solution phase of the MUMPS solver. The PhD aimed at answering industrial needs in application domains where the cost of the solution phase of sparse direct solvers is critical. The PhD was defended on December 10, 2018 [2].

9.2. National Initiatives

9.2.1. ANR

ANR Project SOLHAR (2013-2018), 4,5 years. The ANR Project SOLHAR was launched in November 2013, for a duration of 48 months. It gathers five academic partners (the HiePACS, Cepage, ROMA and Runtime Inria project-teams, and CNRS-IRIT) and two industrial partners (CEA/CESTA and EADS-IW). This project aims at studying and designing algorithms and parallel programming models for implementing direct methods for the solution of sparse linear systems on emerging computers equipped with accelerators.

The proposed research is organized along three distinct research thrusts. The first objective deals with linear algebra kernels suitable for heterogeneous computing platforms. The second one focuses on runtime systems to provide efficient and robust implementation of dense linear algebra algorithms. The third one is concerned with scheduling this particular application on a heterogeneous and dynamic environment.

9.3. International Initiatives

9.3.1. Inria International Labs

9.3.1.1. JLESC — Joint Laboratory on Extreme Scale Computing

The University of Illinois at Urbana-Champaign, Inria, the French national computer science institute, Argonne National Laboratory, Barcelona Supercomputing Center, Jülich Supercomputing Centre and the Riken Advanced Institute for Computational Science formed the Joint Laboratory on Extreme Scale Computing, a follow-up of the Inria-Illinois Joint Laboratory for Petascale Computing. The Joint Laboratory is based at Illinois and includes researchers from Inria, and the National Center for Supercomputing Applications, ANL, BSC and JSC. It focuses on software challenges found in extreme scale high-performance computers.

Research areas include:

- Scientific applications (big compute and big data) that are the drivers of the research in the other topics of the joint-laboratory.
- Modeling and optimizing numerical libraries, which are at the heart of many scientific applications.
- Novel programming models and runtime systems, which allow scientific applications to be updated or reimagined to take full advantage of extreme-scale supercomputers.
- Resilience and Fault-tolerance research, which reduces the negative impact when processors, disk drives, or memory fail in supercomputers that have tens or hundreds of thousands of those components.
- I/O and visualization, which are important part of parallel execution for numerical silulations and data analytics
- HPC Clouds, that may execute a portion of the HPC workload in the near future.

Several members of the ROMA team are involved in the JLESC joint lab through their research on scheduling and resilience. Yves Robert is the Inria executive director of JLESC.

9.3.2. Inria Associate Teams Not Involved in an Inria International Labs

9.3.2.1. Keystone

Title: Scheduling algorithms for sparse linear algebra at extreme scale

International Partner (Vanderbilt University - Department of Electrical Engineering and Computer Science - Padma Raghavan):

Start year: 2016

See also: http://graal.ens-lyon.fr/~abenoit/Keystone

The Keystone project aims at investigating sparse matrix and graph problems on NUMA multicores and/or CPU-GPU hybrid models. The goal is to improve the performance of the algorithms, while accounting for failures and trying to minimize the energy consumption. The long-term objective is to design robust sparse-linear kernels for computing at extreme scale. In order to optimize the performance of these kernels, we plan to take particular care of locality and data reuse. Finally, there are several real-life applications relying on these kernels, and the Keystone project is assessing the performance and robustness of the scheduling algorithms in applicative contexts.

9.3.3. Inria International Partners

9.3.3.1. Declared Inria International Partners

- Anne Benoit, Frederic Vivien and Yves Robert have a regular collaboration with Henri Casanova from Hawaii University (USA). This is a follow-on of the Inria Associate team that ended in 2014.

9.3.4. Cooperation with ECNU

ENS Lyon has launched a partnership with ECNU, the East China Normal University in Shanghai, China. This partnership includes both teaching and research cooperation.

As for teaching, the PROSFER program includes a joint Master of Computer Science between ENS Rennes, ENS Lyon and ECNU. In addition, PhD students from ECNU are selected to conduct a PhD in one of these ENS. Yves Robert is responsible for this cooperation. He has already given two classes at ECNU, on Algorithm Design and Complexity, and on Parallel Algorithms, together with Patrice Quinton (from ENS Rennes).

As for research, the JORISS program funds collaborative research projects between ENS Lyon and ECNU. Anne Benoit and Minsong Chen are leading a JORISS project on scheduling and resilience in cloud computing. Frédéric Vivien and Jing Liu (ECNU) are leading a JORISS project on resilience for real-time applications. In the context of this collaboration two students from ECNU, Li Han and Changjiang Gou, have joined Roma for their PhD.

9.4. International Research Visitors

9.4.1. Visits to International Teams

9.4.1.1. Research Stays Abroad

- Yves Robert has been appointed as a visiting scientist by the ICL laboratory (headed by Jack Dongarra) at the University of Tennessee Knoxville since 2011. He collaborates with several ICL researchers on high-performance linear algebra and resilience methods at scale.
- Anne Benoit and Bora Uçar visited the School of Computational Science and Engineering Georgia Institute of Technology, Atlanta, GA, USA. During their stay August 2017–June 2018, they worked with the research group of Prof. Umit V. Çatalyürek.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Bora Uçar was the general chair of 32nd IEEE IPDPS 2018 (IEEE International Parallel & Distributed Processing Symposium), held in Vancouver, Canada, May 21–25, 2018.

10.1.1.2. Member of the Organizing Committees

- Bora Uçar was a member of the organizing committee of ICGT 2018 (10th International Colloquium on Graph Theory and combinatorics), held in Lyon, July 9–13, 2018

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- Anne Benoit was the program chair of 32nd IEEE IPDPS 2018 (IEEE International Parallel & Distributed Processing Symposium), held in Vancouver, Canada, May 21–25, 2018. She was also the global chair for topic 3: "Scheduling and Load Balancing" of the 24th Int. European Conf. on Parallel and Distributed Computing (EuroPar 2018), held in Torino, Italy, August 27–31, 2018.

10.1.2.2. Member of the Conference Program Committees

- Bora Uçar was a member of the program committee of **IA**^{3}, 2018 The Eight Workshop on Irregular Applications: Architectures and Algorithms, in conjunction with SC'18, November 11–16, 2018, Dallas, Texas, USA; **CSC18**, The 8th SIAM Workshop on Combinatorial Scientific Computing, Bergen, Norway June 6-8, 2018; **HiPC 2018**; 25th IEEE International Conference on High Performance Computing, Data, and Analytics, Bengaluru, India, 17–20 December 2018; **SC18** Doctoral Showcase of The International Conference for High Performance Computing, Networking, Storage, and Analysis, Dallas, TX, USA; **33rd IEEE IPDPS 2019 Workshops Committee**, 33rd IEEE International Parallel and Distributed Processing Symposium, Rio de Janeiro, Brazil, May 20–24, 2019.
- Loris Marchal was a member of the program committee of **IPDPS 2018** , **ICPP 2018** and the workshop of IPDPS **APDCM 2018**.
- Jean-Yves L'Excellent was a member of the program committee of **CSC18**, The 8th SIAM Workshop on Combinatorial Scientific Computing, Bergen, Norway June 6-8, 2018.
- Frédéric Vivien was a member of the program committee of **IPDPS 2018**, **PDP 2018**; **EduPar 18**, and the Poster session of **SC18**.
- Yves Robert was a member of the program committee of the FTXS, Scala and PMBS workshops co-located with SC'18 in Dallas, TX.

10.1.2.3. Reviewer

Bora Uçar reviewed a paper for 33rd IEEE IPDPS 2019.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Anne Benoit is Associate Editor (in Chief) of ParCo, the journal of Parallel Computing: Systems and Applications (from July 2018). She is also a member of the editorial board (Associate Editor) of TPDS, IEEE Transactions on Parallel and Distributed Systems since 2015, and of JPDC, the Journal of Parallel and Distributed Computing, since 2011.
- Bora Uçar is a member of the editorial board of Parallel Computing, April 2016–on going, and SIAM Journal on Matrix Analysis and Applications (SIMAX), May 2018–ongoing.
- Frédéric Vivien is Associate Editor of Parallel Computing (Elsevier) and of JPDC (Elsevier Journal of Parallel and Distributed Computing).
- Yves Robert is Associate Editor of JPDC (Elsevier Journal of Parallel and Distributed Computing) and TOPC (ACM Trans. On Parallel Computing).

10.1.3.2. Reviewer - Reviewing Activities

Bora Uçar reviewed papers for the journals SIAM Journal on Scientific Computing (4 in 2018); ACM Transactions on Mathematical Software (2 in 2018); IEEE Transactions on Parallel and Distributed Systems (1 in 2018); Future Generation Computer Systems (1 in 2018); IEEE Transactions on Signal Processing (1 in 2018); SIAM Journal on Matrix Analysis and applications (1 in 2018);

Anne Benoit, Loris Marchal, Yves Robert and Frédéric Vivien reviewed papers for the journals IEEE Transactions on Parallel and Distributed Systems and Elsevier Journal of Parallel and Distributed Computing.

10.1.4. Invited Talks

- Bora Uçar delivered an invited talk at the Scientific Computing Group's Seminar at the Emory University, Atlanta, USA, September 2017.
- Frédéric Vivien delivered the keynote presentation of the 8th IEEE Workshop PDCO, held in conjunction with IPDPS 2018, in Vancouver, Canada, on Monday May 21, 2018.
- Yves Robert delivered a keynote presentation at SBAC-PAD'2018, the 30th International Symposium on Computer. Architecture and High Performance Computing.
- Yves Robert delivered the keynote presentation at SCALA'2018, the 9th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems, held in conjunction with SC'18

10.1.5. Leadership within the Scientific Community

- Anne Benoit is a member of the Steering Committee of HCW (Heterogeneity in Computing Workshop, co-located with IPDPS) since 2018.
- Yves Robert is a member of the Steering Committee of IPDPS and HCW . He is the liaison between the Steering and Program committees of IPDPS.
- Bora Uçar is a member of the Steering Committee of Combinatorial Scientific Computing (2014–on going); and IPDPS for the years 2017–2019. He is also a vice-chair of IEEE Technical Committee on Parallel Processing (TCPP).

10.1.6. Scientific Expertise

Yves Robert is an expert for the Horizon 2020 program of the European Commission and has reviews two projects in 2018.

10.1.7. Research Administration

Loris Marchal is responsible of the competitive selection of ENS Lyon Student for Computer Science.

Frédéric Vivien is the vice-head of the LIP laboratory since September 2017. He is a member of the scientific council of the École normale supérieure de Lyon and of the academic council of the University of Lyon.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence: Anne Benoit, Responsible of the L3 students at ENS Lyon, France

Licence: Yves Robert, Algorithmique, ENS Lyon, France

Master: Anne Benoit, Parallel and Distributed Algorithms and Programs, 42, M1, ENS Lyon, France

Master : Bora Uçar, Combinatorial Scientific Computing (with Fanny Dufossé), 36, M2 Informatique Fondamentale, ENS Lyon, France.

Master: Yves Robert, Scheduling at scale, 36, M2 Informatique Fondamentale, ENS Lyon, France

Master: Yves Robert, Responsible of M2 Informatique Fondamentale, ENS Lyon, France

Master : Loris Marchal, Complexity and calculability (practicals), 16, M1, Univ. Lyon 1, France.

10.2.2. Supervision

HdR: Loris Marchal, Memory and data aware scheduling, ENS Lyon, March 30, 2018.

PhD in progress: Yiqin Gao, "Replication Algorithms for Real-time Tasks with Precedence Constraints", started in October 2018, ENS Lyon, advisors: Yves Robert and Frédéric Vivien

PhD in progress: Changjiang Gou, Task scheduling on distributed platforms under memory and energy constraints, started in Oct. 2016, supervised by Anne Benoit & Loris Marchal.

PhD in progress: Li Han, "Algorithms for detecting and correcting silent and non-functional errors in scientific workflows", started in September 2016, funding: China Scholarship Council, advisors: Yves Robert and Frédéric Vivien

PhD in progress: Aurélie Kong Win Chang, "Techniques de résilience pour l'ordonnancement de workflows sur plates-formes décentralisées (cloud computing) avec contraintes de sécurité", started in October 2016, funding: ENS Lyon, advisors: Yves Robert, Yves Caniou and Eddy Caron.

PhD in progress: Valentin Le Fèvre, "Scheduling and resilience at scale", started in October 2017, funding: ENS Lyon, advisors: Anne Benoit and Yves Robert.

PhD: Gilles Moreau, On the solution phase of direct methods for sparse linear systems with multiple sparse right-hand sides, ENS Lyon, December 10, 2018, supervised by Jean-Yves L'Excellent and Patrick Amestoy.

PhD in progress: Ioannis Panagiotas, "High performance algorithms for big data graph and hypergraph problems", started in October 2017, funding: Inria, advisors: Frédéric Vivien and Bora Uçar.

PhD in progress: Filip Pawlowski, "High performance tensor computations", started in October 2017, funding: CIFRE, advisors: Yves Robert, Bora Uçar and Albert-Jan Yzelman (Huawei).

PhD: Loïc Pottier, Co-scheduling for large-scale applications: memory and resilience, ENS Lyon, September 18, 2018, supervised by Anne Benoit & Yves Robert.

PhD: Issam Raïs, Discover, model and combine energy leverages for large scale energy efficient infrastructures, ENS Lyon, September 28, 2018, supervised by Laurent Lefèvre & Anne Benoit & Anne-Cécile Orgerie.

PhD: Bertrand Simon, Scheduling task graphs on modern computing platforms, ENS Lyon, July 4, 2018, supervised by Loris Marchal & Frédéric Vivien.

10.2.3. Juries

Yves Robert was a Reviewer for the HDR of Alfredo Buttari (Toulouse) and Head of the Committee for the HDR of Abdou Guermouche and Pierre Ramet (Bordeaux). At ENS Lyon, he was a Committee member for the HDR of Loris Marchal, and for the PhD of Loic Pottier.

10.3. Popularization

10.3.1. Interventions

- Frédéric Vivien took part in the committee which listened to the presentations of high-school students in the scope of a “MATh.en.JEANS” action (December 2018).
- Yves Robert gave the honorary speech for the Honoris Causa Diploma of ENS Lyon awarded to Marc Snir on November 9, 2018.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] L. MARCHAL.*Memory and data aware scheduling*, École Normale Supérieure de Lyon, March 2018, Habilitation à diriger des recherches, https://hal.inria.fr/tel-01934712

[2] G. MOREAU.*On the Solution Phase of Direct Solvers for Sparse Linear Systems with Multiple Sparse Right-Hand Sides*, ENS Lyon ; Université de Lyon, December 2018, https://hal.archives-ouvertes.fr/tel-01959367

[3] L. POTTIER.*Co-scheduling for large-scale applications : memory and resilience*, Université de Lyon, September 2018, https://tel.archives-ouvertes.fr/tel-01892395

[4] B. SIMON.*Scheduling task graphs on modern computing platforms*, Université de Lyon, July 2018, https://tel.archives-ouvertes.fr/tel-01843558

Articles in International Peer-Reviewed Journal

[5] P. R. AMESTOY, A. BUTTARI, J.-Y. L'EXCELLENT, T. MARY.*Performance and Scalability of the Block Low-Rank Multifrontal Factorization on Multicore Architectures*, in "ACM Transactions on Mathematical Software", 2018, https://hal.inria.fr/hal-01955766

[6] P. AMESTOY, J.-Y. L'EXCELLENT, G. MOREAU.*On exploiting sparsity of multiple right-hand sides in sparse direct solvers*, in "SIAM Journal on Scientific Computing", 2018, p. 1-19, https://hal.inria.fr/hal-01955659

[7] G. AUPY, A. BENOIT, S. DAI, L. POTTIER, P. RAGHAVAN, Y. ROBERT, M. SHANTHARAM.*Co-scheduling Amdahl applications on cache-partitioned systems*, in "International Journal of High Performance Computing Applications", 2018, vol. 32, n° 1, p. 123-138, https://hal.inria.fr/hal-01968422

[8] A. BENOIT, L. LEFÈVRE, A.-C. ORGERIE, I. RAÏS.*Reducing the energy consumption of large scale computing systems through combined shutdown policies with multiple constraints*, in "International Journal of High Performance Computing Applications", January 2018, vol. 32, n° 1, p. 176-188 [*DOI :* 10.1177/1094342017714530], https://hal.inria.fr/hal-01557025

[9] H. CASANOVA, J. HERRMANN, Y. ROBERT.*Computing the expected makespan of task graphs in the presence of silent errors*, in "Parallel Computing", July 2018, vol. 75, p. 41-60, https://hal.inria.fr/hal-01968433

[10] F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Further notes on Birkhoff-von Neumann decomposition of doubly stochastic matrices*, in "Linear Algebra and Applications", 2018, vol. 554, p. 68–78 [*DOI :* 10.1016/J.LAA.2018.05.017], https://hal.inria.fr/hal-01586245

[11] L. HAN, L.-C. CANON, H. CASANOVA, Y. ROBERT, F. VIVIEN.*Checkpointing Workflows for Fail-Stop Errors*, in "IEEE Transactions on Computers", February 2018, vol. 67, n^o^ 8, 16 [*DOI :* 10.1109/TC.2018.2801300], https://hal.inria.fr/hal-01701611

[12] O. KAYA, Y. ROBERT.*Computing Dense Tensor Decompositions with Optimal Dimension Trees*, in "Algorithmica", 2018, https://hal.inria.fr/hal-01974471

[13] O. KAYA, B. UÇAR.*Parallel Candecomp/Parafac Decomposition of Sparse Tensors Using Dimension Trees*, in "SIAM Journal on Scientific Computing", 2018, vol. 40, n^o^ 1, p. C99 - C130 [*DOI :* 10.1137/16M1102744], https://hal.inria.fr/hal-01397464

[14] E. KAYAASLAN, C. AYKANAT, B. UÇAR.*1.5D Parallel Sparse Matrix-Vector Multiply*, in "SIAM Journal on Scientific Computing", January 2018, vol. 40, n^o^ 1, p. C25 - C46 [*DOI :* 10.1137/16M1105591], https://hal.inria.fr/hal-01897555

[15] E. KAYAASLAN, T. LAMBERT, L. MARCHAL, B. UÇAR.*Scheduling series-parallel task graphs to minimize peak memory*, in "Theoretical Computer Science", January 2018, vol. 707, p. 1-23 [*DOI :* 10.1016/J.TCS.2017.09.037], https://hal.inria.fr/hal-01891937

[16] L. MARCHAL, B. SIMON, O. SINNEN, F. VIVIEN.*Malleable task-graph scheduling with a practical speed-up model*, in "IEEE Transactions on Parallel and Distributed Systems", June 2018, vol. 29, n^o^ 6, p. 1357-1370 [*DOI :* 10.1109/TPDS.2018.2793886], https://hal.inria.fr/hal-01687189

International Conferences with Proceedings

[17] G. AUPY, A. BENOIT, B. GOGLIN, L. POTTIER, Y. ROBERT.*Co-scheduling HPC workloads on cache-partitioned CMP platforms*, in "IEEE Cluster 2018", Belfast, United Kingdom, Proceedings the 20th IEEE Cluster Conference, September 2018, p. 335-345, https://hal.inria.fr/hal-01874154

[18] G. AUPY, A. GAINARU, V. HONORÉ, P. RAGHAVAN, Y. ROBERT, H. SUN.*Reservation Strategies for Stochastic Jobs*, in "IPDPS 2019 - 33rd IEEE International Parallel and Distributed Processing Symposium", Rio de Janeiro, Brazil, May 2019, https://hal.inria.fr/hal-01968419

[19] O. BEAUMONT, T. LAMBERT, L. MARCHAL, B. THOMAS.*Data-Locality Aware Dynamic Schedulers for Independent Tasks with Replicated Inputs*, in "IPDPSW 2018 IEEE International Parallel and Distributed Processing Symposium Workshops", Vancouver, Canada, IEEE, May 2018, p. 1-8 [*DOI :* 10.1109/IPDPSW.2018.00187], https://hal.inria.fr/hal-01878977

[20] A. BENOIT, A. CAVELAN, F. CIORBA, V. LE FÈVRE, Y. ROBERT.*Combining Checkpointing and Replication for Reliable Execution of Linear Workflows*, in "APDCM'18 workshop, in conjunction with IPDPS'18", Vancouver, Canada, May 2018, https://hal.inria.fr/hal-01963655

[21] A. BENOIT, S. PERARNAU, L. POTTIER, Y. ROBERT.*A performance model to execute workflows on high-bandwidth-memory architectures*, in "ICPP 2018 - 47th International Conference on Parallel Processing",

Eugene, OR, United States, ACM, August 2018, p. 1-10 [*DOI :* 10.1145/3225058.3225110], https://hal.inria.fr/hal-01798726

[22] Y. CANIOU, E. CARON, A. KONG WIN CHANG, Y. ROBERT.*Budget-aware scheduling algorithms for scientific workflows with stochastic task weights on heterogeneous IaaS Cloud platforms*, in "IPDPSW 2018 - IEEE International Parallel and Distributed Processing Symposium Workshops", Vancouver, Canada, IEEE, May 2018, p. 15-26 [*DOI :* 10.1109/IPDPSW.2018.00014], https://hal.inria.fr/hal-01808831

[23] L.-C. CANON, A. KONG WIN CHANG, Y. ROBERT, F. VIVIEN.*Scheduling independent stochastic tasks under deadline and budget constraints*, in "SBAC-PAD 2018 - 30th International Symposium on Computer Architecture and High Performance Computing", Lyon, France, September 2018, p. 1-8, https://hal.inria.fr/hal-01868727

[24] L.-C. CANON, L. MARCHAL, B. SIMON, F. VIVIEN.*Online Scheduling of Task Graphs on Hybrid Platforms*, in "Euro-Par 2018 - 24th International European Conference On Parallel And Distributed Computing", Turin, Italy, August 2018, p. 1-14, https://hal.inria.fr/hal-01828301

[25] F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Approximation algorithms for maximum matchings in undirected graphs*, in "CSC 2018 - SIAM Workshop on Combinatorial Scientific Computing", Bergen, Norway, Proceedings of the Seventh SIAM Workshop on Combinatorial Scientific Computing, SIAM, June 2018, p. 56-65 [*DOI :* 10.1137/1.9781611975215.6], https://hal.archives-ouvertes.fr/hal-01740403

[26] C. GOU, A. BENOIT, M. CHEN, L. MARCHAL, T. WEI.*Reliability-aware energy optimization for throughput-constrained applications on MPSoC*, in "ICPADS - 24th International Conference on Parallel and Distributed Systems", Sentosa, Singapore, IEEE, December 2018, p. 1-10, https://hal.inria.fr/hal-01929927

[27] C. GOU, A. BENOIT, L. MARCHAL.*Memory-aware tree partitioning on homogeneous platforms*, in "PDP 2018 - 26th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing", Cambridge, United Kingdom, March 2018, p. 321-324 [*DOI :* 10.1109/PDP2018.2018.00056], https://hal.inria.fr/hal-01892022

[28] L. HAN, V. LE FÈVRE, L.-C. CANON, Y. ROBERT, F. VIVIEN.*A Generic Approach to Scheduling and Checkpointing Workflows*, in "ICPP 2018 - 47th International Conference on Parallel Processing", Eugene, OR, United States, ACM, August 2018, p. 1-10 [*DOI :* 10.1145/3225058.3225145], https://hal.inria.fr/hal-01798627

[29] *Best Paper*
T. HÉRAULT, Y. ROBERT, A. BOUTEILLER, D. ARNOLD, K. B. FERREIRA, G. BOSILCA, J. DONGARRA.*Optimal Cooperative Checkpointing for Shared High-Performance Computing Platforms*, in "APDCM", Vancouver, Canada, 2018, https://hal.inria.fr/hal-01968441.

[30] V. LE FÈVRE, G. BOSILCA, A. BOUTEILLER, T. HÉRAULT, A. HORI, Y. ROBERT, J. DONGARRA.*Do Moldable Applications Perform Better on Failure-Prone HPC Platforms?*, in "Resilience - EuroPar workshop", Torino, Italy, 2018, p. 787-799, https://hal.inria.fr/hal-01968448

[31] L. MARCHAL, H. NAGY, B. SIMON, F. VIVIEN.*Parallel scheduling of DAGs under memory constraints*, in "IPDPS 2018 - 32nd IEEE International Parallel and Distributed Processing Symposium", Vancouver, Canada, IEEE, May 2018, p. 1-10 [*DOI :* 10.1109/IPDPS.2018.00030], https://hal.inria.fr/hal-01828312

[32] I. RAÏS, M. BOUTIGNY, L. LEFÈVRE, A.-C. ORGERIE, A. BENOIT.*Building the Table of Energy and Power Leverages for Energy Efficient Large Scale Systems*, in "HPCS: International Conference on High Performance Computing & Simulation", Orléans, France, July 2018, p. 284-291 [*DOI :* 10.1109/HPCS.2018.00056], https://hal.archives-ouvertes.fr/hal-01845970

[33] I. RAÏS, L. LEFÈVRE, A.-C. ORGERIE, A. BENOIT.*Exploiting the Table of Energy and Power Leverages*, in "ICA3PP 2018 - 18th International Conference on Algorithms and Architectures for Parallel Processing", Guangzhou, China, November 2018, p. 1-10, https://hal.archives-ouvertes.fr/hal-01927829

[34] A. YASAR, B. UÇAR, U. V. CATALYUREK.*SINA: A Scalable Iterative Network Aligner*, in "2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)", Barcelona, Spain, August 2018, https://hal.inria.fr/hal-01918744

Scientific Books (or Scientific Book chapters)

[35] G. AUPY, Y. ROBERT.*Scheduling for Fault-Tolerance: An Introduction*, in "Topic in parallel and distributed computing: Enhancing the Undergraduate Curriculum: Performance, Concurrency, and Programming on Modern Platforms", Springer International Publishing, September 2018, p. 143-170, https://hal.inria.fr/hal-01968454

Research Reports

[36] P. AMESTOY, A. BUTTARI, J.-Y. L'EXCELLENT, T. MARY.*Bridging the gap between flat and hierarchical low-rank matrix formats: the multilevel BLR format*, University of Manchester, April 2018, https://hal.archives-ouvertes.fr/hal-01774642

[37] P. R. AMESTOY, S. DE LA KETHULLE DE RYHOVE, J.-Y. L'EXCELLENT, G. MOREAU, D. V. SHANTSEV.*Efficient use of sparsity by direct solvers applied to 3D controlled-source EM problems*, Inria Grenoble Rhône-Alpes ; LIP - ENS Lyon, November 2018, n^{o} RR-9220, 26, https://hal.inria.fr/hal-01912713

[38] G. AUPY, A. BENOIT, B. GOGLIN, L. POTTIER, Y. ROBERT.*Co-scheduling HPC workloads on cache-partitioned CMP platforms*, Inria, February 2018, n^{o} RR-9154, https://hal.inria.fr/hal-01719728

[39] G. AUPY, A. GAINARU, V. HONORÉ, P. RAGHAVAN, Y. ROBERT, H. SUN.*Reservation Strategies for Stochastic Jobs (Extended Version)*, Inria & Labri, Univ. Bordeaux ; Department of EECS, Vanderbilt University, Nashville, TN, USA ; Laboratoire LIP, ENS Lyon & University of Tennessee Knoxville, Lyon, France, October 2018, n^{o} RR-9211, p. 1-38, https://hal.inria.fr/hal-01903592

[40] A. BENOIT, A. CAVELAN, F. CIORBA, V. LE FÈVRE, Y. ROBERT.*Combining Checkpointing and Replication for Reliable Execution of Linear Workflows with Fail-Stop and Silent Errors*, ROMA (Inria Rhône-Alpes / LIP Laboratoire de l'Informatique du Parallélisme) ; LIP - Laboratoire de l'Informatique du Parallélisme, December 2018, p. 1-32, https://hal.inria.fr/hal-01955859

[41] A. BENOIT, A. CAVELAN, F. CIORBA, V. LE FÈVRE, Y. ROBERT.*Combining Checkpointing and Replication for Reliable Execution of Linear Workflows*, Inria - Research Centre Grenoble – Rhône-Alpes, February 2018, n^{o} RR-9152, p. 1-36, https://hal.inria.fr/hal-01714978

[42] A. BENOIT, S. PERARNAU, L. POTTIER, Y. ROBERT.*A performance model to execute workflows on high-bandwidth memory architectures*, ENS Lyon ; Inria Grenoble Rhône-Alpes ; University of Tennessee

Knoxville ; Georgia Institute of Technology ; Argonne National Laboratory, April 2018, n^{o} RR-9165, p. 1-28, https://hal.inria.fr/hal-01767888

[43] G. BOSILCA, A. BOUTEILLER, T. HÉRAULT, V. LE FÈVRE, Y. ROBERT, J. J. DONGARRA.*Distributed Termination Detection for HPC Task-Based Environments*, Inria - Research Centre Grenoble – Rhône-Alpes, June 2018, n^{o} RR-9181, p. 1-28, https://hal.inria.fr/hal-01811823

[44] L.-C. CANON, A. KONG WIN CHANG, Y. ROBERT, F. VIVIEN.*Scheduling independent stochastic tasks deadline and budget constraints*, Inria - Research Centre Grenoble – Rhône-Alpes, June 2018, n^{o} RR-9178, p. 1-34, https://hal.inria.fr/hal-01811885

[45] L.-C. CANON, L. MARCHAL, B. SIMON, F. VIVIEN.*Online Scheduling of Sequential Task Graphs on Hybrid Platforms*, LIP - ENS Lyon, February 2018, n^{o} RR-9150, https://hal.inria.fr/hal-01720064

[46] F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Effective heuristics for matchings in hypergraphs*, Inria Grenoble Rhône-Alpes, November 2018, n^{o} RR-9224, p. 1-18, https://hal.archives-ouvertes.fr/hal-01924180

[47] F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Scaling matrices and counting the perfect matchings in graphs*, Inria Grenoble Rhône-Alpes, March 2018, n^{o} RR-9161, p. 1-22, https://hal.inria.fr/hal-01743802

[48] C. GOU, A. BENOIT, M. CHEN, L. MARCHAL, T. WEI.*Reliability-aware energy optimization for throughput-constrained applications on MPSoC*, Laboratoire LIP, École Normale Supérieure de Lyon & CNRS & Inria, France ; Shanghai Key Lab. of Trustworthy Computing, East China Normal University, China ; Georgia Institute of Technology, USA, April 2018, n^{o} RR-9168, p. 1-35, https://hal.inria.fr/hal-01766763

[49] L. HAN, V. LE FÈVRE, L.-C. CANON, Y. ROBERT, F. VIVIEN.*A Generic Approach to Scheduling and Checkpointing Workflows*, Inria, April 2018, n^{o} RR-9167, p. 1-29, https://hal.inria.fr/hal-01766352

[50] J. HERRMANN, M. YUSUF ÖZKAYA, B. UÇAR, K. KAYA, U. V. CATALYUREK.*Acyclic partitioning of large directed acyclic graphs*, Inria - Research Centre Grenoble – Rhône-Alpes, March 2018, n^{o} RR-9163, https://hal.inria.fr/hal-01744603

[51] V. LE FÈVRE, G. BOSILCA, A. BOUTEILLER, T. HÉRAULT, A. HORI, Y. ROBERT, J. J. DONGARRA.*Do moldable applications perform better on failure-prone HPC platforms?*, Inria Grenoble Rhône-Alpes, May 2018, n^{o} RR-9174, p. 1-24, https://hal.inria.fr/hal-01799498

[52] L. MARCHAL, B. SIMON, F. VIVIEN.*Limiting the memory footprint when dynamically scheduling DAGs on shared-memory platforms*, Inria Grenoble Rhône-Alpes, December 2018, n^{o} RR-9231, p. 1-41, https://hal.inria.fr/hal-01948462

[53] M. Y. ÖZKAYA, A. BENOIT, B. UÇAR, J. HERRMANN, U. V. CATALYUREK.*A scalable clustering-based task scheduler for homogeneous processors using DAG partitioning*, Inria Grenoble Rhône-Alpes, June 2018, n^{o} RR-9185, p. 1-30, https://hal.inria.fr/hal-01817501

References in notes

[54] *Blue Waters Newsletter*, dec 2012

[55] *Blue Waters Resources*, 2013, https://bluewaters.ncsa.illinois.edu/data

[56] *The BOINC project*, 2013, http://boinc.berkeley.edu/

[57] *Final report of the Department of Energy Fault Management Workshop*, December 2012, https://science.energy.gov/~/media/ascr/pdf/program-documents/docs/FaultManagement-wrkshpRpt-v4-final.pdf

[58] *System Resilience at Extreme Scale: white paper*, 2008, DARPA, https://pdfs.semanticscholar.org/9fcb/154d6afce23cd9951fd7c116b86255d91b5c.pdf

[59] *Top500 List - November*, 2011, http://www.top500.org/list/2011/11/

[60] *Top500 List - November*, 2012, http://www.top500.org/list/2012/11/

[61] M. AMARIS, G. LUCARELLI, C. MOMMESSIN, D. TRYSTRAM.*Generic Algorithms for Scheduling Applications on Hybrid Multi-core Machines*, in "Euro-Par 2017: Parallel Processing", 2017, p. 220–231

[62] I. ASSAYAD, A. GIRAULT, H. KALLA.*Tradeoff exploration between reliability power consumption and execution time*, in "Proceedings of SAFECOMP, the Conf. on Computer Safety, Reliability and Security", Washington, DC, USA, 2011

[63] H. AYDIN, Q. YANG.*Energy-aware partitioning for multiprocessor real-time systems*, in "IPDPS'03, the IEEE Int. Parallel and Distributed Processing Symposium", 2003, p. 113–121

[64] N. BANSAL, T. KIMBREL, K. PRUHS.*Speed Scaling to Manage Energy and Temperature*, in "Journal of the ACM", 2007, vol. 54, n^{o} 1, p. 1 – 39, http://doi.acm.org/10.1145/1206035.1206038

[65] A. BENOIT, L. MARCHAL, J.-F. PINEAU, Y. ROBERT, F. VIVIEN.*Scheduling concurrent bag-of-tasks applications on heterogeneous platforms*, in "IEEE Transactions on Computers", 2010, vol. 59, n^{o} 2, p. 202-217

[66] S. BLACKFORD, J. CHOI, A. CLEARY, E. D'AZEVEDO, J. DEMMEL, I. DHILLON, J. DONGARRA, S. HAMMARLING, G. HENRY, A. PETITET, K. STANLEY, D. WALKER, R. C. WHALEY.*ScaLAPACK Users' Guide*, SIAM, 1997

[67] S. BLACKFORD, J. DONGARRA.*Installation Guide for LAPACK*, LAPACK Working Note, June 1999, n^{o} 41, originally released March 1992

[68] R. A. BRUALDI.*Notes on the Birkhoff algorithm for doubly stochastic matrices*, in "Canadian Mathematical Bulletin", 1982, vol. 25, n^{o} 2, p. 191–199

[69] A. BUTTARI, J. LANGOU, J. KURZAK, J. DONGARRA.*Parallel tiled QR factorization for multicore architectures*, in "Concurrency: Practice and Experience", 2008, vol. 20, n^{o} 13, p. 1573-1590

[70] J.-J. CHEN, T.-W. KUO.*Multiprocessor energy-efficient scheduling for real-time tasks*, in "ICPP'05, the Int. Conference on Parallel Processing", 2005, p. 13–20

[71] S. DONFACK, L. GRIGORI, W. GROPP, L. V. KALE.*Hybrid Static/dynamic Scheduling for Already Optimized Dense Matrix Factorization*, in "Parallel Distributed Processing Symposium (IPDPS), 2012 IEEE 26th International", 2012, p. 496-507, http://dx.doi.org/10.1109/IPDPS.2012.53

[72] J. DONGARRA, J.-F. PINEAU, Y. ROBERT, Z. SHI, F. VIVIEN.*Revisiting Matrix Product on Master-Worker Platforms*, in "International Journal of Foundations of Computer Science", 2008, vol. 19, n^{o} 6, p. 1317-1336

[73] J. DONGARRA, J.-F. PINEAU, Y. ROBERT, F. VIVIEN.*Matrix Product on Heterogeneous Master-Worker Platforms*, in "13th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming", Salt Lake City, Utah, February 2008, p. 53–62

[74] I. S. DUFF, J. K. REID.*The multifrontal solution of indefinite sparse symmetric linear systems*, in ""ACM Transactions on Mathematical Software"", 1983, vol. 9, p. 302-325

[75] I. S. DUFF, J. K. REID.*The multifrontal solution of unsymmetric sets of linear systems*, in "SIAM Journal on Scientific and Statistical Computing", 1984, vol. 5, p. 633-641

[76] L. GRIGORI, J. W. DEMMEL, H. XIANG.*Communication avoiding Gaussian elimination*, in "Proceedings of the 2008 ACM/IEEE conference on Supercomputing", Piscataway, NJ, USA, SC '08, IEEE Press, 2008, 29:1, http://dl.acm.org/citation.cfm?id=1413370.1413400

[77] B. HADRI, H. LTAIEF, E. AGULLO, J. DONGARRA.*Tile QR Factorization with Parallel Panel Processing for Multicore Architectures*, in "IPDPS'10, the 24st IEEE Int. Parallel and Distributed Processing Symposium", 2010

[78] J. W. H. LIU.*An application of generalized tree pebbling to sparse matrix factorization*, in "SIAM Journal on Algebraic and Discrete Methods", 1987, vol. 8, n^{o} 3, p. 375–395

[79] J. W. H. LIU.*The multifrontal method for sparse matrix solution: Theory and Practice*, in "SIAM Review", 1992, vol. 34, p. 82–109

[80] R. MELHEM, D. MOSSÉ, E. ELNOZAHY.*The Interplay of Power Management and Fault Recovery in Real-Time Systems*, in "IEEE Transactions on Computers", 2004, vol. 53, n^{o} 2, p. 217-231

[81] A. J. OLINER, R. K. SAHOO, J. E. MOREIRA, M. GUPTA, A. SIVASUBRAMANIAM.*Fault-aware job scheduling for bluegene/l systems*, in "IPDPS'04, the IEEE Int. Parallel and Distributed Processing Symposium", 2004, p. 64–73

[82] G. QUINTANA-ORTÍ, E. QUINTANA-ORTÍ, R. A. VAN DE GEIJN, F. G. V. ZEE, E. CHAN.*Programming Matrix Algorithms-by-Blocks for Thread-Level Parallelism*, in "ACM Transactions on Mathematical Software", 2009, vol. 36, n^{o} 3

[83] Y. ROBERT, F. VIVIEN.*Algorithmic Issues in Grid Computing*, in "Algorithms and Theory of Computation Handbook", Chapman and Hall/CRC Press, 2009

[84] G. ZHENG, X. NI, L. V. KALE.*A scalable double in-memory checkpoint and restart scheme towards exascale*, in "Dependable Systems and Networks Workshops (DSN-W)", 2012, http://dx.doi.org/10.1109/DSNW.2012.6264677

[85] D. ZHU, R. MELHEM, D. MOSSÉ.*The effects of energy management on reliability in real-time embedded systems*, in "Proc. of IEEE/ACM Int. Conf. on Computer-Aided Design (ICCAD)", 2004, p. 35–40

Project-Team SOCRATE

Software and Cognitive radio for telecommunications

IN COLLABORATION WITH: Centre of Innovation in Telecommunications and Integration of services

IN PARTNERSHIP WITH:
Institut national des sciences appliquées de Lyon

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Networks and Telecommunications

Table of contents

Project-Team SOCRATE

Creation of the Team: 2012 January 01, updated into Project-Team: 2013 July 01

Keywords:

Computer Science and Digital Science:

A1.1.2. - Hardware accelerators (GPGPU, FPGA, etc.)
A1.1.10. - Reconfigurable architectures
A1.1.12. - Non-conventional architectures
A1.2.5. - Internet of things
A1.2.6. - Sensor networks
A1.5.2. - Communicating systems
A2.3.1. - Embedded systems
A2.6.1. - Operating systems
A5.9. - Signal processing
A8.6. - Information theory

Other Research Topics and Application Domains:

B6.2. - Network technologies
B6.2.2. - Radio technology
B6.4. - Internet of things
B6.6. - Embedded systems

1. Team, Visitors, External Collaborators

Research Scientist

Samir Perlaza [Inria, Researcher]

Faculty Members

Malcolm Egan [INSA Lyon, Associate Professor]
Florent Dupont de Dinechin [INSA Lyon, Professor, HDR]
Jean-Marie Gorce [INSA Lyon, Professor, HDR]
Claire Goursaud [INSA Lyon, Associate Professor, HDR]
Florin Hutu [INSA Lyon, Associate Professor]
Kevin Marquet [INSA Lyon, Associate Professor]
Tanguy Risset [Team leader, INSA Lyon, Professor, HDR]
Guillaume Salagnac [INSA Lyon, Associate Professor]
Leonardo Sampaio Cardoso [INSA Lyon, Associate Professor]
Guillaume Villemaud [INSA Lyon, Associate Professor, HDR]

External Collaborator

Nikolai Lebedev [Ecole supérieure de chimie physique électronique de Lyon]

Technical Staff

Olivier Nicolas [INSA Lyon]
Pascal Girard [INSA Lyon]
Othmane Oubejja [Inria]

PhD Students

Dadja Anade Akpo [INSA Lyon]

Gautier Berthou [Inria]
Andrea Bocco [CEA]
Tristan Delizy [INSA Lyon]
Diane Duchemin [Inria]
Luc Forget [INSA Lyon, from Oct 2018]
Yanni Zhou [Insa Lyon, from sept 2018]
Tarik Lassouaoui [Insa Lyon, from sept 2018]
Hassan Kallam [INSA Lyon]
Nizar Khalfet [Inria]
David Kibloff [Inria]
Yuqi Mo [INSA Lyon]
Cyrille Morin [Inria]
Yohann Uguen [INSA Lyon]
Régis Rousseau [INSA Lyon]

Administrative Assistant

Florence Maillard [Inria, from Feb 2018]

2. Overall Objectives

2.1. Introduction

The success of radio networking relies on a small set of rules: *i)* protocols are completely defined beforehand, *ii)* resource allocation policies are mainly designed in a static manner and *iii)* access network architectures are planned and controlled. Such a model obviously lacks adaptability and also suffers from a suboptimal behavior and performance.

Because of the growing demand for radio resources, several heterogeneous standards and technologies have been introduced by the standard organizations or industry by different workgroups within the IEEE (802 family), ETSI (GSM), 3GPP (3G, 4G) or the Internet Society (IETF standards) leading to the almost saturated usage of several frequency bands (see Fig. 1).

These two facts, obsolescence of current radio networking rules on one hand, and saturation of the radio frequency band on the other hand, are the main premises for the advent of a new era of radio networking that will be characterized by self-adaptive mechanisms. These mechanisms will rely on software radio Åtechnologies, distributed algorithms, end-to-end dynamic routing protocols and therefore require a cross-layer vision of "cognitive wireless networking": *Getting to the meet of Cognition and Cooperation, beyond the inherent communication aspects: cognition is more than cognitive radio and cooperation is not just relaying. Cognition and cooperation have truly the potential to break new ground for mobile communication systems and to offer new business models.* [56]

From a social perspective, pervasive communications and ambient networking are becoming part of more and more facets of our daily life. Probably the most popular usage is mobile Internet access, which is made possible by numerous access technologies, e.g. cellular mobile networks, WiFi, Bluetooth, etc. The access technology itself is becoming *transparent for the end user*, who does not care about how to access the network but is only interested in the services available and in the quality of this service.

Beyond simple Internet access, many other applications and services are built on the basis of pervasive connectivity, for which the communication is just a mean, and not a finality. Thus, the wireless link is expected to even be *invisible to the end user* and constitutes the first element of the Future Internet of Things [54], to develop a complete twin virtual world fully connected to the real one.

The way radio technologies have been developed until now is far from offering a real wireless convergence [43]. The current development of the wireless industry is surely slowed down by the lack of radio resources and the lack of systems flexibility.

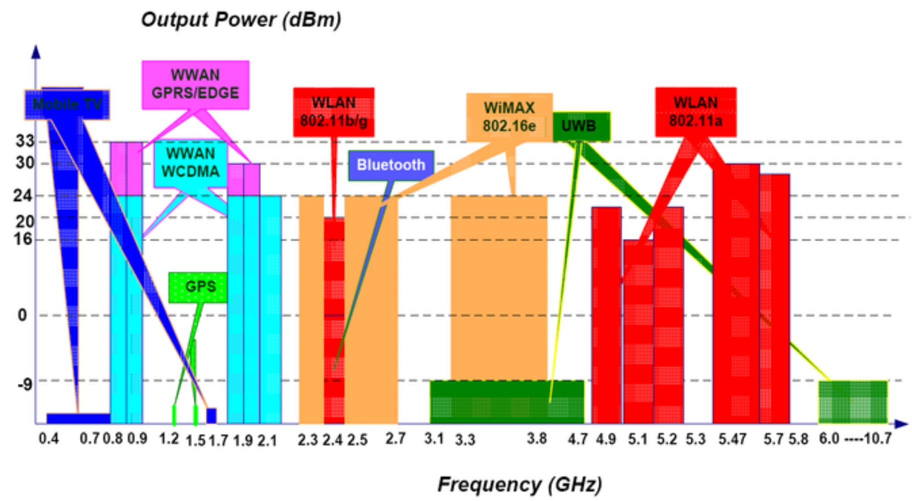


Figure 1. The most recent standards for wireless communications are developed in the UHF and VHF bands. These bands are mostly saturated (source: WPAN/WLAN/WWAN Multi-Radio Coexistence, IEEE 802 Plenary, Atlanta, USA, Nov.2007)

One can get rid of this technological bottleneck by solving three complementary problems: *terminal flexibility, agile radio resource management* and *autonomous networking*. These three objectives are subsumed by the concept of *Software Radio*, a term coined by J. Mitola in his seminal work during the early 90's [51], [52]. While implementing everything in software nodes is still an utopia, many architectures now hitting the market include some degree of programmability; this is called Software-Defined Radio. The word "defined" has been added to distinguish from the ideal software radio. A software *defined* radio is a software radio which is defined for a given frequency range and a maximal bandwidth.

In parallel, the development of new standards is threatened by the radio spectrum scarcity. As illustrated in Fig. 1, the increasing number of standards already causes partial saturation of the UHF band, and will probably lead to its full saturation in the long run. However, this saturation is only "virtual" because all equipments are fortunately not emitting all the time [43]. A good illustration is the so-called "white spaces", i.e. frequency bands that are liberated by analog television disappearing and can be re-used for other purposes, different rules are set up in different countries. In this example, a solution for increasing the real capacity of the band originates from *self-adaptive behavior*. In this case, flexible terminals will have to implement agile algorithms to share the radio spectrum and to avoid interference. In this context, cooperative approaches are even more promising than simple resource sharing algorithms.

With Software-Defined Radio technology, terminal flexibility is at hand, many questions arise that are related to the software layer of a software radio machine: how will this kind of platform be programmed? How can we write programs that are portable from one terminal to another? Autonomous networking will only be reached after a deep understanding of network information theory. Thus, given that there will be many ways for transmitting data from one point to another, what is the most efficient way in terms of throughput? power consumption? etc. Last but not least, agile Radio Resource sharing is addressed by studying MIMO and multi-standard radio front-end. This new technology is offering a wide range of research problems. These three topics: software programming of a software radio machine, distributed algorithms for radio resource management and multi-standard radio front-end constitute the research directions of Socrate.

2.2. Technological State of the Art

A Software-Defined Radio (SDR) system is a radio communication system in which computations that in the past were typically implemented in hardware (e.g. mixers, filters, amplifiers, modulators/demodulators, detectors, etc.) are instead implemented as software programs [51], [44].

2.2.1. SDR Technology

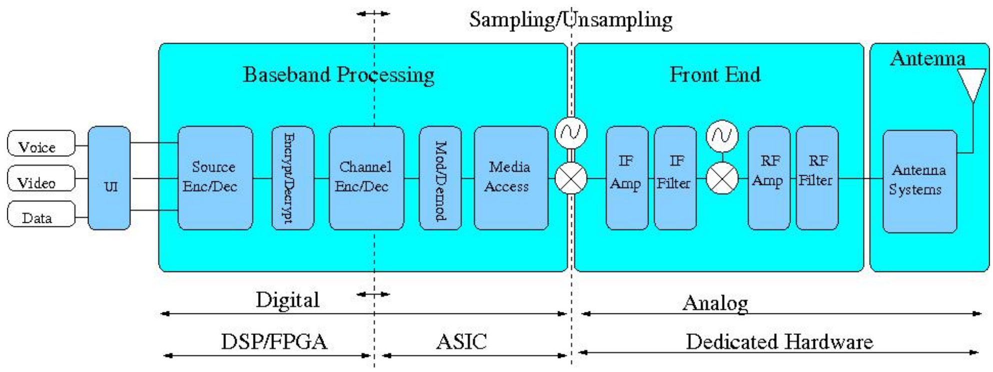


Figure 2. Radio Block Diagram, highlighting separation between digital and analog parts, as well as programmable, configurable and fixed hardware parts.

The different components of a radio system are illustrated in Fig. 2. Of course, all of the digital components may not be programmable, but the bigger the programmable part (DSP/FPGA part on Fig. 2), the more *software* the radio. Dedicated IPs. In this context, IP stand for *Intellectual Properties, this term is widely used to designated dedicated special-purpose circuit blocks implemented in various technologies: Asic,* FPGA, DSP, *etc.* are needed, for these IP it is more suitable to use the term *configurable* than programmable. In a typical SDR, the analog part is limited to a frequency translation down to an intermediate band which is sampled and all the signal processing is done digitally.

2.2.2. SDR Forum Classification

To encourage a common meaning for the term "SDR" the SDR Forum (recently renamed *Wireless Innovation Forum* (http://www.wirelessinnovation.org) proposes to distinguish five tiers:

- *Tier 0 – Hardware Radio:* The radio parameters cannot be changed, radio is implemented only with hardware components.
- *Tier 1 – Software Controlled Radio:* A radio where only the control functions are implemented in software, baseband processing is still performed in hardware, the radio is able to switch between different hardware.
- *Tier 2 – Software-Defined Radio:* The most popularly understood definition of SDR: the radio includes software control of modulation, bandwidth, frequency range and frequency bands. Conversion to digital domain still occurs after frequency conversion. It is currently implemented using a wide range of technologies: Asics, FPGAs, DSPs, etc.
- *Tier 3 – Ideal Software Radio:* Digital conversion occurs directly at the antenna, programmability extends to the whole system.
- *Tier 4 – Ultimate Software Radio:* Same reconfigurability capabilities as in Tier 3, but with a switching between two configurations in less than one millisecond.

The main restriction to build an ideal software radio is sampling rate: sampling at a high rate is not an easy task. Following the Shannon-Nyquist theorem, sampling the RF signal at a rate greater than twice the frequency of the signal is sufficient to reconstruct the signal. Sampling can be done at lower rate (decimation), but errors can be introduced (aliasing) that can be corrected by filtering (dirty radio concept). Building an SDR terminal implies a trade-of between sampling frequency and terminal complexity. For instance, sampling at 4.9 GHz would require a 12-bit resolution ADC with at least 10GHz sample rate which is today not available with reasonable power consumption (several hundreds Watt).

2.2.3. Cognitive Radio

SDR technology enables *over the air programming* (Otap) which consists in describing methods for distributing new software updates through the radio interface. However, as SDR architectures are heterogeneous, a standard distribution method has not emerged yet.

Cognitive Radio is a wireless communication system that can sense the air, and decide to configure itself in a given mode, following a local or distributed decision algorithm. Although Tier 3 SDR would be an ideal platform for cognitive radio implementation, cognitive radios do not have to be SDR.

Cognitive Radio is currently a very hot research topic as show the dozens of sessions in research conferences dedicated to it. In 2009, the American National Science Foundation (NSF) held a workshop on "Future Directions in Cognitive Radio Network Research" [53]. The purpose of the workshop was to explore how the transition from cognitive radios to cognitive radio *networks* can be made. The resulting report indicated the following:

- Emerging cognitive radio technology has been identified as a high impact disruptive technology innovation, that could provide solutions to the *radio traffic jam* problem and provide a path to scaling wireless systems for the next 25 years.
- Significant new research is required to address the many technical challenges of cognitive radio networking. These include dynamic spectrum allocation methods, spectrum sensing, cooperative communications, incentive mechanisms, cognitive network architecture and protocol design, cognitive network security, cognitive system adaptation algorithms and emergent system behavior.

The report also mentioned the lack of cognitive radio testbeds and urged *"The development of a set of cognitive networking test-beds that can be used to evaluate cognitive networks at various stages of their development"*, which, in some sense strengthens the creation of the Socrate team and its implication in the FIT project [46].

2.3. Scientific Challenges

Having a clear idea of relevant research areas in SDR is not easy because many parameters are not related to economical cost. For instance, military research has made its own development of SDR for its particular needs: US military SDR follows the SCA communication architecture [49] but this is usually not considered as a realistic choice for a commercial SDR handset. The targeted frequency band has a huge impact as sampling at high rates is very expensive, and trade-offs between flexibility, complexity, cost and power consumption have a big influence on the relative importance of the hot research topics.

Here are the relevant research domains where efforts are needed to help the deployment of SDR:

- *Antennas and RF Front-Ends:* This is a key issue for reducing interference, increasing capacity and reusing frequency. Hot topics such as wake-up radio or multi protocol parallel radio receivers are directly impacted by research on Antennas. Socrate has research work going on in this area.
- *Analog to Digital Converters:* Designing low-power high frequency ADC is still a hot topic rather studied by micro-electronics laboratories (Lip6 for instance in France).
- *Architecture of SDR systems:* The ideal technology for embedded SDR still has to be defined. Hardware prototypes are built using FPGAs, Asics and DSPs, but the real challenge is to handle a Hardware/Software design which includes radio and antennas parts.

- *Middleware for* SDR *systems:* How to manage, reconfigure, update and debug SDR systems is still an open question which is currently studied for each SDR platform prototype. Having a common programming interface for SDR systems in one research direction of Socrate.
- *Distributed signal processing:* Cognitive, smart or adaptive radios will need complex decision algorithms which, most of the time will need to be solved in a distributed manner. Socrate has clearly a strong research effort in that direction. Distributed information theory is also a hot research topic that Socrate wishes to study.

3. Research Program

3.1. Research Axes

In order to keep young researchers in an environment close to their background, we have structured the team along the three research axes related to the three main scientific domains spanned by Socrate. However, we insist that a *major objective* of the Socrate team is to *motivate the collaborative research between these axes*. The first one is entitled “Flexible Radio Front-End” and will study new radio front-end research challenges brought up by the arrival of MIMO technologies, and reconfigurable front-ends. The second one, entitled “Multi-user communication”, will study how to couple the self-adaptive and distributed signal processing algorithms to cope with the multi-scale dynamics found in cognitive radio systems. The last research axis, entitled “Software Radio Programming Models” is dedicated to embedded software issues related to programming the physical protocols layer on these software radio machines. Figure 3 illustrates the three regions of a transceiver corresponding to the three Socrate axes.

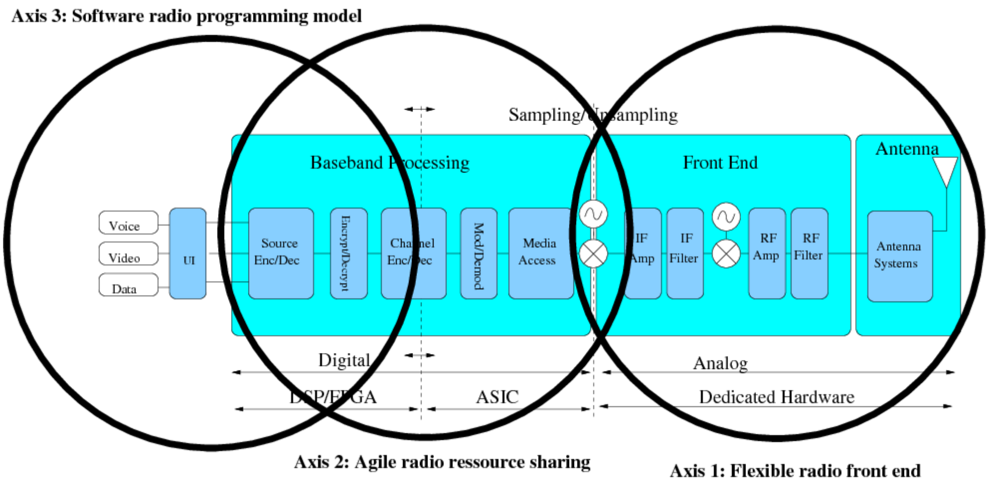


Figure 3. Center of interest for each of the three Socrate research axes with respect to a generic software radio terminal.

3.2. Flexible Radio Front-End

Participants: Guillaume Villemaud, Florin Hutu.

This axis mainly deals with the radio front-end of software radio terminals (right of Fig 3). In order to ensure a high flexibility in a global wireless network, each node is expected to offer as many degrees of freedom as possible. For instance, the choice of the most appropriate communication resource (frequency channel, spreading code, time slot,...), the interface standard or the type of antenna are possible degrees of freedom. The *multi-** paradigm denotes a highly flexible terminal composed of several antennas providing MIMO features to enhance the radio link quality, which is able to deal with several radio standards to offer interoperability and efficient relaying, and can provide multi-channel capability to optimize spectral reuse. On the other hand, increasing degrees of freedom can also increase the global energy consumption, therefore for energy-limited terminals a different approach has to be defined.

In this research axis, we expect to demonstrate optimization of flexible radio front-end by fine grain simulations, and also by the design of home made prototypes. Of course, studying all the components deeply would not be possible given the size of the team, we are currently not working in new technologies for DAC/ADC and power amplifiers which are currently studied by hardware oriented teams. The purpose of this axis is to build system level simulation taking into account the state of the art of each key component.

3.3. Multi-User Communications

Participants: Jean Marie Gorce, Claire Goursaud, Samir Perlaza, Leonardo Sampaio Cardoso, Malcolm Egan.

While the first and the third research axes deal with the optimization of the cognitive radio nodes themselves from system and programming point of view, an important complementary objective is to consider the radio nodes in their environments. Indeed, cognitive radio does not target the simple optimization of point to point transmissions, but the optimization of simultaneous concurrent transmissions. The tremendous development of new wireless applications and standards currently observed calls for a better management of the radio spectrum with opportunistic radio access, cooperative transmissions and interference management. This challenge has been identified as one of the most important issue for 5G to guarantee a better exploitation of the spectrum. In addition, mobile internet is going to support a new revolution that is the *tactile internet*, with real time interactions between the virtual and the real worlds, requiring new communication objectives to be met such as low latency end to end communications, distributed learning techniques, in-the-network computation, and many more. The future network will be heterogeneous in terms of technologies, type of data flows and QoS requirements. To address this revolution two work directions have naturally formed within the axis. The first direction concerns the theoretical study of fundamental limits in wireless networks. Introduced by Claude Shannon in the 50s and heavily developed up to today, Information Theory has provided a theoretical foundation to study the performance of wireless communications, not from a practical design view point, but using the statistical properties of wireless channels to establish the fundamental trade-offs in wireless communications. Beyond the classical *energy efficiency - spectral efficiency* tradeoff, information theory and its many derivations, i.e., network information theory, may also help to address additional questions such as determining the optimal rates under decentralized policies, asymptotic behavior when the density of nodes increases, latency controled communication with finite block-length theory, etc. In these cases, information theory is often associated to other theoretical tools such as game theory, stochastic geometry, control theory, graph theory and many others.

Our first research direction consists in evaluating specific mulit-user scenarios from a network information theory perspective, inspired by practical scenarios from various applicative frameworks (e.g. 5G, Wifi, sensor networks, IoT, etc.), and to establish fundamental limits for these scenarios. The second research direction is related to algorithmic and protocol design (PHY/MAC), applied to practical scenarios. Exploiting signal processing, linear algebra inspired models and distributed algorithms, we develop and evaluate various distributed algorithms allowing to improve many QoS metrics such as communication rates, reliability, stability, energy efficiency or computational complexity.

It is clear that both research directions are symbiotic with respect to each other, with the former providing theoretical bounds that serves as a reference to the performance of the algorithms created in the later. In the other way around, the later offers target scenarios for the former, through identifying fundamental problems that are interesting to be studied from the fundamental side. Our contributions of the year in these two directions are summarized further in the document.

3.4. Software Radio Programming Model

Participants: Tanguy Risset, Kevin Marquet, Guillaume Salagnac, Florent de Dinechin.

Finally the third research axis is concerned with software aspect of the software radio terminal (left of Fig 3). We have currently two actions in this axis, the first one concerns the programming issues in software defined radio devices, the second one focusses on low power devices: how can they be adapted to integrate some reconfigurability.

The expected contributions of Socrate in this research axis are :

- The design and implementation of a "middleware for SDR", probably based on a Virtual Machine.
- Prototype implementations of novel software radio systems, using chips from Leti and/or Lyrtech software radio boards.
- Development of a *smart node*: a low-power Software-Defined Radio node adapted to WSN applications.
- Methodology clues and programming tools to program all these prototypes.

3.5. Evolution of the Socrate team

In 2018 the Socrate team has decided to split in two teams: the Maracas team will consist of the activities of Socrate Axis 2 and be directed by Jean-Marie Gorce, and the Socrate team which will consist in the Axis 1 and 3 of the current version of Socrate. This change is explicit since september 2018 as the Maracas team is created.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Various

Two new workshop organized by the team in relation with CorteXlab:

- First French GNU Radio days: https://gnuradio-fr-18.sciencesconf.org/
- ISP-IoT : First Winter School on Information Theory and Signal Processing for Internet of Things : https://isp-iot.sciencesconf.org

4.1.2. Awards

The PhD of Victor Quintero (former PhD in Socrate) received the best PhD award in the area of digital society in Nov 2018.

Samir M. Perlaza is Visiting Research Collaborator (Honorific Position) Oct. 2018; Term 2018 - 2019. Department of Electrical Engineering, Princeton University

Samir M. Perlaza has been awarded a "Make our Planet Great Again" Fellowship, Sep. 2018 by Embassy of France in the United States of America and Thomas Jefferson Foundation in New York, NY.

The article *Karatsuba with Rectangular Multipliers for FPGAs* , presented by Florent de Dinechin, obtained the Best Paper Award of the Arith 2018 conference in Amherst, MA.
BEST PAPERS AWARDS :
[17]
M. KUMM, O. GUSTAFSSON, F. DE DINECHIN, J. KAPPAUF, P. ZIPF.*Karatsuba with Rectangular Multipliers for FPGAs*, in "ARITH 2018 - 25th IEEE International Symposium on Computer Arithmetic", Amherst, United States, IEEE, June 2018, p. 13-20, Best paper award [*DOI :* 10.1109/ARITH.2018.8464809], https://hal.inria.fr/hal-01773447

5. New Software and Platforms

5.1. fftweb

KEYWORDS: Experimentation - Data visualization - SDR (Software Defined Radio)
FUNCTIONAL DESCRIPTION: fftweb is a real-time spectral (FFT) visualization of one or several signal, embedded in a web page. The FFT is computed in a GNURadio block, then sent to a gateway server, which serves the web page, associated javascripts, and signal websockets. The end user only has to use the GNURadio block and and the web page, and doesn't need to bother about the internal details of the system. fftweb has been developped specially for the CorteXlab testbed but with minor adaptations, it can be used in other contexts, and also can be used to draw more generic real-time graphs, not only FFTs. Technologies: GNURadio, python, python-gevent, Javascript, D3JS

- Contact: Matthieu Imbert

5.2. FloPoCo

Floating-Point Cores, but not only
KEYWORD: Synthesizable VHDL generator
FUNCTIONAL DESCRIPTION: The purpose of the open-source FloPoCo project is to explore the many ways in which the flexibility of the FPGA target can be exploited in the arithmetic realm.
NEWS OF THE YEAR: FloPoCo has been enhanced in 2019 with a last-bit accurate IIR filter generator and a generator of parallel FFT cores, among others.

- Participants: Florent Dupont De Dinechin and Luc Forget
- Partners: CNRS - ENS Lyon - UCBL Lyon 1 - UPVD
- Contact: Florent Dupont De Dinechin
- URL: http://flopoco.gforge.inria.fr/

5.3. minus

KEYWORDS: Experimentation - SDR (Software Defined Radio)
FUNCTIONAL DESCRIPTION: Minus is an experiment control system able to control, the whole lifecycle of a radio experiment in CorteXlab or any other testbed inspired by it. Minus controls and automates the whole experiment process starting from node power cycling, experiment deployment, experiment start and stop, and results collection and transfer. Minus is also capable of managing multiple queues of experiments which are executed simultaneously in the testbed.

- Contact: Matthieu Imbert

5.4. WiPlan

FUNCTIONAL DESCRIPTION: Wiplan is a software including an Indoor propagation engine and a wireless LAN optimization suite, which has been registered by INSA-Lyon. The heart of this software is the propagation simulation core relying on an original method, MR-FDPF (multi-resolution frequency domain ParFlow), proposed by JM Gorce in 2001 and further extended. The discrete ParFlow equations are translated in the Fourier domain providing a large linear system, solved in two steps taking advantage of a multi- resolution approach. The first step computes a cell-based tree structure referred to as the pyramid. In the second phase, a radiating source is simulated, taking advantage of the pre-processed pyramidal structure. Using of a full-space discrete simulator instead of classical ray-tracing techniques is a challenge due to the inherent high computation requests. However, we have shown that the use of a multi-resolution approach allows the main computational load to be restricted to a pre-processing phase. Extensive works have been done to make predictions more realistic.

- Contact: Tanguy Risset

5.5. Sytare

KEYWORDS: Embedded systems - Operating system - Non volatile memory
FUNCTIONAL DESCRIPTION: Sytare is an embedded operating system targeting tiny platforms with intermittent power. In order to make power failures transparent for the application, the system detects imminent failures and saves a checkpoint of program state to non-volatile memory. Hardware peripherals are also made persistent without requiring developer attention.

- Authors: Tristan Delizy, Gautier Berthou, Guillaume Salagnac, Kevin Marquet and Tanguy Risset
- Contact: Guillaume Salagnac
- Publication: Peripheral State Persistence For Transiently Powered Systems
- URL: https://hal.inria.fr/hal-01460699

6. New Results

6.1. Multi-User Communications

Activities in axis 2 primarily focus on communicating multi-user systems. They represent the core of the research activity that will be pursued in Maracas team.

The first pillar of our research concerns the evaluation of fundamental limits of wireless systems (e.g. capacity) often express as a fundamental tradeoff : energy efficiency - spectral efficiency tradeoff, rate versus reliability, information versus energy transfert,... Our work relies mostly on information theory, signal processing, estimation theory and game theory.

The second pillar concerns the evaluation of real systems and their performance is confronted to the above mentioned fundamental limits. These activities rely on strong collaborations with industry (Nokia, Orange, SigFox, Sequans, SPIE-ICS,...) We also manage the FIT/CorteXlab testbed offering a remote access to a worldwide unique platform.

Beyond these two pillars, we also explore new research areas where our background is relevant. These prospective activities are performed with external collaborations and prepare the future activity of Maracas team. This year we explored molecular communications (supported by an Inria exploratory project), smart grids in collaboration with Sheffield, VLC in association with Agora team or Privacy preservation in collaboration with Privatics team.

6.1.1. Fundamental limits in communications

6.1.1.1. Variations on point to point capacity and related tools

In [31] discrete approximations of the capacity are introduced where the input distribution is constrained to be discrete in addition to any other constraints on the input. For point-to-point memoryless additive noise channels, rates of convergence to the capacity of the original channel are established for a wide range of channels for which the capacity is finite. These results are obtained by viewing discrete approximations as a capacity sensitivity problem, where capacity losses are studied when there are perturbations in any of the parameters describing the channel. In particular, it is shown that the discrete approximation converges arbitrarily close to the channel capacity at rate $O(\Delta)$, where Δ is the discretization level of the approximation. Examples of channels where this rate of convergence holds are also given, including additive Cauchy and inverse Gaussian noise channels.

In [30] the properties of finite frames are explored. Finite frames are sequences of vectors in finite dimensional Hilbert spaces that play a key role in signal processing and coding theory. We studied the class of tight unit-norm frames for $\mathbb{C}^d$ that also form regular schemes, called tight regular schemes (TRS). Many common frames that arise in applications such as equiangular tight frames and mutually unbiased bases fall in this class. We investigate characteristic properties of TRSs and prove that for many constructions, they are intimately connected to weighted 1-designs—arising from quadrature rules for integrals over spheres in $\mathbb{C}^d$ with weights dependent on the Voronoi regions of each frame element.

6.1.1.2. Interference channel with feedback

The interference channel is a well-known model used to represent simultaneous transmissions in a wireless environment. In the framework of Victor Quintero's PhD, we explored the performance of this model with noisy feedbacks.

In [35], an achievable η-Nash equilibrium (η-NE) region for the two-user Gaussian interference channel with noisy channel-output feedback is presented for all $\eta \geq 1$. This result is obtained in the scenario in which each transmitter-receiver pair chooses its own transmit-receive configuration in order to maximize its own individual information transmission rate. At an η-NE, any unilateral deviation by either of the pairs does not increase the corresponding individual rate by more than η bits per channel use.

In [6], the capacity region of the linear deterministic interference channel with noisy channel-output feedback (LD-IC-NF) is fully characterized. The proof of achievability is based on random coding arguments and rate splitting; block-Markov superposition coding; and backward decoding. The proof of converse reuses some of the existing outer bounds and includes new ones obtained using genie-aided models. Following the insight gained from the analysis of the LD-IC-NF, an achievability region and a converse region for the two-user Gaussian interference channel with noisy channel-output feedback (GIC-NF) are presented. Finally, the achievability region and the converse region are proven to approximate the capacity region of the G-IC-NF to within 4.4 bits.

6.1.1.3. Wiretap channel

The Wiretap channel allows to address the secrecy constraint in an information theory framework. In [13], an analysis of an input distribution that achieves the secrecy capacity of a general degraded additive noise wiretap channel is presented. In particular, using convex optimization methods, an input distribution that achieves the secrecy capacity is characterized by conditions expressed in terms of integral equations. The new conditions are used to study the structure of the optimal input distribution for three different additive noise cases: vector Gaussian; scalar Cauchy; and scalar exponential.

6.1.1.4. Simultaneous Information and Energy Transmission

Simultaneous information and energy transmission (SIET) is an active research problem and aims at providing energy and information simultaneously from transmitters to receivers. We explore the optimal trade-offs in different settings.

In [34], a non-asymptotic analysis of the fundamental limits of simultaneous energy and information transmission (SEIT) is presented. The notion of information-capacity region, i.e., the largest set of simultaneously achievable information and energy rates, is revisited in a context in which transmissions occur within a finite number of channel uses and strictly positive error decoding probability and energy shortage probability are tolerated. The focus is on the case of one transmitter, one information receiver and one energy harvester communicating through binary symmetric memoryless channels. In this case, the information-capacity region is approximated and the trade-off between information rate and energy rate is thoroughly studied.

In [5], the fundamental limits of simultaneous information and energy transmission (SIET) in the two-user Gaussian interference channel (G-IC) with and without perfect channel-output feedback are approximated by two regions in each case, i.e., an achievable region and a converse region. When the energy transmission rate is normalized by the maximum energy rate the approximation is within a constat gap. In the proof of achievability, the key idea is the use of power-splitting between two signal components: an information-carrying component and a no-information component. The construction of the former is based on random coding arguments, whereas the latter consists in a deterministic sequence known by all transmitters and receivers. The proof of the converse is obtained via cut-set bounds, genie-aided channel models, Fano's inequality and some concentration inequalities considering that channel inputs might have a positive mean. Finally, the energy transmission enhancement due to feedback is quantified and it is shown that feedback can at most double the energy transmission rate at high signal to noise ratios.

6.1.1.5. Modeling Interference in Large-Scale Uplink SCMA

Massive connectivity is a fundamental challenge for IoT, as discussed in the next section from a practical perspective. From a theoretical perspective, we propose to relax the assumption of Gaussian interference.

Fast varying active transmitter sets with very short length transmissions arise in communications for the Internet of Things. As a consequence, the interference is dynamic, leading to non-Gaussian statistics. At the same time, the very high density of devices is motivating non-orthogonal multiple access (NOMA) techniques, such as sparse code multiple access (SCMA). In [2], we study the statistics of the dynamic interference from devices using SCMA. In particular, we show that the interference is $\alpha-$stable with non-trivial dependence structure for large scale networks modeled via Poisson point processes. Moreover, the interference on each frequency band is shown to be sub-Gaussian $\alpha-$stable in the special case of disjoint SCMA codebooks. We investigate the impact of the $\alpha-$stable interference on achievable rates and on the optimal density of devices. Our analysis suggests that ultra dense networks are desirable even with $\alpha-$stable interference.

This contribution is a good introduction of the next section where the performance of IoT access techniques are evaluated.

6.1.1.6. General Massive Machine Type Communications Uplink

Non Orthogonal Multiple Access (NOMA) is expected to play an important role for IoT networks, allowing to reduce signaling overheads and to maximize the capacity of dense networks with multiple packets simultaneous transmission. In the uplink, NOMA can improve significantly the performance of an ALOHA random access if the receiver implements a multiuser detection algorithm. In [11], we compared the performance of a code domain NOMA with a classical ALOHA protocol, through simulations. The code domain NOMA uses random Gaussian codes at the transmitters and exploits compressive sensing at the receiver to maximize users detection and to minimize symbol error rates.

As the number of machine type communications increases at an exponential rate, new solutions have to be found in order to deal with the uplink traffic. At the same time, new types of Base Stations (BS) that use a high number of antennas are being designed, and their beamforming capabilities can help to separate signals that have different angles of arrivals. In [15], we consider a network where a BS serves a high number of nodes that lacks a receive chain, and we analyze the evolution of the outage probability as a function of the number of antennas at the BS. We then study the effect of an angle offset between the main beam and the desired node's direction in order to provide realistic results in a beam-switching scenario.

6.1.1.7. Multiple Base Stations Diversity for UNB Systems

In the framework of the long-term collaboration with Sigfox, the PhD of Yuqi Mo defended mast December, explored the performance of Ultra Narrow Band (UNB) with a focus on sophisticated signal processing techniques such as multi-BS processing or successive interference cancellation (SIC). UNB (Ultra Narrow Band) is one of the technologies dedicated to low-power wide-area communication for IoT, currently exploited by SigFox

In [33], [18], the specificity of UNB is the Aloha-type channel access scheme, asynchronous in both time and frequency domain. This randomness can cause partial spectral interference. In this paper, we take advantage of the spatial diversity of multiple base stations to improve the UNB performance, by using selection combining. In the presence of pathloss and spectral randomness of UNB, the channels are considered correlated. A theoretical analysis of outage probability is demonstrated by considering this correlation, for the case of 2 base stations. This methodology of probability computing can be extended to K BSs. The diversity of multiple receivers is proved to be beneficial in enhancing the performance of UNB networks. This gain is shown to be related to the density of the base stations, as well as the distance between each of them. In [8], we propose to apply signal combining and interference cancellation technologies across multiple base stations in UNB networks, in order to take advantage of their spatial diversity. We evaluate and compare the performance enhancement of each technology, compared to single BS case. These technologies exploiting multi-BS diversity are proved to be significantly beneficial in improving UNB networks' scalability. We can gain until 28 times better performance with one iteration global SIC. We highlight that these results provide us a choice among the technologies according to the improvement needs and the implementation complexity.

6.1.2. Contributions in other application fields

6.1.2.1. Molecular communications

Molecular communications is emerging as a technique to support coordination in nanonetworking, particularly in biochemical systems. In complex biochemical systems such as in the human body, it is not always possible to view the molecular communication link in isolation as chemicals in the system may react with chemicals used for the purpose of communication. There are two consequences: either the performance of the molecular communication link is reduced; or the molecular link disrupts the function of the biochemical system. As such, it is important to establish conditions when the molecular communication link can coexist with a biochemical system. In [4], we develop a framework to establish coexistence conditions based on the theory of chemical reaction networks. We then specialize our framework in two settings: an enzyme-aided molecular communication system; and a low-rate molecular communication system near a general biochemical system. In each case, we prove sufficient conditions to ensure coexistence. In [29], we develop a general framework for the coexistence problem by drawing an analogy to the cognitive radio problem in wireless communication systems. For the particularly promising underlay strategy, we propose a formalization and outline key consequences.

Another key challenge in nanonetworking is to develop a means of coordinating a large number of nanoscale devices. Devices in molecular communication systems—once information molecules are released— are typically viewed as passive, not reacting chemically with the information molecules. While this is an accurate model in diffusion-limited links, it is not the only scenario. In particular, the dynamics of molecular communication systems are more generally governed by reaction-diffusion, where the reaction dynamics can also dominate. This leads to the notion of reaction-limited molecular communication systems, where the concentration profiles of information molecules and other chemical species depends largely on reaction kinetics. In this regime, the system can be approximated by a chemical reaction network. In [14], we exploit this observation to design new protocols for both point-to-point links with feedback and networks for event detection. In particular, using connections between consensus and advection theory and reaction networks lead to simple characterizations of equilibrium concentrations, which yield simple—but accurate— design rules even for networks with a large number of devices.

6.1.2.2. Smart Grids

Smart grids is another application field where information theory and signal processing can be useful. During 2018, we addressed security issues. In [41], random attacks that jointly minimize the amount of information acquired by the operator about the state of the grid and the probability of attack detection are presented. The attacks minimize the information acquired by the operator by minimizing the mutual information between the observations and the state variables describing the grid. Simultaneously, the attacker aims to minimize the probability of attack detection by minimizing the Kullback-Leibler (KL) divergence between the distribution when the attack is present and the distribution under normal operation. The resulting cost function is the weighted sum of the mutual information and the KL divergence mentioned above. The trade-off between the probability of attack detection and the reduction of mutual information is governed by the weighting parameter on the KL divergence term in the cost function. The probability of attack detection is evaluated as a function of the weighting parameter. A sufficient condition on the weighting parameter is given for achieving an arbitrarily small probability of attack detection. The attack performance is numerically assessed on the IEEE 30-Bus and 118-Bus test systems.

6.1.2.3. Privacy and tracking

In a joint work with Privatics team, we presented in [40] the analysis of an Ultrasound-based tracking application. By analyzing several mobile applications along with the network communication and sample of the original audio signal, we were able to reverse engineer the ultrasonic communications and some other elements of the system. Based on those finding we show how arbitrary ultrasonic signal can be generated and how to perform jamming. Finally we analyze a real world deployment and discuss privacy implications.

6.1.2.4. VLC

In a joint work with Agora, we present in [12] our efforts to design a communication system between an ordinary RGB light emitting diode and a smart-phone. This work in progress presents our preliminary findings obtained investigating this poorly known and unusual channel. We give engineering insights on driving an

RGB light emitting diode for camera communication and discuss remaining challenges. Finally, we propose possible solutions to cope with these issues that are blockers for a user ready implementation.

6.1.2.5. Intelligent Transport

On-demand transport has been disrupted by Uber and other providers, which are challenging the traditional approach adopted by taxi services. Instead of using fixed passenger pricing and driver payments, there is now the possibility of adaptation to changes in demand and supply. Properly designed, this new approach can lead to desirable tradeoffs between passenger prices, individual driver profits and provider revenue. However, pricing and allocations - known as mechanisms - are challenging problems falling in the intersection of economics and computer science. In [3], we develop a general framework to classify mechanisms in on-demand transport. Moreover, we show that data is key to optimizing each mechanism and analyze a dataset provided by a real-world on-demand transport provider. This analysis provides valuable new insights into efficient pricing and allocation in on-demand transport.

6.2. Flexible Radio Front-End

Activities in this axis could globally be divided in two main topics: low-power wireless sensors (with applications in wearable devices, guided propagation for ventilation systems, and tag-to-tag RFID), and optimization of waveforms (for wake-up radio receivers and wireless power transfer).

6.2.1. Low-Power WSN

Wearable sensors for health monitoring can enable the early detection of various symptoms, and hence rapid remedial actions may be undertaken. In particular, the monitoring of cardiac events by using such wearable sensors can provide real-time and more relevant diagnosis of cardiac arrhythmia than classical solutions. However, such devices usually use batteries, which require regular recharging to ensure long-term measurements. In the framework of a local collaborative project, we therefore designed and evaluated a connected sensor for the ambulatory monitoring of cardiac events, which can be used as an autonomous device without the need of a battery. Even when using off-the-shelf, low-cost integrated circuits, by optimizing both the hardware and software embedded in the device, we were able to reduce the energy consumption of the entire system to below 0.4 mW while measuring and storing the ECG on a non-volatile memory. Moreover, in this project, a power-management circuit able to store energy collected from the radio communication interface is proposed, able to make the connected sensor fully autonomous. Initial results show that this sensor could be suitable for a truly continuous and long-term monitoring of cardiac events [32].

In collaboration with Atlantic, we have done here a preliminary study [37], [23] of wireless transmissions using the ventilation metallic ducts as waveguides. Starting from the waveguide theory, we deeply studied in simulation the actual attenuation encountered by radiowaves in such a specific medium. This kind of wireless link appears to be really efficient, and therefore highly promising to implement Internet of Things (IoT) in old buildings to make them smarter. This study also expresses a very simple empirical model in order to ease dimensioning a wireless network in such conditions and a specific antenna design enabling both good performance and high robustness to the influence of the environment.

The Spie ICS- INSA Lyon chair on IoT has granted us for a PhD thesis on Scatter Radio and RFID tag-to-tag communications. Some seminal results have shown that it is actually possible to create a communication between two RFID tags, just using ambient radiowaves or a dedicated distant radio source, without the need of generating a signal from the tag itself.

6.2.2. Optimization of waveforms for wake-up radio and energy harvesting

First Filter Bank Multi Carrier (FBMC) signals are employed in order to improve the performance of a quasi-passive wake-up radio receiver (WuRx) for which the addressing is performed by the means of a frequency fingerprint. The feasibility of such kind of WuRx was already demonstrated by using orthogonal frequency-division multiplexing (OFDM) signals to form the identifiers. Together with the main advantage of this approach (i.e. no base band processing needed and consequently a reduced energy consumption), one of the drawbacks is their low sensitivity. Through a set of circuit-system co-simulations, it is shown that by their characteristics, especially high Peak to Average Power Ratio (PAPR) and high out of band attenuation, FBMC signals manage to boost the sensitivity and moreover to enhance the robustness of this kind of WuRx. Moreover, we introduced robust wake-up IDs for quasi-passive wake-up receivers in an Internet of Things context[16]. These IDs can address single devices and are based on the Hadamard codes. Further a novel wake-up threshold is implemented to make the device more sensitive and robust against false wake-ups (FWUs). The wake-up procedure is simulated with a tap delay line (TDL) model for a line of sight (LOS) channel and a non line of sight (NLOS) channel. In both scenarios sufficient wake-up distances are reached with low false wake-up probabilitys (FWUPs). Additionally, the system is tested against the influence of an external bandwidth use. Finally, a recommendation for the global system is given.

In [21], we are proposing a way to maximize the DC power collected in the case of a wireless power transfer (WPT) scenario. Three main aspects are taken into account: the RF (radio frequency) source, the propagation channel and the rectifier as the main part of the energy collecting circuit. This problem is formulated as a convex optimization one. Then, as a first step towards solving this problem, a rectifier circuit was simulated by using Keysight's ADS software and, by using a classical model identification strategy i.e. Vector Fitting algorithm, the state-space model of the passive parts of this rectifier were extracted. In order to verify the extracted model, S11 input reflection coefficients and DC output voltages of the original circuit and the state-space model are compared.

6.2.3. UWB for localization

Ultra Wide Band (UWB) is a wireless communication technology that is characterized, in its *impulse radio* scheme [55], by very short duration waveforms called *pulses* (in the order of few nanoseconds), using a wide band and low power spectral density. Among the many advantages offered by this technology is the fact that the arrival time of a pulse can be determined quite precisely, giving the opportunity to measure the distance between two communicating devices by estimating the flight time of the signal.

Although this technology has been known for a long time, it is only recently that cheap UWB chips have been commercialized for civilian applications. As the UWB technology is sensitive to many parameters, the effective performance of localization systems based on UWB may vary a lot compared to what is announced in datasheets. Some accuracy studies have been performed [47], [48] but few of them focus on rapid movement of the transceivers.

Indeed, indoor ranging is in itself dependent on many parameter and very difficult to evaluate objectively, but when the transceivers are moving fast (say as if they were attached to dancer's wrists), more parameters are to be taken into account: transceiver calibration, random errors, presence of obstacle, antenna orientation etc.

In [20], we study experimentally the precision of UWB ranging for rapid movements in an indoor environment, based on the technology proposed by Decawave (DW1000 [45]) whose chips have already been integrated in many commercial devices. We show in particular how to improve the precision of the distance measured by averaging the ranging over successive samples.

6.3. Software Radio Programming Model

6.3.1. Non Uniform Memory Access Analyzer

Non Uniform Memory Access (NUMA) architectures are nowadays common for running High-Performance Computing (HPC) applications. In such architectures, several distinct physical memories are assembled to create a single shared memory. Nevertheless, because there are several physical memories, access times to these memories are not uniform depending on the location of the core performing the memory request and on the location of the target memory. Hence, threads and data placement are crucial to efficiently exploit such architectures. To help in taking decision about this placement, profiling tools are needed. In [36], we propose NUMA MeMory Ana-lyzer (NumaMMA), a new profiling tool for understanding the memory access patterns of HPC applications. NumaMMA combines efficient collection of memory traces using hardware mechanisms with original visualization means allowing to see how memory access patterns evolve over time. The information reported by NumaMMA allows to understand the nature of these access patterns inside each object allocated by the application. We show how NumaMMA can help understanding the memory patterns of several HPC applications in order to optimize them and get speedups up to 28% over the standard non optimized version.

6.3.2. Environments for transiently powered devices

An important research initiative is being followed in Socrate today: the study of the new NVRAM technology and its use in ultra-low power context. NVRAM stands for Non-Volatile Radom Access Memory. Non-Volatile memory has been extising for a while (Nand Flash for instance) but was not sufficiently fast to be used as main memory. Many emerging technologies are forseen for Non-Volatile RAM to replace current RAM [50].

Socrate has started a work on the applicability of NVRAM for *transiantly powered systems*, i.e. systems which may undergo power outage at any time. This study resulted in the Sytare software presented at the NVMW conference [25] and also to the starting of an Inria Project Lab [39]: ZEP.

The Sytare software introduces a checkpointing system that takes into account peripherals (ADC, leds, timer, radio communication, etc.) present on all embedded system. Checkpointing is the natural solution to power outage: regularly save the state of the system in NVRAM so as to restore it when power is on again. However, no work on checkpointing took into account the restoration of the states of peripherals, Sytare provides this possibility. A complete description of Sytare has been accepted to IEEE Transaction on Computers [1], special issue on NVRAM.

6.3.3. Dynamic memory allocation for heterogeneous memory systems

In a low power system-on-chip the memory hierarchy is traditionally composed of Static RAM (SRAM) and NOR flash. The main feature of SRAM is a fast access time, while Flash memory is dense, and also non-volatile i.e. it does not require power to retain data. Because of its low writing speed, Flash memory is mostly used in a read-only fashion (e.g. for code) and the amount of SRAM is kept to a minimum in order to lower leakage power.

Emerging memory technologies exhibit different trade-offs and more heterogeneity. Non-Volatile RAM technologies like MRAM (Magnetic RAM) or RRAM (Resistive RAM) open new perspectives on power-management since they can be switched on or off at very little cost. Their characteristics are very dependent on the technology used, but it is now widely known that they will provide a high integration density and fast read access time to persistent data. NVRAM is usually not as fast as SRAM and some technologies have a limited endurance hence are not suited to store frequently modified data. In addition, most NVRAM technologies have asymmetric access times, writes being slower than reads.

In the context of embedded systems, the hardware architecture is evolving towards a model where different memory banks, with different hardware characteristics, are directly exposed to software, as it has been the case for scratchpad memories (SPM). This raises questions including:

- What is the expected performance impact of adding fast memory to a system based on NVRAM? In particular: will the addition of a small amount of fast memory result in significant performance improvement?
- How should one adapt and optimize their software memory management to leverage these new technologies?

In [10], [28], we study these questions in the perpective of dynamic memory allocation. In this first study we show, with extensive profiling how much can be gained with a clever dynamic memory allocation in the context of heterogeneous memory. We limit the study to two different memories, RAM and NVRAM for instance. This gain can go up to 15% of performance, depending of course of the performances of the different memories used. These results will be helpfull to design a clever dynamic allocator for these new architectures and also will help in the design process of new architecture for low power systems that will include NVRAM for normaly-off systems for instance.

6.3.4. Arithmetic for signal processing

Linear Time Invariant (LTI) filters are often specified and simulated using high-precision software, before being implemented in low-precision fixed-point hardware. A problem is that the hardware does not behave exactly as the simulation due to quantization and rounding issues. The article [7] advocates the construction of LTI architectures that behave as if the computation was performed with infinite accuracy, then converted to the low-precision output format with an error smaller than its least significant bit. This simple specification guarantees the numerical quality of the hardware, even for critical LTI systems. Besides, it is possible to derive the optimal values of all the internal data formats that ensure that the specification is met. This requires a detailed error analysis that captures not only the quantization and rounding errors, but also their infinite accumulation in recursive filters. This generic methodology is detailed for the case of low-precision LTI filters in the Direct Form I implemented in FPGA logic. It is demonstrated by a fully automated and open-source architecture generator tool integrated in FloPoCo, and validated on a range of Infinite Impulse Response filters.

6.3.5. Karatsuba multipliers on modern FPGAs

The Karatsuba method is a well-known technique to reduce the complexity of large multiplications. However it is poorly suited to the rectangular 17x25-bit multipliers embedded in recent Xilinx FPGAs: The traditional Karatsuba approach must under-use them as square 18x18 ones. In [17], the Karatsuba method is extended to efficiently use such rectangular multipliers to build larger multipliers. Rectangular multipliers can be efficiently exploited if their input word sizes have a large greatest common divider. In the Xilinx FPGA case, this can be obtained by using the 17x25 embedded multipliers as 16x24. The obtained architectures are implemented with due detail to architectural features such as the pre-adders and post-adders available in Xilinx DSP blocks. They are synthesized and compared with traditional Karatsuba, but also with (non-Karatsuba) state-of-the-art tiling techniques that make use of the full rectangular multipliers. The proposed technique improves resource consumption and performance for multipliers of numbers larger than 64 bits.

6.3.6. PyGA: a Python to FPGA compiler prototype

In a collaboration with Intel, Yohann Uguen has worked on a compiler of Python to FPGA [22]. Based on the Numba Just-In-Time (JIT) compiler for Python and the Intel FPGA SDK for OpenCL, it allows any Python user to use a FPGA card as an accelerator for Python seamlessly, albeit with limited performance so far.

6.3.7. General computer arithmetic

A second edition of the Handbook for Floating-Point Arithmetic has been published [38].

With colleagues from Aric, we have worked on a critical review [42] of the Posit system, a proposed alternative to the prevalent floating-point format.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants with Industry

7.1.1. Research Contract with Atlantic 2016-2018

Socrate (Guillaume Villemaud, Florin Hutu, Guillaume Salagnac and Tanguy Risset) are collborating with Atlantic to prototype guided wireless communications in ventilation ducts with low energy consumption. The project will lead to a shift to wireless communications in HVAC ducts.

7.1.2. Research Contract with SigFox 2015-2018

Socrate explored the performance of UNB networks with an emphasis on robust signal processing techniques (PhD defended on Dec 2018).

7.1.3. Research Contract with Orange 2016-2018

Socrate explored in this partnership the theoretical limits of IoT access networks by combining information theory and stochastic geometry.

7.1.4. Research Contract with Nokia 2017-2021

Socrate contributes to two research actions in the Nokia Bell Labs - Inria common lab. The first ADR is on Network Information Theory devoted to to the modeling of IoT networks, and which relies on our academic work in the ANR Arburst. We collaborate with Agora, Infine and Eva teams.

The second ADR is on machine learning for wireless networks. Our contribution is on designing new PHY layer protocols with machine learning, with an experimental assessment of these techniques on FIT/CorteXlab.

7.1.5. Research Contract with Bosch 2018

In collaboration with Aric, Socrate worked with Bosch on the impementation of some elementary functions in an embedded context.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. Equipex FIT- Future Internet of Things

The FIT projet is a national equipex (*equipement d'excellence*), headed by the Lip6 laboratory. As a member of Inria, Socrate is in charge of the development of an Experimental Cognitive Radio platform that is used as test-bed for SDR terminals and cognitive radio experiments. This has been operational since 2014 and is maintained for a duration of 7 years. To give a quick view, the user will have a way to configure and program through Internet several SDR platforms (MIMO, SISO, and baseband processing nodes).

8.1.2. Insa-Spie IoT Chair

The Insa-Spie IoT Chair http://www.citi-lab.fr/chairs/iot-chair/ relies on the expertise of the CITI Lab. The skills developed within the different teams of the lab integrate the study, modelling, conception and evaluation of technologies for communicating objects and dedicated network architectures. It deals with network, telecom and software matters as well as societal issues such as privacy. The chair will also lean on the skills developed at INSA Lyon or in IMU LabEx.

8.1.3. Inria Project Lab: ZEP

The ZEP project addresses the issue of designing tiny computing objects with no battery by combining non-volatile memory (NVRAM), energy harvesting, micro-architecture innovations, compiler optimizations, and static analysis. The main application target is Internet of Things (IoT) where small communicating objects will be composed of this computing part associated to a low-power wake-up radio system. The ZEP project gathers four Inria teams that have a scientific background in architecture, compilation, operating system and low power together with the CEA Lialp and Lisan laboratories of CEA LETI & LIST. The major outcomes of the project will be a prototype harvesting board including NVRAM and the design of a new microprocessor associated with its optimizing compiler and operating system.

The scientific work (in progress) is organized around three fields :

- specific NVRAM-based architecture
- dedicated compiler pass that computes a worst-case energy consumption
- operating system managing NVRAM and energy, ensuring memory consistency across power outages

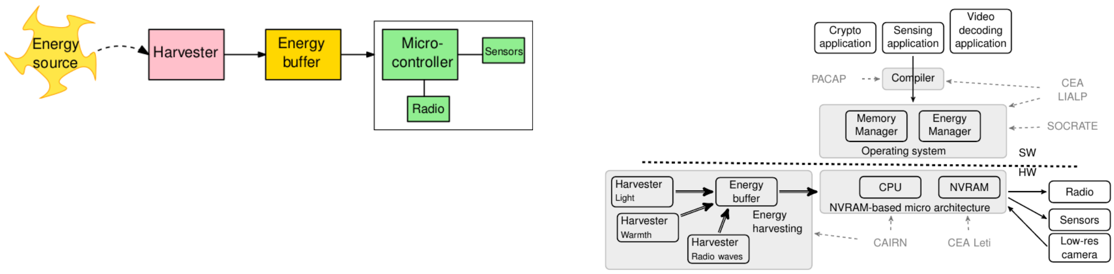


Figure 4. Example of system targeted by the ZEP project on the left, and on the right: the ZEP research program.

The project is illustrated by the figure 4, where PACAP, SOCRATE, CORSE, and CAIRN are the teams involved in the project.

Another important goal of the project is to structure the research and innovation that should occur within Inria to prepare the important technological shift brought by NVRAM technologies.

8.1.4. ANR - Imprenum

The objective of this project (INSA-Lyon, École Normale Supérieure de Lyon, CEA LETI) is to promote **accuracy as a first class concern** in all the levels of a computing system:

- at the hardware level, with better support for lower-than-standard and higher-than-standard precisions;
- at the level of run-time support software, in particular answering the memory management challenges entailed by adaptive precision;
- at the lower level of mathematical libraries (kernel level), for instance BLAS for linear algebra, enhancing well established libraries with precision and accuracy control;
- at the higher level of mathematical libraries (solver level, including algebraic linear solvers such as LAPACK, ad hoc steppers for Ordinary Differential Equation, eigenvalues kernels, triangularization problems for computational geometry, etc.) Here, accuracy and precision control of the lower levels should enable higher-level properties such as convergence and stability;
- at the compiler level, enhancing optimising compilers with novel optimisations related to precision and accuracy;
- at the language level, embedding accuracy specification and control in existing languages, and possibly defining domain-specific languages with accuracy-aware semantics for some classes of applications.

8.1.5. ADT CorteXlab

The Socrate project-team is in charge of the FIT/CorteXlab platform. This platform (ADT Inria 2015-2017) makes use of many complex technologies from signal processing to computer science through micro-electornics and FPGA. The objectiv of the CorteXlab ADT is to maintain a support to the user of the FPGA-based platform of CorteXlab and to provide tutorial and running experiment that will help them in builing experimentation using the PicoSDR machines.

8.1.6. ANR - Ephyl

The general objective of the project EPHYL - “Enhanced PHY for Cellular Low Power Communication IoT” (2016-2019, 183 keuros) is to investigate coming and future LPWA technologies with the aim to improve coverage, data rate and connectivity while keeping similar level of complexity and power consumption at the node for the access. New waveforms enablers will be investigated and trialled in order to increase the efficiency of future systems and to provide efficient and fair access to the radio resource. The proposed new waveforms should comply with system constraints and with the coexistence of multiple communications.

8.1.7. ANR - Arburst

In this project Arburst - "Achievable region of bursty wireless networks" (2016-2020, 195 KEuros), we propose an original approach complementary to other existing projects. Instead of proposing one specific technical solution, our objective is to define a unified theoretical framework devoted to the study of IoT networks fundamental limits. We aim at establishing the fundamental limits for a decentralized system in a bursty regime which includes short packets of information and impulsive interference regime. We are targeting the fundamental limits, their mathematical expression (according to the usual information theory framework capturing the capacity region by establishing a converse and achievability theorems). We will use the recent results relative to finite block-length information theory and we will evaluate the margin for improvement between existing approaches and these limits and we will identify the scientific breakthrough that may bring significant improvements for IoT/M2M communications. This project will contribute to draw the roadmap for the development of IoT/M2M networks and will constitute a unified framework to compare existing techniques, and to identify the breakthrough concepts that may afford the industry the leverage to deploy IoT/M2M technical solutions.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. COM-MED

Title: COMMunication systems with renewable Energy micro-griD

Programm: H2020

Duration: October 2016 - October 2019

Coordinator: Inria

Inria contact: Samir M. Perlaza

A smart micro-grid is a small-scale power-grid system consisting of a number of distributed energy sources and loads which is responsible to ensure power sufficiency in a small area. The effectiveness of a smart micro-grid depends on the proper implementation of a communications and networking system which monitors, controls and manages the grid's operations. Due to the ever growing worldwide energy consumption, the need of an efficient framework for managing the way power is distributed and utilized has increased. The main objective of the project COM-MED is to study the fundamental interplay between communications and power networks in the context of smart micro-grids and renewable energy sources. On one hand, we study advanced signal processing techniques and communications methods to optimize the operation of smart micro-grid systems. On the other hand, we focus on mobile communications networks with renewable energy base-stations (BSs) and we investigate communications and networking techniques that take into account both data traffic and energy profiles to support high quality-of-service (QoS). The objectives of each technical WP have been assigned in such a way as to ensure that the project's target is realized during the project's time period. The theoretical results derived from the WPs 3, 4 and 5 will be tested using the telecommunication network of MTN in Cyprus but also the state-of-the-art equipment of the CITI/Inria research lab in France. The outcome of this project will provide a theoretical framework for the optimal cooperation between communications networks and power networks in the context of smart micro-grids and renewable energy sources. This is in line with the objectives of the call's theme "Renewable Energy" and is of paramount importance for the Mediterranean area. The consortium of the project has the expertise and the infrastructure to implement the objectives set and to bring the project to a successful end.

8.2.2. Collaborations in European Programs, Except FP7 & H2020

Socrate is very active in COST IRACON CA15104: Guillaume Villemaud is National Delegate (Alt.) and FIT/CorteXlab is identify as one of the COST platform: .

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners

Socrate has strong collaborations with several international partners.

- **Princeton University**, School of Applied Science, Department of Electrical Engineering, NJ. USA. This cooperation with Prof. H. Vincent Poor is on topics related to decentralized wireless networks. Samir M. Perlaza has been appointed as Visiting Research Collaborator at the EE Department for the academic period 2016-2017. Scientific-Leaders at Inria: Samir M. Perlaza and Jean-Marie Gorce.
- **Technical University of Berlin**, Dept. of Electrical Engineering and Computer Science, Germany. This cooperation with Prof. Rafael Schaffer is on secrecy and covert communications. Scientific-Leaders at Inria: Samir M. Perlaza.
- **National University Singapore (NUS)**, Department of Electrical and Computer Engineering, Singapore. This collaboration with Prof. Vincent Y. F. Tan is on the study of finite block-length transmissions in multi-user channels and the derivation of asymptotic capacity results with non-vanishing error probabilities. Scientific-Leaders at Inria: Samir M. Perlaza
- **University of Sheffield**, Department of Automatic Control and Systems Engineering, Sheffield, UK. This cooperation with Prof. Inaki Esnaola is on topics related to information-driven energy systems and multi-user information theory. Scientific-in-charge at Inria: Samir M. Perlaza.
- **Rutgers University**, Winlab, Orbit testbed. This cooperation with Ivan Seskar is related to experimental wireless testbed. Orbit has been one of the first wireless testbeds of its type. Tanguy Risset and Leonardo Sampaio-Cardoso have visited Winlab and I. Seskar visited the Socrate team for one week. Their collaboration is on the development of tools to ease experiment handling on wireless testbeds: visualisation, synchronization etc. Scientific-Leader at Inria: Tanguy Risset
- **University of Arizona**, Department of Electrical and Computer Engineering, Tucson, AZ, USA. This cooperation with Prof. Ravi Tandon is on topics related to channel-output feedback in wireless networks. Scientific-Leader at Inria: Samir M. Perlaza.
- **University of Cyprus**, Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus. This cooperation with Prof. Ioannis Krikidis is on topics related to energy-harvesting and wireless communications systems. Scientific-Leaders at Inria: Guillaume Villemaud and Samir M. Perlaza.
- **Universidade Federal do Ceará**, GTEL, Departamento de Teleinformática, Fortaleza, Brazil. This recently started cooperation with Prof. Tarcisio Ferreira Maciel is on topics related to the optimization of radio ressources for massive MIMO in 5G and 5G-like wireless communications systems. Scientific-in-charge at Inria: Leonardo Sampaio-Cardoso.
- **Universidad Nacional del Sur**, LaPSyC laboratory, Bahía Blanca , Argentina. This cooperation with Prof. Juan Cousseau is on topics related to Full-Duplex communications and Interference Alignment. Scientific-in-charge at Inria: Guillaume Villemaud.
- **Bell Labs New Jersey, USA**, This cooperation with Prof. Antonia Tulino (affiliated to Bell Labs and to University of Napoli, Italy) is on caching in wireless networks. The objective is to demonstrate the efficiency of caching at the edge of wireless networks through experimentations on CorteXlab. This work will be published in 2017 in a special issue of IEEE Communication magazine (Yasser Fadlallah, Antonia M. Tulino, Dario Barone, Giuseppe Vettigli, Jaime Llorca and Jean-Marie Gorce: Coding for caching in 5G networks, IEEE Communication Magazine, 2017, accepted for publication). Scientific leader at Inria : Jean-Marie Gorce.
- **Technical University "Gh. Asachi" of Iasi, Romania**, Department of Electronics, Telecommunications and Information Technology. This recent collaboration has started on topics related on the theoretical aspects of the ultra-low power radio communications. Scientific-in-charge at Inria: Florin Hutu

- **Queen's University Belfast**, UK. This collaboration is on molecular communication and massive MIMO with Prof. Trung Q. Duong. Scientific-in-charge at Inria: Malcolm Egan
- **Czech Technical University in Prague**, Czech Republic. This collaboration is on optimisation methods related to machine learning with Dr. Vyacheslav Kungurtsev. Scientific-in-charge at Inria: Malcolm Egan
- **TUMCREATE**, Singapore. This collaboration is on signal processing in communications with Dr. Ido Nevat. Scientific-in-charge at Inria: Malcolm Egan
- **telecommunications department of UMNG (Universidad Militar de Nueva Granada), Bogota, Colombia. Ongoing collaboration on security for GSM networks using deep learning. Scientific-in-charge at Inria: Leonardo Sampaio-Cardoso.**

8.4. International Research Visitors

8.4.1. Visits from International Teams

- Prof. Edward Guillen and his joint PhD student with Leonardo Sampaio-Cardoso, José Rugeles, came to Lyon in June 2018 for a 1-month academic stay, to develop work on FIT/CorteXlab for the security for GSM networks using deep learning project

8.4.2. Visits to International Teams

8.4.2.1. Sabbatical programme

Samir M. Perlaza has been on Sabatical year at Princeton University up to septembre 2018.

Malcolm Egan has been a visiting research collaborator in Prof. Poor's group in Princeton University March-April 2018.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

The Socrate team launched two new workshops in relation with FIT/CorteXlab:

- First French GNU Radio days: https://gnuradio-fr-18.sciencesconf.org/. The workshop aimed at the development of radio prototyping why GNU Radio in France and in Europe, by Leonardo Sampaio-Cardoso and Tanguy Risset.
- ISP-IoT : First Winter School on Information Theory and Signal Processing for Internet of Things : https://isp-iot.sciencesconf.org by Jean-Marie Gorce.

Socrate organized *Journée plateformes d'évaluations radio*, 28-03-2018, joint event between the GDR ASR and the GDR ISIS. Co-organized by Leonardo Sampaio-Cardoso and Jean-Marie Gorce with Thalès.

9.1.1.1. Member of the Organizing Committees

- Tanguy Risset, Leonardo Sampaio-Cardoso, Guillaume Villemaud and Jean-Marie Gorce were members of the organizing committee of the first French GNU Radio days.
- Samir M. Perlaza, Claire Goursaud, Leonardo Sampaio-Cardoso were members of the organizing committee of ISP-IOT, the first Winter School on Information Theory and Signal Processing for Internet of Things.
- Seed meeting *Dynamic Modeling and Simulation for Molecular Communication Networks*. Workshop funded by the French Embassy in the UK, London. Co-organized by Malcolm Egan. 17th and 18th September 2018, London UK.

9.1.1.2. Chair of Conference Program Committees

- Jean-Marie Gorce (chair) was co-chair of the ISP-IOT, the first Winter School on Information Theory and Signal Processing for Internet of Things.
- Samir M. Perlaza is chair of the special session on “Energy Harvesting and Wireless Powered Communications” hosted at the 20th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), July 2-5, 2019, Cannes, France.
- Samir M. Perlaza is co-chair of the Workshop on “Mathematical Tools for IoT Networks Modeling (MOTION)” hosted at the IEEE Wireless Communications and Networking Conference (WCNC). April, 15-18 2019, Marrakech, Morocco.
- Samir M. Perlaza is Publication Chair of the International Symposium on Information Theory (ISIT), July, 2019, Paris, France.

9.1.1.3. Member of Conference Program Committees

Tanguy Risset was a member of the following technical program committees:

- IEEE Computer Society Annual Symposium on VLSI (ISVLSI) 2018
- Design Automation and Test in Europe (DATE) 2018
- International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM) 2018
- International Conference on Advances in Cognitive Radio (COCORA) 2018.
- PIMRC 2018

Guillaume Villemaud was a member of the following technical program committees:

- CROWNCOM2018
- PIMRC 2018
- EUCAP 2018.

Florent de Dinechin was a member of the following technical program committees:

- 25th IEEE Symposium on Computer Arithmetic (Arith 2018)
- Design Automation and Test in Europe (DATE 2018) (as co-chair of track D11),
- 26th IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM 2018)

Malcolm Egan was a member of the following technical program committees:

- IEEE Global Communications Conference (Globecom), 9-13 December 2018, Abu Dhabi, UAE.
- International Conference on Advanced Technologies for Communications (ATC) 2018
- International Conference on Recent Advances on Signal Processing, Telecommunications & Computing (SigTelCom) 2018.

Jean-Marie Gorce was a member of the following technical program committees:

- IEEE International Conference on Communications (ICC), 20-24 May 2018, Kansas City, MO, USA.
- IEEE Global Communications Conference (Globecom), 9-13 December 2018, Abu Dhabi, UAE.
- IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 9-12 September 2018, Bologna, Italy.
- IEEE 5G World Forum 2018 (WF-5G), 9-11 July 2018, Santa Clara, California, USA
- International Conference on Telecommunications(ICT), June, 26-28, 2018, Saint-Malo, France.
- IEEE Wireless Communications and Networking Conference (WCNC). April, 15-18, 2018, Barcelona, Spain.
- IEEE 5G World Forum 2019 (WF-5G), 2019.
- International Conference on Telecommunications(ICT), 2019.

Samir M. Perlaza was a member of the following program committees :

- International Conference on Computing, Networking and Communications (ICNC), 18-21 February, 2019, Honolulu, Hawaii, USA.
- IEEE International Conference on Communications (ICC), 20-24 May 2019, Shanghai, China.
- Workshop on Green and Sustainable 5G Wireless Networks at IEEE Global Communications Conference (Globecom), 9-13 December 2018, Abu Dhabi, UAE.
- IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 9-12 September 2018, Bologna, Italy.

Claire Goursaud was a member of the following technical program committees:

- IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) 2018
- EAI International Conference on Body Area Networks (BodyNet) 2018

Leonardo Sampaio-Cardoso was a member of the following technical program committees:

- IEEE Wireless Communications and Networking Conference (WCNC) 2019
- IEEE International Conference on Communications (ICC) 2019

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

Guillaume Villemaud is an associate editor of Annals of Telecommunications (Springer).

Malcolm Egan is an associate editor of IEEE Communications Letters and guest editor for IEEE Access special section on Molecular Communication Networks.

Jean-Marie Gorce is an an associate editor of Journal of Wireless Communications and Networking.

Claire Goursaud is associate editor for ETT and ITL, Wiley.

Samir M. Perlaza is editor of the IEEE Transactions on Communications for the term 2018-2021 in the area of Simultaneous Information and Energy Transmission, Feedback and Applications of Game Theory in Wireless Communications.

Samir M. Perlaza is Associate Editor of the IET Smart Grid for the term 2018-2021.

9.1.3. Scientific Expertise

Guillaume Villemaud served as Research Expert for the European commission for the H2020-MSCA-NIGHT-2018.

Tanguy Risset is member of the Administration council (Conseil d'administration) of the GRAME institute (centre national de création musicale).

Guillaume Villemaud is a member of the Delphi Expert Panel on Software Defined Networks (SDN) and Network Functions Virtualisation (NFV).

Jean-Marie Gorce was appointed as scientific expert for Haute école spécialisée de Suisse occidentale and for FNRS (Belgium).

Jean-Marie Gorce is member of the Administration council (Conseil d'admistration) of ESISAR, Valence.

Jean-Marie Gorce was member of the jury for prix de thèse en signal-image de EEA-ISIS-GRETSI 2017.

Jean-Marie Gorce was the chair of a recruitment committee for a full professor position at Insa.

9.1.4. Invited Talks

Malcolm Egan gave invited talks in the Poor group in Princeton University and Prof. Erkip's group in NYU. With Samir M. Perlaza, he had an invited paper at CISS 2018.

Jean-Marie Gorce gave invited talks at IEMN/Ircica, invited by Prof. Laurent Clavier (June 2018), and at Orange Labs (Sept 2019).

Samir M. Perlaza gave a keynote on “Key Technologies in the IoT: Simultaneous Wireless Information and Energy Transmission” at the First Winter School on Information Theory and Signal Processing for Internet of Things, Villeurbanne, France, November 20, 2018.

Florent de Dinechin gave two talks at the FPGA days organized at LIP, Lyon, june 21, "Introduction to FPGA computing", then "computing just right on FPGAs"

Florent de Dinechin gave a talk at the RiscV days organized at CEA, October 16, Grenoble, "Dark silicon: a computer arithmetic perspective"

Florent de Dinechin gave a talk at the RiscV days organized at CEA, October 16, Grenoble, "Dark silicon: a computer arithmetic perspective"

Florent de Dinechin gave an invited lecture at the Advanced Workshop on FPGA-based Systems-On-Chip for Scientific Instrumentation and Reconfigurable Computing, organized by the International center for Theoretical Physics in Trieste, Italy, November 29, 2018.

Florent de Dinechin gave a talk at the Institute for Systems Programming of the Russian Academy of Science in Moscow, Russia, December 24, 2018.

Guillaume Salagnac and Tanguy Risset gave an invited talk at Spintec (Grenoble, FR), October 23rd, 2018, untitled "Peripheral State Persistence and Interrupt Management For Transiently Powered Systems".

Guillaume Villemaud gave an invited talk in the special session on Wake-up radio at the IEEE ICECS 2018 conference in december 2018

9.1.5. Research Administration

Jean-Marie Gorce is Adjoint to Scientific Director of Inria Rhônes-Alpes center.

Tanguy Risset is Vice-director of the FIL (CNRS Computer Science Research Federation of Lyon/Saint-Etienne).

Florent de Dinechin is director of the Citi-Lab.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Tanguy Risset and Jean-Marie Gorce are professors at the Telecommunications Department of Insa Lyon.

Florent de Dinechin is a professor at the Computer Science Department of Insa Lyon.

Claire Goursaud and Malcolm Egan are an associate professor at the Telecommunications department of Insa Lyon.

Leonardo Sampaio-Cardoso is an associate professor at the FIMI department as well as the Telecommunications Department of Insa Lyon

Guillaume Salagnac and Kevin Marquet are associate professors at the Computer Science Department of Insa Lyon.

Guillaume Villemaud and Florin Hutu are associate professor at the Electrical Engineering Department of Insa Lyon.

Samir M. Perlaza and Jean-Marie Gorce teach the course on Network Information Theory at École Normale Supérieure de Lyon.

9.2.2. Supervision

PhD in progress **Gautier Berthou***Operating system for transiently powered systems*, Inria, (IPL ZEP) since 01/2018.

PhD in progress **Tristan Delizy***memory management for normally-of NV-RAM based systems*, Insa-Lyon, (Region ARC6) since 09/2016.

PhD in progress **Yohan Uguen***Synthesis of arithmetic operators*, Insa-Lyon, (Mininstry of research) since 09/2016.

PhD in progress **David Kibloff***New strategy for Physical Layer Security in wireless networks: self-jamming using Full-Duplex Transceivers*, École Doctorale EEA de Lyon, funded by Inria-DGA grant since 10/2015.

PhD in progress **Nizar Khalfet***Stochastic Energy Sources to Power Communication Systems*, École Doctorale EEA de Lyon, funded by EU Project COM-MED since 10/2016.

PhD in progress : **Andrea Bocco**: *Porposition d'une unité de caclul U-NUM pour le calcul scientifique*, ANR Metalibm grant, since 12/2016.

PhD in progress : **Hassan Kallam**: *Topology aided multi-user interference management in wireless network* , Fed4PMR Insavalor project grant, since 01/2017.

PhD in progress : **Anade Akpo Dadja**: *Non assymptotic fundemental limits of bursty communications*, ANR Arburst grant, since 10/2017.

PhD in progress : **Diane Duchemin**: *Distributed coding in dense IoT Network*, ANR Ephyl grant, since 01/2017.

PhD starting **Luc Forget** : *Algèbre linéaire calculant au plus juste*, ANR Imprenum, since 10/2018.

PhD starting **Yanni Zhou** : *Full Duplex and spatial modulation*

PhD starting **Tarik Lassouaoui**: *Tag 2 Tag communication*

PhD starting **Regis Rousseau**: *Wireless Power Transfer*

PhD defended: **Yuqi Mo***Scaling of Iot Communication issues*, École Doctorale EEA,sept. 2018.

9.2.3. Juries

- Guillaume Salagnac was examiner in the jury of the PhD Nadir Cherify "Assistance au développement de logiciels embarqués contraints en énergie"
- Florin Hutu was examiner int the PhD Jury of Frank Itoua, defended on March 12, 2018 at Limoges University.
- Tanguy Risset was a member of the jury of the following theses:
  - Maxime France-Pillois (U. Grenoble)
  - Perrin Njoyah Ntafam (U. Grenoble)
  - Simon Rokicki (U. Rennes)
- Florent de Dinechin was a reviewer in the jury of Hugues de Lassus Saint-Genis, defended on May 17 2018 at Université de Perpignan Via Domitia.
- Malcolm Egan was a co-supervisor in the jury of Mauro de Freitas, defended at Université de Lille 1.
- Claire Goursaud and Jean-Marie Gorce were co-supervisors in the jury of Yuqi Mo, defended at Insa Lyon.
- Jean-Marie Gorce was :
  - a reviewer of the HdR jury of Guillaume Andrieux (IETR, Nantes U., December 2018).
  - a reviewer of the PhD jury of Philippe Ezran (Supelec Paris Saclay U., January 2018).
  - a reviewer of the PhD jury of Antony Pottier (IMT Atlantique, Nov 2018).
  - a reviewer of the PhD jury of Luca Feltrin (Bologna U., Italy, December 2018).
  - an examiner of the PhD jury of Xavier Leturc (IMT Paris, U. Paris Saclay, December 2018),

- Guillaume Villemaud was a reviewer for the PhD defense of
  - Andreina Liendo (Univ. Grenoble)
  - Mai-Thanh TRAN (Univ. Rennes 1)
  - Manuel Milla Peinado (Univ. Poitiers)
- Claire Goursaud was a member in the jury of Alaa Khreis, defended at l'Université Paris Saclay, prepared at Télécom ParisTech .
- Claire Goursaud was a reviewer in the jury of Remi Bonnefoi, defended at CentraleSupelec

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] G. BERTHOU, T. DELIZY, K. MARQUET, T. RISSET, G. SALAGNAC.*Sytare: a Lightweight Kernel for NVRAM-Based Transiently-Powered Systems*, in "IEEE Transactions on Computers", 2018, p. 1-14, https://hal.archives-ouvertes.fr/hal-01954979

[2] M. EGAN, L. CLAVIER, C. ZHENG, M. DE FREITAS, J.-M. GORCE.*Dynamic Interference for Uplink SCMA in Large-Scale Wireless Networks without Coordination*, in "EURASIP Journal on Wireless Communications and Networking", August 2018, vol. 2018, n^{o} 1, p. 1-14 [*DOI :* 10.1186/S13638-018-1225-Z], https://hal.archives-ouvertes.fr/hal-01871576

[3] M. EGAN, J. DRCHAL, J. MRKOS, M. JAKOB.*Towards Data-Driven On-Demand Transport*, in "EAI Endorsed Transactions on Industrial Networks and Intelligent Systems", June 2018, vol. 5, n^{o} 14, p. 1-10 [*DOI :* 10.4108/EAI.27-6-2018.154835], https://hal.archives-ouvertes.fr/hal-01839452

[4] M. EGAN, T. C. MAI, T. Q. DUONG, M. DI RENZO.*Coexistence in Molecular Communications*, in "Nano Communication Networks", February 2018, vol. 16, p. 37-44 [*DOI :* 10.1016/J.NANCOM.2018.02.006], https://hal.archives-ouvertes.fr/hal-01650966

[5] N. KHALFET, S. M. PERLAZA.*Simultaneous Information and Energy Transmission in the Two-User Gaussian Interference Channel*, in "IEEE Journal on Selected Areas in Communications", September 2018, p. 1-15 [*DOI :* 10.1109/JSAC.2018.2872365], https://hal.archives-ouvertes.fr/hal-01874019

[6] V. QUINTERO, S. PERLAZA, I. ESNAOLA, J.-M. M. GORCE.*Approximate Capacity Region of the Two-User Gaussian Interference Channel with Noisy Channel-Output Feedback*, in "IEEE Transactions on Information Theory", July 2018, vol. 64, n^{o} 7, p. 5326-5358, Part of this work was presented at the IEEE International Workshop on Information Theory (ITW), Cambridge, United Kingdom, September 2016 and IEEE International Workshop on Information Theory (ITW), Jeju Island, Korea, October, 2015. Parts of this work appear in Inria Technical Report Number 0456, 2015, and Inria Research Report Number 8861. [*DOI :* 10.1109/TIT.2018.2827076], https://hal.archives-ouvertes.fr/hal-01397118

[7] A. VOLKOVA, M. ISTOAN, F. DE DINECHIN, T. HILAIRE.*Towards Hardware IIR Filters Computing Just Right: Direct Form I Case Study*, in "IEEE Transactions on Computers", 2018 [*DOI :* 10.1109/TC.2018.2879432], https://hal.sorbonne-universite.fr/hal-01561052

Invited Conferences

[8] Y. MO, C. GOURSAUD, J.-M. GORCE.*Uplink Multiple Base Stations Diversity for UNB based IoT networks*, in "CAMA 2018 - IEEE International Conference on Antenna Measurement and Applications", Västerås, Sweden, September 2018, p. 1-4, https://hal.inria.fr/hal-01887640

[9] M. S. WIDMAIER, F. HUTU, G. VILLEMAUD.*Efficiency of Orthogonal Codes for Quasi-passive Wake-Up Radio Receivers using Frequency Footprint IDs*, in "25th IEEE International Conference on Electronics Circuits and Systems", Bordeaux, France, December 2018, https://hal.archives-ouvertes.fr/hal-01973114

International Conferences with Proceedings

[10] T. DELIZY, S. GROS, K. MARQUET, M. MOY, T. RISSET, G. SALAGNAC.*Estimating the Impact of Architectural and Software Design Choices on Dynamic Allocation of Heterogeneous Memories*, in "RSP 2018 - 29th International Symposium on Rapid System Prototyping", Turin, Italy, October 2018, p. 1-7, https://hal.archives-ouvertes.fr/hal-01891599

[11] D. DUCHEMIN, J.-M. GORCE, C. GOURSAUD.*Code Domain Non Orthogonal Multiple Access versus ALOHA: a simulation based study*, in "ICT 2018 - 25th International Conference on Telecommunications", Saint-Malo, France, 25th international Conference on Communications, IEEE, June 2018, p. 445-450 [*DOI :* 10.1109/ICT.2018.8464836], https://hal.inria.fr/hal-01801103

[12] A. DUQUE, R. STANICA, H. RIVANO, C. GOURSAUD, A. DESPORTES.*Poster: Insights into RGB-LED to Smartphone Communication*, in "EWSN 2018 - International Conference on Embedded Wireless Systems and Networks", Madrid, Spain, ACM, February 2018, p. 173-174, https://hal.inria.fr/hal-01683605

[13] A. DYTSO, M. EGAN, S. PERLAZA, H. V. POOR, S. SHAMAI.*Optimal Inputs for Some Classes of Degraded Wiretap Channels*, in "2018 IEEE Information Theory Workshop (ITW)", Guangzhou, China, November 2018, https://hal.inria.fr/hal-01884159

[14] M. EGAN, T. C. MAI, T. Q. DUONG, M. DI RENZO.*Coordination via Advection Dynamics in Nanonetworks with Molecular Communication*, in "ICC 2018 - IEEE International Conference on Communications", Kansas City, United States, IEEE, May 2018, p. 1-6 [*DOI :* 10.1109/ICC.2018.8422573], https://hal.archives-ouvertes.fr/hal-01696268

[15] M. GOUTAY, L. CARDOSO, C. GOURSAUD.*Massive Machine Type Communications Uplink Traffic: Impact of Beamforming at the Base Station*, in "ICT 2018 - 25th International Conference on Telecommunications", Saint Malo, France, IEEE, June 2018, p. 493-497 [*DOI :* 10.1109/ICT.2018.8464894], https://hal.inria.fr/hal-01875596

[16] F. HUTU, G. VILLEMAUD.*On the use of the FBMC modulation to increase the performance of a wake-up radio*, in "IEEE Radio and Wireless Symposium (RWS 2018)", Anaheim, CA, United States, IEEE, January 2018, p. 139-142 [*DOI :* 10.1109/RWS.2018.8304968], https://hal.archives-ouvertes.fr/hal-01845177

[17] *Best Paper*
M. KUMM, O. GUSTAFSSON, F. DE DINECHIN, J. KAPPAUF, P. ZIPF.*Karatsuba with Rectangular Multipliers for FPGAs*, in "ARITH 2018 - 25th IEEE International Symposium on Computer Arithmetic", Amherst, United States, IEEE, June 2018, p. 13-20, Best paper award [*DOI :* 10.1109/ARITH.2018.8464809], https://hal.inria.fr/hal-01773447.

[18] Y. MO, C. GOURSAUD, J.-M. GORCE.*Multiple Base Stations Diversity for UNB Systems: Theoretical Analysis and Performances*, in "ISNCC 2018 - International Symposium on Networks, Computers and Communications", Rome, Italy, June 2018, p. 1-6, https://hal.inria.fr/hal-01887619

[19] S. POOJARY, R. EL-AZOUZI, E. ALTMAN, A. SUNNY, I. TRIKI, M. HADDAD, T. JIMENEZ, S. VALENTIN, D. TSILIMANTOS.*Analysis of QoE for adaptive video streaming over wireless networks*, in "WiOpt 2018 - 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks", Shanghai, China, May 2018, https://hal.inria.fr/hal-01972647

[20] T. RISSET, C. GOURSAUD, X. BRUN, K. MARQUET, F. MEYER.*UWB Ranging for Rapid Movements*, in "IPIN 2018 - International Conference on Indoor Positioning and Indoor Navigation", Nantes, France, September 2018, p. 1-8 [*DOI :* 10.1109/IPIN.2018.8533820], https://hal.inria.fr/hal-01951454

[21] R. ROUSSEAU, F. HUTU, G. VILLEMAUD.*On the Use of Vector Fitting and State-Space Modeling to Maximize the DC Power Collected by a Wireless Power Transfer System*, in "AT-RASC 2018 - 2nd URSI Atlantic Radio conference", Grande Canarie, Spain, May 2018, p. 1-4, https://hal.archives-ouvertes.fr/hal-01903660

[22] Y. UGUEN, E. PETIT.*PyGA: a Python to FPGA compiler prototype*, in "AI-SEPS 2018 - 5th ACM SIGPLAN International Workshop on Artificial Intelligence and Empirical Methods for Software Engineering and Parallel Computing Systems", Boston, United States, ACM, November 2018, p. 11-15 [*DOI :* 10.1145/3281070.3281072], https://hal.inria.fr/hal-01959155

[23] G. VILLEMAUD, F. HUTU, P. BELLOCHE, F. KNINECH.*Wireless Transmission in Ventilation (HVAC) Ducts for the Internet of Things and Smarter Buildings: Proof of Concept and Specific Antenna Design*, in "12th European Conference on Antennas and Propagation", Londres, United Kingdom, April 2018, https://hal.archives-ouvertes.fr/hal-01973176

National Conferences with Proceeding

[24] F. DE DINECHIN, M. DARRIN, A. DUDERMEL, S. MICHELLAND, A. REYNAUD.*Une architecture minimisant les échanges entre processeur et mémoire*, in "ComPAS 2018 - Conférence d'informatique en Parallélisme, Architecture et Système", Toulouse, France, July 2018, p. 1-8, https://hal.inria.fr/hal-01959855

Conferences without Proceedings

[25] G. BERTHOU, T. DELIZY, K. MARQUET, T. RISSET, G. SALAGNAC.*Peripheral State Persistence and Interrupt Management For Transiently Powered Systems*, in "NVMW 2018 - 9th Annual Non-Volatile Memories Workshop", San Diego, United States, March 2018, p. 1-2, https://hal.archives-ouvertes.fr/hal-01943919

[26] D. BUJOREANU, Y. M. BENANE, B. NICOLAS, H. LIEBGOTT, J.-M. GORCE, L. CARDOSO, D. FRIBOULET.*Jointly Optimized Modulatiion/Filtering Technique for Pseudo-Orthogonal Binary Sequences*, in "2018 IEEE International Ultrasonics Symposium (IUS)", Kobe, Japan, IEEE, October 2018, p. 1-5, https://hal.archives-ouvertes.fr/hal-01987969

[27] M. CUNCHE, L. SAMPAIO CARDOSO.*Analyzing Ultrasound-based Physical Tracking Systems*, in "GreHack 2018", Grenoble, France, November 2018, https://hal.inria.fr/hal-01927513

[28] T. DELIZY, S. GROS, K. MARQUET, M. MOY, T. RISSET, G. SALAGNAC.*Quels objets en NVRAM ? Placement en mémoires de travail hétérogènes*, in "Compas 2018 - Conférence d'informatique en Parallélisme, Architecture et Système", Toulouse, France, July 2018, p. 1-8, https://hal.archives-ouvertes.fr/hal-01891398

[29] M. EGAN, T. Q. DUONG, M. DI RENZO, J.-M. GORCE, I. NEVAT, V. LOSCRI.*Cognitive Molecular Communication*, in "2018 -3rd Workshop on Molecular Communications", Ghent, Belgium, April 2018, p. 1-2, https://hal.archives-ouvertes.fr/hal-01728766

[30] M. EGAN.*On Regular Schemes and Tight Frames*, in "SETA 2018 - Sequences and Their Applications", Hong Kong, Hong Kong SAR China, October 2018, p. 1-12, https://hal.archives-ouvertes.fr/hal-01839448

[31] M. EGAN, S. PERLAZA.*Capacity Approximation of Continuous Channels by Discrete Inputs*, in "CISS 2018 - 52nd Annual Conference on Information Sciences and Systems", Princeton, United States, IEEE, March 2018, p. 1-6 [*DOI :* 10.1109/CISS.2018.8362269], https://hal.archives-ouvertes.fr/hal-01686036

[32] B. MASSOT, F. HUTU, C. GEHIN, N. NOURY.*Design and Optimization of an Autonomous, Ambulatory Cardiac Event Monitor*, in "IEEE International Conference on E-health Networking, Application & Services", Ostrava, Czech Republic, September 2018, https://hal.archives-ouvertes.fr/hal-01879334

[33] Y. MO, C. GOURSAUD, J.-M. GORCE.*Bénéfice de l'Annulation Successive d'Interférence pour des Réseaux Ultra Narrow Band : Théorie et Application à l'IoT*, in "ALGOTEL 2018 - 20èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications", Roscoff, France, May 2018, p. 1-4, https://hal.archives-ouvertes.fr/hal-01783923

[34] S. PERLAZA, A. TAJER, H. VINCENT POOR.*Simultaneous Information and Energy Transmission: A Finite Block-Length Analysis*, in "SPAWC 2018 - IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications", Kalamata, Greece, IEEE, June 2018, p. 1-5 [*DOI :* 10.1109/SPAWC.2018.8446028], https://hal.archives-ouvertes.fr/hal-01722794

[35] V. QUINTERO, S. M. PERLAZA, J.-M. GORCE, H. VINCENT POOR.*Approximate Nash Region of the Gaussian Interference Channel with Noisy Output Feedback*, in "ITW 2018 - IEEE Information Theory Workshop", Guangzhou, China, IEEE, November 2018, p. 1-5, https://hal.archives-ouvertes.fr/hal-01876880

[36] F. TRAHAY, M. SELVA, L. MOREL, K. MARQUET.*NumaMMA: NUMA MeMory Analyzer*, in "ICPP 2018 - 47th International Conference on Parallel Processing", Eugene, United States, August 2018, p. 1-10 [*DOI :* 10.1145/3225058.3225094], https://hal-cea.archives-ouvertes.fr/cea-01854072

[37] G. VILLEMAUD, F. HUTU, P. BELLOCHE, F. KNINECH.*Wireless Transmission in Ventilation (HVAC) Ducts for the Internet of Things and Smarter Buildings: Proof of Concept and Specific AntennaDesign*, in "IRACON 6th MCM and 6th Technical Meeting", Nicosie, Cyprus, January 2018, https://hal.archives-ouvertes.fr/hal-01973399

Scientific Books (or Scientific Book chapters)

[38] J.-M. MULLER, N. BRUNIE, F. DE DINECHIN, C.-P. JEANNEROD, M. JOLDES, V. LEFÈVRE, G. MELQUIOND, N. REVOL, S. TORRES.*Handbook of Floating-point Arithmetic (2nd edition)*, Birkhäuser Basel, July 2018, p. 1-627 [*DOI :* 10.1007/978-3-319-76526-6], https://hal.inria.fr/hal-01766584

Other Publications

[39] G. BERTHOU, A. CARER, H.-P. CHARLES, S. DERRIEN, K. MARQUET, I. MIRO-PANADES, D. PALA, I. PUAUT, F. RASTELLO, T. RISSET, E. ROHOU, G. SALAGNAC, O. SENTIEYS, B. YARAHMADI.*The Inria ZEP project: NVRAM and Harvesting for Zero Power Computations*, March 2018, 1, NVMW 2018 - 10th Annual Non-Volatile Memories Workshop, Poster, https://hal.inria.fr/hal-01941766

[40] M. CUNCHE, L. S. CARDOSO.*Analysis of an Ultrasound-Based Physical Tracking System*, May 2018, working paper or preprint, https://hal.inria.fr/hal-01798091

[41] K. SUN, I. ESNAOLA, S. PERLAZA, H. VINCENT POOR.*Stealth Attacks on the Smart Grid*, August 2018, working paper or preprint, https://hal.inria.fr/hal-01857366

[42] F. DE DINECHIN, L. FORGET, J.-M. MULLER, Y. UGUEN.*Posits: the good, the bad and the ugly*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01959581

References in notes

[43] I. AKYILDIZ, W.-Y. LEE, M. VURAN, S. MOHANTY.*A survey on spectrum management in cognitive radio networks*, in "Communications Magazine, IEEE", April 2008, vol. 46, n^{o} 4, p. 40 -48

[44] K. BIESECKER, J. DOBIAC, N. FEKADU, M. JONES, C. KAIN, K. RAMAN.*Software Defined Radio Roadmap*, 2008, Noblis Technical Report, for National Institute of Justice, USA

[45] DECAWAVE.*DW1000 Datasheet*, 2014 V2.04, 1

[46] EQUIPEX.*Future internet of things*, 2011, http://fit-equipex.fr/

[47] F. HAMMER, R. YUDANTO, K. NEUMANN, M. PICHLER, J. COCKX, C. NIESTROJ, F. PETRÉ.*Performance Evaluation of 3D-Position Estimation Systems*, in "IEEE Sensors Journal", Aug 2016, vol. 16, n^{o} 16, p. 6416-6424

[48] A. R. JIMÉNEZ, F. SECO.*Comparing Decawave and Bespoon UWB location systems: Indoor/outdoor performance analysis*, in "2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)", Oct 2016, p. 1-8

[49] JOINT PROGRAM EXECUTIVE OFFICE (JPEO), JOINT TACTICAL RADIO SYSTEM (JTRS).*Software Communications Architecture Specification*, 2006, JTRS Standards, version 2.2.2

[50] J. S. MEENA, S. M. SZE, U. CHAND, T.-Y. TSENG.*Overview of emerging nonvolatile memory technologies*, in "Nanoscale Research Letters", 2014, vol. 9, 526

[51] J. MITOLA III.*The software radio*, in "IEEE National Telesystems Conference", 1992

[52] J. MITOLA III.*Software radios: Survey, critical evaluation and future directions*, in "Aerospace and Electronic Systems Magazine, IEEE", Apr 1993, vol. 8, n^{o} 4, p. 25-36

[53] P. STEENKISTE, D. SICKER, G. MINDEN, D. RAYCHAUDHURI.*Future Directions in Cognitive Radio Network Research*, 2009, NSF Workshop Report

[54] M. WEISER.*The computer for the 21st Century*, in "Pervasive Computing, IEEE", January 2002, vol. 99, n^{o} 1, p. 19 -25

[55] M. Z. WIN, R. A. SCHOLTZ.*Impulse radio: how it works*, in "IEEE Communications Letters", 1998, vol. 2, p. 36–38

[56] Q. ZHANG, F. H. P. FITZEK, V. B. IVERSEN.*Cognitive radio MAC protocol for WLAN*, in "Proceedings of the IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC", 2008, p. 1-6

Project-Team SPADES

Sound Programming of Adaptive Dependable Embedded Systems

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

IN PARTNERSHIP WITH:
Institut polytechnique de Grenoble

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Embedded and Real-time Systems

Table of contents

Project-Team SPADES

Creation of the Team: 2013 January 01, updated into Project-Team: 2015 July 01

Keywords:

Computer Science and Digital Science:

A1.1.1. - Multicore, Manycore
A1.1.9. - Fault tolerant systems
A1.3. - Distributed Systems
A2.1.1. - Semantics of programming languages
A2.1.6. - Concurrent programming
A2.1.9. - Synchronous languages
A2.3. - Embedded and cyber-physical systems
A2.3.1. - Embedded systems
A2.3.2. - Cyber-physical systems
A2.3.3. - Real-time systems
A2.4.1. - Analysis
A2.4.3. - Proofs
A2.5.2. - Component-based Design
A7.3. - Calculability and computability

Other Research Topics and Application Domains:

B5.2.1. - Road vehicles
B6.3.3. - Network Management
B6.4. - Internet of things
B6.6. - Embedded systems

1. Team, Visitors, External Collaborators

Research Scientists

Gregor Goessler [Team leader, Inria, Researcher, HDR]
Pascal Fradet [Inria, Researcher, HDR]
Alain Girault [Inria, Senior Researcher, HDR]
Sophie Quinton [Inria, Researcher]
Jean-Bernard Stefani [Inria, Senior Researcher]

Faculty Member

Xavier Nicollin [Institut polytechnique de Grenoble, Associate Professor]

Technical Staff

Souha Ben Rayana [Inria, from May 2018]

PhD Students

Xiaojie Guo [Univ. Grenoble Alpes]
Maxime Lesourd [Univ. Grenoble Alpes]
Stephan Plassart [Univ. Grenoble Alpes]
Christophe Prévot [Thales until Oct 2018, Inria from Nov 2018]
Arash Shafiei [Orange Labs]
Martin Vassor [Inria]

Post-Doctoral Fellows
Jia Jie Wang [Inria, from May 2018]
Nicolas Hili [IRT Saint-Exupery]

Visiting Scientist
Ismail Assayad [Univ. Hassan II, Casablanca, Sep 2018]

Administrative Assistant
Helen Pouchot-Rouge-Blanc [Inria]

2. Overall Objectives

2.1. Overall Objectives

The SPADES project-team aims at contributing to meet the challenge of designing and programming dependable embedded systems in an increasingly distributed and dynamic context. Specifically, by exploiting formal methods and techniques, SPADES aims to answer three key questions:

1. How to program open networked embedded systems as dynamic adaptive modular structures?
2. How to program reactive systems with real-time and resource constraints on multicore architectures?
3. How to program reliable, fault-tolerant embedded systems with different levels of criticality?

These questions above are not new, but answering them in the context of modern embedded systems, which are increasingly distributed, open and dynamic in nature [24], makes them more pressing and more difficult to address: the targeted system properties – dynamic modularity, time-predictability, energy efficiency, and fault-tolerance – are largely antagonistic (*e.g.*, having a highly dynamic software structure is at variance with ensuring that resource and behavioral constraints are met). Tackling these questions together is crucial to address this antagonism, and constitutes a key point of the SPADES research program.

A few remarks are in order:

- We consider these questions to be central in the construction of future embedded systems, dealing as they are with, roughly, software architecture and the provision of real-time and fault-tolerance guarantees. Building a safety-critical embedded system cannot avoid dealing with these three concerns.
- The three questions above are highly connected. For instance, composability along time, resource consumption and reliability dimensions are key to the success of a component-based approach to embedded systems construction.
- For us, “Programming” means any constructive process to build a running system. It can encompass traditional programming as well as high-level design or “model-based engineering” activities, provided that the latter are supported by effective compiling tools to produce a running system.
- We aim to provide semantically sound programming tools for embedded systems. This translates into an emphasis on formal methods and tools for the development of provably dependable systems.

3. Research Program

3.1. Introduction

The SPADES research program is organized around three main themes, *Design and Programming Models*, *Certified real-time programming*, and *Fault management and causal analysis*, that seek to answer the three key questions identified in Section 2.1. We plan to do so by developing and/or building on programming languages and techniques based on formal methods and formal semantics (hence the use of *“sound programming”* in the project-team title). In particular, we seek to support design where correctness is obtained by construction, relying on proven tools and verified constructs, with programming languages and programming abstractions designed with verification in mind.

3.2. Design and Programming Models

Work on this theme aims to develop models , languages and tools to support a "correct-by-construction" approach to the development of embedded systems.

On the programming side, we focus on the definition of domain specific programming models and languages supporting static analyses for the computation of precise resource bounds for program executions. We propose dataflow models supporting dynamicity while enjoying effective analyses. In particular, we study parametric extensions where properties such as liveness and boundedness remain statically analyzable.

On the design side, we focus on the definition of component-based models for software architectures combining distribution, dynamicity, real-time and fault-tolerant aspects. Component-based construction has long been advocated as a key approach to the "correct-by-construction" design of complex embedded systems [49]. Witness component-based toolsets such as PTOLEMY [38], BIP [30], or the modular architecture frameworks used, for instance, in the automotive industry (AUTOSAR) [22]. For building large, complex systems, a key feature of component-based construction is the ability to associate with components a set of *contracts*, which can be understood as rich behavioral types that can be composed and verified to guarantee a component assemblage will meet desired properties.

Formal models for component-based design are an active area of research. However, we are still missing a comprehensive formal model and its associated behavioral theory able to deal *at the same time* with different forms of composition, dynamic component structures, and quantitative constraints (such as timing, fault-tolerance, or energy consumption).

We plan to develop our component theory by progressing on two fronts: a semantical framework and domain-specific programming models. The work on the semantical framework should, in the longer term, provide abstract mathematical models for the more operational and linguistic analysis afforded by component calculi. Our work on component theory will find its application in the development of a COQ-based toolchain for the certified design and construction of dependable embedded systems, which constitutes our first main objective for this axis.

3.3. Certified Real-Time Programming

Programming real-time systems (*i.e.*, systems whose correct behavior depends on meeting timing constraints) requires appropriate languages (as exemplified by the family of synchronous languages [32]), but also the support of efficient scheduling policies, execution time and schedulability analyses to guarantee real-time constraints (*e.g.*, deadlines) while making the most effective use of available (processing, memory, or networking) resources. Schedulability analysis involves analyzing the worst-case behavior of real-time tasks under a given scheduling algorithm and is crucial to guarantee that time constraints are met in any possible execution of the system. Reactive programming and real-time scheduling and schedulability for multiprocessor systems are old subjects, but they are nowhere as mature as their uniprocessor counterparts, and still feature a number of open research questions [28], [36], in particular in relation with mixed criticality systems. The main goal in this theme is to address several of these open questions.

We intend to focus on two issues: multicriteria scheduling on multiprocessors, and schedulability analysis for real-time multiprocessor systems. Beyond real-time aspects, multiprocessor environments, and multicore ones in particular, are subject to several constraints *in conjunction*, typically involving real-time, reliability and energy-efficiency constraints, making the scheduling problem more complex for both the offline and the online cases. Schedulability analysis for multiprocessor systems, in particular for systems with mixed criticality tasks, is still very much an open research area.

Distributed reactive programming is rightly singled out as a major open issue in the recent, but heavily biased (it essentially ignores recent research in synchronous and dataflow programming), survey by Bainomugisha et al. [28]. For our part, we intend to focus on devising synchronous programming languages for distributed systems and precision-timed architectures.

3.4. Fault Management and Causal Analysis

Managing faults is a clear and present necessity in networked embedded systems. At the hardware level, modern multicore architectures are manufactured using inherently unreliable technologies [33], [43]. The evolution of embedded systems towards increasingly distributed architectures highlighted in the introductory section means that dealing with partial failures, as in Web-based distributed systems, becomes an important issue.

In this axis we intend to address the question of *how to cope with faults and failures in embedded systems?*. We will tackle this question by exploiting reversible programming models and by developing techniques for fault ascription and explanation in component-based systems.

A common theme in this axis is the use and exploitation of causality information. Causality, *i.e.*, the logical dependence of an effect on a cause, has long been studied in disciplines such as philosophy [53], natural sciences, law [54], and statistics [55], but it has only recently emerged as an important focus of research in computer science. The analysis of logical causality has applications in many areas of computer science. For instance, tracking and analyzing logical causality between events in the execution of a concurrent system is required to ensure reversibility [52], to allow the diagnosis of faults in a complex concurrent system [47], or to enforce accountability [51], that is, designing systems in such a way that it can be determined without ambiguity whether a required safety or security property has been violated, and why. More generally, the goal of fault-tolerance can be understood as being to prevent certain causal chains from occurring by designing systems such that each causal chain either has its premises outside of the fault model (*e.g.*, by introducing redundancy [45]), or is broken (*e.g.*, by limiting fault propagation [57]).

4. Application Domains

4.1. Industrial Applications

Our applications are in the embedded system area, typically: transportation, energy production, robotics, telecommunications, the Internet of things (IoT), systems on chip (SoC). In some areas, safety is critical, and motivates the investment in formal methods and techniques for design. But even in less critical contexts, like telecommunications and multimedia, these techniques can be beneficial in improving the efficiency and the quality of designs, as well as the cost of the programming and the validation processes.

Industrial acceptance of formal techniques, as well as their deployment, goes necessarily through their usability by specialists of the application domain, rather than of the formal techniques themselves. Hence, we are looking to propose domain-specific (but generic) realistic models, validated through experience (*e.g.*, control tasks systems), based on formal techniques with a high degree of automation (*e.g.*, synchronous models), and tailored for concrete functionalities (*e.g.*, code generation).

4.2. Industrial Design Tools

The commercially available design tools (such as UML with real-time extensions, MATLAB/ SIMULINK/ dSPACE[0]) and execution platforms (OS such as VXWORKS, QNX, real-time versions of LINUX ...) start now to provide, besides their core functionalities, design or verification methods. Some of them, founded on models of reactive systems, come close to tools with a formal basis, such as for example STATEMATE by iLOGIX.

Regarding the synchronous approach, commercial tools are available: SCADE[0] (based on LUSTRE), CONTROLBUILD and RT-BUILDER (based on SIGNAL) from GEENSYS[0] (part of DASSAULTSYSTEMES), specialized environments like CELLCONTROL for industrial automatism (by the Inria spin-off ATHYS– now part of DASSAULTSYSTEMES). One can observe that behind the variety of actors, there is a real consistency of the synchronous technology, which makes sure that the results of our work related to the synchronous approach are not restricted to some language due to compatibility issues.

[0] http://www.dspaceinc.com
[0] http://www.esterel-technologies.com
[0] http://www.geensoft.com

4.3. Current Industrial Cooperations

Regarding applications and case studies with industrial end-users of our techniques, we cooperate with Thales on schedulability analysis for evolving or underspecified real-time embedded systems, with Orange Labs on software architecture for cloud services and with Daimler on reduction of nondeterminism and analysis of deadline miss models for the design of automotive systems.

5. New Software and Platforms

5.1. pyCPA_TWCA

Analysis tool for weakly-hard real-time systems
KEYWORDS: Real time - Scheduling analyses
FUNCTIONAL DESCRIPTION: pyCPA_TWCA is a pyCPA plugin for Typical Worst-Case Analysis. pyCPA is an open-source Python implementation of Compositional Performance Analysis developed at TU Braunschweig, which allows in particular response-time analysis. pyCPA_TWCA is an extension of that tool that is co-developed by Sophie Quinton and Zain Hammadeh at TU Braunschweig. It allows in particular the computation of weakly-hard guarantees for real-time tasks, i.e. number of deadline misses out of a sequence of executions. So far, pyCPA_TWCA is restricted to uniprocessor systems of independent tasks. pyCPA_TWCA can handle the following scheduling policies: Fixed Priority Preemptive, Fixed Priority Non-Preemptive, Weighted Round-Robin, Earliest Deadline First.

- Contact: Sophie Quinton

5.2. CertiCAN

Certifier of CAN bus analysis results
KEYWORDS: Certification - CAN bus - Real time - Static analysis
FUNCTIONAL DESCRIPTION: CertiCAN is a tool, produced using the Coq proof assistant, allowing the formal certification of the correctness of CAN bus analysis results. Result certification is a process that is lightweight and flexible compared to tool certification, which makes it a practical choice for industrial purposes. The analysis underlying CertiCAN, which is based on a combined use of two well-known CAN analysis techniques, is computationally efficient. Experiments demonstrate that CertiCAN is able to certify the results of RTaW-Pegase, an industrial CAN analysis tool, even for large systems. Furthermore, CertiCAN can certify the results of any other RTA tool for the same analysis and system model (periodic tasks with offsets in transactions).

- Contact: Xiaojie Guo

6. New Results

6.1. Design and Programming Models

Participants: Pascal Fradet, Alain Girault, Gregor Goessler, Xavier Nicollin, Christophe Prévot, Sophie Quinton, Arash Shafiei, Jean-Bernard Stefani, Martin Vassor, Souha Ben Rayana.

6.1.1. A multiview contract theory for cyber-physical system design and verification

The design and verification of critical cyber-physical systems is based on a number of models (and corresponding analysis techniques and tools) representing different viewpoints such as function, timing, security and many more. Overall correctness is guaranteed by mostly informal, and therefore basic, arguments about the relationship between these viewpoint-specific models. More precisely, the assumptions that a viewpoint-specific analysis makes on the other viewpoints remain mostly implicit, and whenever explicit they are handled mostly manually. In [11], we argue that the current design process over-constrains the set of possible system designs and that there is a need for methods and tools to formally relate viewpoint-specific models and corresponding analysis results. We believe that a more flexible contract-based approach could lead to easier integration, to relaxed assumptions, and consequently to more cost efficient systems while preserving the current modelling approach and its tools.

The framework we have in mind would provide viewpoint specific contract patterns guaranteeing inter-viewpoint consistency in a flexible manner. At this point, most of the work remains to be done. On the application side, we need a more complete picture of existing inter-viewpoint models. We also need the theory required for the correctness proofs, but it should be based on the needs on the application side.

6.1.2. End-to-end worst-case latencies of task chains for flexibility analysis

In collaboration with Thales, we address the issue of change during design and after deployment in safety-critical embedded system applications. More precisely, we focus on timing aspects with the objective to anticipate, at design time, future software evolutions and identify potential schedulability bottlenecks. The work presented in this section is the PhD topic of Christophe Prévot, in the context of a collaboration with Thales TRT, and our algorithms are being implemented in the Thales tool chain, in order to be used in industry.

This year, we have completed our work on the analysis of end-to-end worst-case latencies of task chains [10] that was needed to extend our approach for quantifying the flexibility, with respect to timing, of real-time systems made of chains of tasks. In a nutshell, flexibility is the property of a given system to accommodate changes in the future, for instance the modification of some of the parameters of the system, or the addition of a new task in the case of a real-time system.

One major issue that hinders the use of performance analysis in industrial design processes is the pessimism inherent to any analysis technique that applies to realistic system models (*e.g.*, , systems with task chains). Indeed, such analyses may conservatively declare unschedulable systems that will in fact never miss any deadlines. The two main avenues for improving this are (i) computing tighter upper bounds on the worst-case latencies, and (ii) measuring the pessimism, which requires to compute also guaranteed lower bounds. A lower bound is guaranteed by providing an actual system execution exhibiting a behavior as close to the worst case as possible. As a first step, we focus in [10] on uniprocessor systems executing a set of sporadic or periodic hard real-time task chains. Each task has its own priority, and the chains are scheduled according to the fixed-priority preemptive scheduling policy. Computing the worst-case end-to-end latency of each chain is complex because of the intricate relationship between the task priorities. Compared to state of the art analyses, we propose here tighter upper bounds, as well as lower bounds on these worst-case latencies. Our experiments show the relevance of lower bounds on the worst-case behavior for the industrial design of real-time embedded systems.

Based on our end-to-end latency analysis for task chains, we have also proposed an extension of the concept of slack to task chains and shown how it can be used to perform flexibility analysis and sensitivity analysis. This solution is particularly relevant for industry as it provides means by which the system designer can anticipate the impact on timing of software evolutions, at design time as well as after deployment.

6.1.3. Location graphs

We have introduced the location graph model [58] as an expressive framework for the definition of component-based models able to deal with dynamic software configurations with sharing and encapsulation constraints. We have completed a first study of the location graph behavioral theory (under submission), initiated its formalization in Coq, and an implementation of the location framework with an emphasis of the expression of different isolation and encapsulation constraints.

We are now studying conservative extensions to the location graph framework to support the compositional design of heterogeneous hybrid dynamical systems and their attendant notions of approximate simulations [60].

In collaboration with the Spirals team at Inria Lille – Nord Europe, we have applied the location framework for the definition of a pivot model for the description of software configurations in a cloud computing environment. We have shown how to interpret in our pivot model several configuration management models and languages including TOSCA, OCCI, Docker Compose, Aeolus, OpenStack HOT.

6.1.4. Dynamicity in dataflow models

Recent dataflow programming environments support applications whose behavior is characterized by dynamic variations in resource requirements. The high expressive power of the underlying models (*e.g.*, Kahn Process Networks or the CAL actor language) makes it challenging to ensure predictable behavior. In particular, checking *liveness* (*i.e.*, no part of the system will deadlock) and *boundedness* (*i.e.*, the system can be executed in finite memory) is known to be hard or even undecidable for such models. This situation is troublesome for the design of high-quality embedded systems. In the past few years, we have proposed several parametric dataflow models of computation (MoCs) [40], [31], we have written a survey providing a comprehensive description of the existing parametric dataflow MoCs [34], and we have studied *symbolic* analyses of dataflow graphs [35]. More recently, we have proposed an original method to deal with lossy communication channels in dataflow graphs [39].

We are now studying models allowing dynamic reconfigurations of the *topology* of the dataflow graphs. In particular, many modern streaming applications have a strong need for reconfigurability, for instance to accommodate changes in the input data, the control objectives, or the environment.

We have proposed a new MoC called Reconfigurable Dataflow (RDF) [15]. RDF extends SDF with transformation rules that specify how the topology and actors of the graph may be reconfigured. Starting from an initial RDF graph and a set of transformation rules, an arbitrary number of new RDF graphs can be generated at runtime. The major quality of RDF is that it can be statically analyzed to guarantee that all possible graphs generated at runtime will be connected, consistent, and live. This is the research topic of Arash Shafiei's PhD, in collaboration with Orange Labs.

6.1.5. Monotonic prefix consistency in distributed systems

We have studied the issue of data consistency in distributed systems. Specifically, we have considered a distributed system that replicates its data at multiple sites, which is prone to partitions, and which is assumed to be available (in the sense that queries are always eventually answered). In such a setting, strong consistency, where all replicas of the system apply synchronously every operation, is not possible to implement. However, many weaker consistency criteria that allow a greater number of behaviors than strong consistency, are implementable in available distributed systems. We have focused on determining the strongest consistency criterion that can be implemented in a convergent and available distributed system that tolerates partitions, and we have shown that no criterion stronger than Monotonic Prefix Consistency (MPC [61], [44]) can be implemented [18].

6.2. Certified Real-Time Programming

Participants: Pascal Fradet, Alain Girault, Gregor Goessler, Xavier Nicollin, Sophie Quinton, Xiaojie Guo, Maxime Lesourd.

6.2.1. Time predictable programming languages and architectures

Time predictability (PRET) is a topic that emerged in 2007 as a solution to the ever increasing unpredictability of today's embedded processors, which results from features such as multi-level caches or deep pipelines [37]. For many real-time systems, it is mandatory to compute a strict bound on the program's execution time. Yet, in general, computing a tight bound is extremely difficult [64]. The rationale of PRET is to simplify both the programming language and the execution platform to allow more precise execution times to be easily computed [27].

We have extended the PRET-C compiler [25] in order to make it energy aware. To achieve this, we use dynamic voltage and frequency scaling (DFVS) and we insert DVFS control points in the control flow graph of the PRET-C program. Several difficulties arise: (i) the control flow graph is concurrent, (ii) the resulting optimization problem is a time and energy multi-criteria problem, and (iii) since we consider PRET-C programs, we actually address the Worst-Case Execution Time (WCET) and the Worst-Case Energy Consumption (WCEC). Thanks to a novel ILP formulation and to a bicriteria heuristic, we are able to address the two objectives jointly and to compute, for each PRET-C program, the Pareto front of the non-dominated solutions in the 2D space (WCET,WCEC) [63]. We have recently improved this result to reduce the complexity of the algorithm and to produce the *optimal* Pareto front. This is the topic of Jia Jie Wang's postdoc.

Moreover, within the CAPHCA project, we have proposed a new approach for predictable inter-core communication between tasks allocated on different cores. Our approach is based on the execution of synchronous programs written in the FOREC programming language on deterministic architectures called PREcision Timed. The originality resides in the time-triggered model of computation and communication that allows for a very precise control over the thread execution. Synchronisation is done via configurable Time Division Multiple Access (TDMA) arbitrations (either physical or conceptual) where the optimal size and offset of the time slots are computed to reduce the inter-core synchronization costs. Results show that our model guarantees time-predictable inter-core communication, the absence of concurrent accesses (without relying on hardware mechanisms), and allows for optimized execution throughput. This is the topic of Nicolas Hili's postdoc.

6.2.2. Schedulability of weakly-hard real-time systems

We focus on the problem of computing tight deadline miss models for real-time systems, which bound the number of potential deadline misses in a given sequence of activations of a task. In practical applications, such guarantees are often sufficient because many systems are in fact not hard real-time [4]. A weakly-hard real-time guarantee specifies an upper bound on the maximum number m of deadline misses of a task in a sequence of k consecutive executions. Based on our previous work on Typical Worst-Case Analysis [4], [8], we have introduced in [13] the first verification method which is able to provide weakly-hard real-time guarantees for tasks and task chains in systems with multiple resources under partitioned scheduling with fixed priorities. All existing weakly-hard real-time verification techniques are restricted today to systems with a single resource. Our verification method is applied in the context of switched networks with traffic streams between nodes, and we demonstrate its practical applicability on an automotive case study.

6.2.3. Synthesis of switching controllers using approximately bisimilar multiscale abstractions

The use of discrete abstractions for continuous dynamics has become standard in hybrid systems design (see *e.g.*, [60] and the references therein). The main advantage of this approach is that it offers the possibility to leverage controller synthesis techniques developed in the areas of supervisory control of discrete-event systems [56]. The first attempts to compute discrete abstractions for hybrid systems were based on traditional systems behavioral relationships such as simulation or bisimulation, initially proposed for discrete systems most notably in the area of formal methods. These notions require inclusion or equivalence of observed behaviors which is often too restrictive when dealing with systems observed over metric spaces. For such systems, a more natural abstraction requirement is to ask for closeness of observed behaviors. This leads to the notions of approximate simulation and bisimulation introduced in [42]. These approaches are based on sampling of time and space where the sampling parameters must satisfy some relation in order to obtain abstractions of a prescribed precision. In particular, the smaller the time sampling parameter, the finer the lattice used for approximating the state-space; this may result in abstractions with a very large number of states when the sampling period is small. However, there are a number of applications where sampling has to be fast; though this is generally necessary only on a small part of the state-space.

We are currently investigating an approach using mode sequences as symbolic states for our abstractions. By using mode sequences of variable length we are able to adapt the granularity of our abstraction to the dynamics of the system, so as to automatically trade off precision against controllability of the abstract states.

6.2.4. A Markov Decision Process approach for energy minimization policies

In the context of independent real-time sporadic jobs running on a single-core processor equipped with Dynamic Voltage and Frequency Scaling (DVFS), we have proposed a Markov Decision Process approach (MDP) to compute the scheduling policy that dynamically chooses the voltage and frequency level of the processor such that each job meets its deadline and the total energy consumption is minimized. We distinguish two cases: the finite case (there is a fixed time horizon) and the infinite case. In the finite case, several *offline* solutions exist, which all use the complete knowledge of all the jobs that will arrive within the time horizon [65], *i.e.*, their size and deadlines. But clearly this is unrealistic in the embedded context where the characteristics of the jobs are not known in advance. Then, an optimal offline policy called Optimal Available (OA) has been proposed in [65]. Our goal was to improve this result by taking into account the *statistical characteristics* of the upcoming jobs. When such information is available (for instance by profiling the jobs based on execution traces), we have proposed several speed policies that optimize the *expected* energy consumption. We have shown that this general constrained optimization problem can be modeled as an unconstrained MDP by choosing a proper state space that also encodes the constraints of the problem. In particular, this implies that the optimal speed at each time can be computed using a *dynamic programming* algorithm (under a finite horizon), and that the optimal speed at any time t will be a deterministic function of the current state at time t [41]. Under an infinite horizon, we use a *Value Iteration* algorithm.

This work led us to compare several existing speed policies with respect to their feasibility. Indeed, the policies (OA) [65], (AVR) [65], and (BKP) [29] all assume that the maximal speed S_{max} available on the processor is infinite, which is an unrealistic assumption. For these three policies and for our (MDP) policy, we have established necessary and sufficient conditions on S_{max} guaranteeing that no job will ever miss its deadline.

This is the topic of Stephan Plassart's PhD, funded by the CASERM Persyval project.

6.2.5. Formal proofs for schedulability analysis of real-time systems

We have started to lay the foundations for computer-assisted formal verification of real-time systems analyses. Specifically, we contribute to Prosa [23], a Coq library of reusable concepts and proofs for real-time systems analysis. A key scientific challenge is to achieve a modular structure of proofs, *e.g.*, for response time analysis. Our goal is to use this library for:

1. a better understanding of the role played by some assumptions in existing proofs;
2. a formal verification and comparison of different analysis techniques; and
3. the certification of results of existing (*e.g.*, industrial) analysis tools.

Our first major result [16] is a task model that generalizes the digraph model [59] and its corresponding analysis for fixed-priority scheduling with limited preemption. The motivation for this work, which is not yet fully proven in Coq, is to obtain a formally verified schedulability analysis for a very expressive task model. In the context of computer assisted verification, it permits to factorize the correctness proofs of a large number of analyses. The digraph task model seems a good candidate due to its powerful expressivity. Alas, its ability to capture dependencies between arrival and execution times of jobs of different tasks is very limited. Our extended model can capture dependencies between jobs of the same task as well as jobs of different tasks. We provide a correctness proof of the analysis that is written in a way amenable to its formalization in the Coq proof assistant. Despite being much more general, the Response Time Analysis (RTA) for our model is not significantly more complex than the original one. Also, it underlines similarities between existing analyses, in particular the analysis for the digraph model and Tindell's offset model [62].

A second major result is CertiCAN, a tool produced using Coq for the formal certification of CAN analysis results. Result certification is a process that is light-weight and flexible compared to tool certification, which makes it a practical choice for industrial purposes. The analysis underlying CertiCAN is based on a combined use of two well-known CAN analysis techniques [62] that makes it computationally efficient. Experiments demonstrate that CertiCAN is able to certify the results of RTaW-Pegase, an industrial CAN analysis tool, even for large systems. This result paves the way for a broader acceptance of formal tools for the certification of real-time systems analysis results. Beyond CertiCAN, we believe that this work is significant in that it demonstrates the advantage of result certification over tool certification for the RTA of CAN buses. In addition, the underlying technique can be reused for any other system model for which there exist RTAs with different levels of precision. This work will be presented at RTAS 2019.

In parallel, we have completed and published in [17] a Coq formalization of Typical Worst-Case Analysis (TWCA) [4], [8], an analysis technique for weakly-hard real-time systems. Our generic analysis is based on an abstract model that characterizes the exact properties needed to make TWCA applicable to any system model.

Our results are formalized and checked using the Coq proof assistant along with the Prosa schedulability analysis library. This work opens up new research directions for TWCA by providing a formal framework for the trade-off that must be found between time efficiency and precision of the analysis. Hopefully, our generic proof will make it easier to extend TWCA to more complex models in the future. In addition, our experience with formalizing real-time systems analyses shows that it is not only a way to increase confidence in the results of the analyses; it also helps understanding their key intermediate steps, the exact assumptions required, and how they can be generalized.

6.2.6. Logical execution time

In collaboration with TU Braunschweig and Daimler, we have worked on the application of the Logical Execution Time (LET) paradigm [50], according to which data are read and written at predefined time instants, to the automotive industry. The LET paradigm was considered until recently by the automotive industry as not efficient enough in terms of buffer space and timing performance. The shift to embedded multicore processors has represented a game changer: The design and verification of multicore systems is a challenging area of research that is still very much in progress. Predictability clearly is a crucial issue which cannot be tackled without changes in the design process. Several OEMs and suppliers have come to the conclusion that LET might be a key enabler and a standardization effort is already under way in the automotive community to integrate LET into AUTOSAR. We have organized a Dagstuhl seminar [9] to discuss and sketch solutions to the problems raised by the use of LET in multicore systems. A white paper on the topic is under preparation.

So far, LET has been applied only at the ECU (Electronic Control Unit) level by the automotive industry. Recent developments in electric powertrains and autonomous vehicle functions raise parallel programming from the multicore level to the vehicle level where the standard LET approach cannot apply directly. We have proposed System Level LET [21], an extension of LET with relaxed synchronization requirements which allows separating network design from ECU design and makes LET applicable to automotive distributed systems.

6.2.7. Scheduling under multiple constraints and Pareto optimization

We have continued our work on multi-criteria scheduling, in two directions. First, in the context of dynamic applications that are launched and terminated on an embedded homogeneous multi-core chip, under execution time and energy consumption constraints, we have proposed a two layer adaptive scheduling method [26]. In the first layer, each application (represented as a DAG of tasks) is scheduled statically on subsets of cores: 2 cores, 3 cores, 4 cores, and so on. For each size of these sets (2, 3, 4, ...), there may be only one topology or several topologies. For instance, for 2 or 3 cores there is only one topology (a “line”), while for 4 cores there are three distinct topologies (“line”, “square”, and “T shape”). Moreover, for each topology, we generate statically several schedules, each one subject to a different total energy consumption constraint, and consequently with a different Worst-Case Reaction Time (WCRT). Coping with the energy consumption constraints is achieved thanks to Dynamic Frequency and Voltage Scaling (DVFS). In the second layer, we use these pre-generated static schedules to reconfigure dynamically the applications running on the multi-core each time a new application is launched or an existing one is stopped. The goal of the second layer is to perform a dynamic global optimization of the configuration, such that each running application meets a pre-defined quality-of-service constraint (translated into an upper bound on its WCRT) and such that the total energy consumption be minimized. For this, we *(i)* allocate a sufficient number of cores to each active application, *(ii)* allocate the unassigned cores to the applications yielding the largest gain in energy, and *(iii)* choose for each application the best topology for its subset of cores (*i.e.*, better than the by default “line” topology). This is a joint work with Ismail Assayad (U. Casablanca, Morocco) who visited the team in 2018.

Second, we have proposed the first of its kind multi-criteria scheduling heuristics for a DAG of tasks onto an homogeneous multi-core chip. Given an application modeled as a Directed Acyclic Graph (DAG) of tasks and a multicore architecture, we produce a set of non-dominated (in the Pareto sense) static schedules of this DAG onto this multicore. The criteria we address are the execution time, reliability, power consumption, and peak temperature. These criteria exhibit complex antagonistic relations, which make the problem challenging. For instance, improving the reliability requires adding some redundancy in the schedule, which penalizes the execution time. To produce Pareto fronts in this 4-dimension space, we transform three of the four criteria into constraints (the reliability, the power consumption, and the peak temperature), and we minimize the fourth one (the execution time of the schedule) under these three constraints. By varying the thresholds used for the three constraints, we are able to produce a Pareto front of non-dominated solutions. Each Pareto optimum is a static schedule of the DAG onto the multicore. We propose two algorithms to compute static schedules. The first is a ready list scheduling heuristic called ERPOT (Execution time, Reliability, POwer consumption and Temperature). ERPOT actively replicates the tasks to increase the reliability, uses Dynamic Voltage and Frequency Scaling to decrease the power consumption, and inserts cooling times to control the peak temperature. The second algorithm uses an Integer Linear Programming (ILP) program to compute an optimal schedule. However, because our multi-criteria scheduling problem is NP-complete, the ILP algorithm is limited to very small problem instances. Comparisons showed that the schedules produced by ERPOT are on average only 10% worse than the optimal schedules computed by the ILP program, and that ERPOT outperforms the PowerPerf-PET heuristic from the literature on average by 33%. This is a joint work with Athena Abdi and Hamid Zarandi from Amirkabir University in Tehran, Iran.

6.3. Fault Management and Causal Analysis

Participants: Pascal Fradet, Alain Girault, Gregor Goessler, Jean-Bernard Stefani, Martin Vassor.

6.3.1. Fault Ascription in Concurrent Systems

The failure of one component may entail a cascade of failures in other components; several components may also fail independently. In such cases, elucidating the exact scenario that led to the failure is a complex and tedious task that requires significant expertise.

The notion of causality *(did an event e cause an event e'?)* has been studied in many disciplines, including philosophy, logic, statistics, and law. The definitions of causality studied in these disciplines usually amount to variants of the counterfactual test "e is a cause of e' if both e and e' have occurred, and in a world that is as close as possible to the actual world but where e does not occur, e' does not occur either". In computer science, almost all definitions of logical causality — including the landmark definition of [48] and its derivatives — rely on a causal model that. However, this model may not be known, for instance in presence of black-box components. For such systems, we have been developing a framework for blaming that helps us establish the causal relationship between component failures and system failures, given an observed system execution trace. The analysis is based on a formalization of counterfactual reasoning [6].

We are currently working on a revised version of our general semantic framework for fault ascription in [46] that satisfies a set of formally stated requirements — such as its behavior under several notions of abstraction and refinement —, and on its instantiation to acyclic models of computation, in order to compare our approach with the standard definition of *actual causality* proposed by Halpern and Pearl.

6.3.2. Fault Management in Virtualized Networks

From a more applied point of view we are investigating, in the context of Sihem Cherrared's PhD thesis, approaches for fault explanation and localization in virtualized networks. In essence, Network Function Virtualization (NFV), widely adopted by the industry and the standardization bodies, is about running network functions as software workloads on commodity hardware to optimize deployment costs and simplify the life-cycle management of network functions. However, it introduces new fault management challenges including dynamic topology and multi-tenant fault isolation that we discuss in [14]. As a first step to tackle those challenges, we have extended the classical fault management process to the virtualized functions by introducing LUMEN: a Global Fault Management Framework. Our approach aims at providing the availability and reliability of the virtualized 5G end-to-end service chain. LUMEN includes the canonical steps of the fault management process and proposes a monitoring solution for all types of Network virtualization Environments. Our framework is based on open source solutions and could easily be integrated with other existing autonomic management models.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- Inria and Orange Labs have established in 2015 a joint virtual research laboratory, called I/O LAB. We have been heavily involved in the creation of the laboratory and are actively involved in its operation (Jean-Bernard Stefani is one of the two co-directors of the lab). I/O LAB focuses on the network virtualization and cloudification. As part of the work of I/O LAB, we have cooperated with Orange Lab, as part of a cooperative research contract funded by Orange, on defining architectural principles and frameworks for network cloud infrastructures encompassing control and management of computing, storage and network resources.
- With Daimler (subcontracting via iUTBS): We have proposed, in collaboration with TU Braunschweig, an extension of the LET paradigm [50], called *System-level LET, to accommodate the specific needs of the design process in the automotive industry, in which the network structure must be made explicit in the LET program.*

7.2. Bilateral Grants with Industry

With Thales: Early performance assessment for evolving and variable cyber-physical systems. This CIFRE grant funds the PhD of Christophe Prévot.

With Orange: Programming IoT and sofware defined radio with dynamic dataflow models of computation. This CIFRE grant funds the PhD of Arash Shafiei.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. CASERM (PERSYVAL-Lab project)

Participants: Pascal Fradet, Alain Girault, Gregor Goessler, Xiaojie Guo, Maxime Lesourd, Xavier Nicollin, Stephan Plassart, Sophie Quinton, Jean-Bernard Stefani, Martin Vassor.

Despite recent advances, there exists currently no integrated formal methods and tools for the design and analysis of reconfigurable multi-view embedded systems. This is the goal of the CASERM project.

The CASERM project represents a significant effort towards a COQ-based design method for reconfigurable multi-view embedded systems, in order to formalize the structure and behavior of systems and to prove their main properties. The use of a proof assistant to support such a framework is motivated by the fact that the targeted systems are both extremely complex and critical. The challenges addressed are threefold:

1. to model software architectures for embedded systems taking into account their dynamicity and multiple constraints (functional as well as non functional);
2. to propose novel scheduling techniques for dynamically reconfiguring embedded systems; and
3. to advance the state of the art in automated proving for such systems.

The objectives of CASERM that address these challenges are organized in three tasks. They consist respectively in designing an architecture description framework based on a process calculus, in proposing online optimization methods for dynamic reconfiguration systems (this is the topic of Stephan Plassart's PhD), and in developing a formal framework for real-time analysis in the COQ proof assistant (this is the topic of Xiaojie Guo's and Maxime Lesourd's PhD). A fourth task focuses on common case studies for the evaluation of the obtained results.

The CASERM consortium gathers researchers from the LIG and VERIMAG laboratories who are reknowned specialists in these fields. The project started in November 2016 and will last three years.

8.2. National Initiatives

8.2.1. ANR

8.2.1.1. RT-Proofs

Participants: Pascal Fradet, Xiaojie Guo, Maxime Lesourd, Sophie Quinton.

RT-Proofs is an ANR/DFG project between Inria, MPI-SWS, Onera, TU Braunschweig and Verimag, running from 2018 until 2020.

The overall objective of the RT-Proofs project is to lay the foundations for computer-assisted formal verification of timing analysis results. More precisely, the goal is to provide:

1. a strong formal basis for schedulability, blocking, and response-time analysis supported by the Coq proof assistant, that is as generic, robust, and modular as possible;
2. correctness proofs for new and well-established generalized response-time analysis results, and a better, precise understanding of the role played by key assumptions and formal connections between competing analysis techniques;
3. an approach for the generation of proof certificates so that analysis results – in contrast to analysis tools – can be certified.

8.2.1.2. DCore

Participants: Gregor Goessler, Jean-Bernard Stefani.

DCORE is an ANR project between Inria project teams ANTIQUE, FOCUS and SPADES, and the IRIF lab, running from 2019 to 2023.

The overall objective of the project is to develop a semantically well-founded, novel form of concurrent debugging, which we call *causal debugging*, that aims to alleviate the deficiencies of current debugging techniques for large concurrent software systems. The causal debugging technology developed by DCORE will comprise and integrate two main novel engines:

1. *a reversible execution engine* that allows programmers to backtrack and replay a concurrent or distributed program execution, in a way that is both precise and efficient (only the exact threads involved by a return to a target anterior or posterior program state are impacted);
2. a *causal analysis engine* that allows programmers to analyze concurrent executions, by asking questions of the form “what caused the violation of this program property?”, and that allows for the precise and efficient investigation of past and potential program executions.

8.2.2. Institute of Technology (IRT)

8.2.2.1. CAPHCA

Participants: Alain Girault, Nicolas Hili.

CAPHCA is a project within the Antoine de Saint Exupéry IRT. The general objective of the project is to provide methods and tools to achieve performance and determinism on modern, high-performance, multi-core and FPGA-enabled SOCs. Our specific contribution lies withing work pakacges dedicated to the design of novel PRET architectures and programming languages (see Section 6.2.1).

8.3. European Initiatives

8.3.1. Collaborations in European Programs, Except FP7 & H2020

Program: Celtic-Plus

Project acronym: SENDATE

Project title: Secure Networking for a Data center cloud in Europe

Duration: April 2016 - March 2019

Coordinator: Nokia France

Other partners: Nokia, Orange, IMT, Inria

Abstract: The SENDATE project aims to develop a clean-slate architecture for converged telecommunications networks and distributed data centers supporting 5G cellular networks and the needs from the Industrial Internet and the Internet of Things. It aims to provide scientific and technical solutions for intra and inter data centrers security, control, management and orchestration, placement and management of virtual network functions, as well as high-speed transport networks for data centers access and interconnection.

8.3.2. Collaborations with Major European Organizations

We have a strong collaboration with the Technische Universität Braunschweig in Germany. In particular, Sophie Quinton is involved in the CCC project (http://ccc-project.org/) to provide methods and mechanisms for the verification of software updates after deployment in safety-critical systems, and in the TypicalCPA project which aims at computing deadline miss models for distributed systems.

We also have a recent collaboration with the MPI-SWS in Kaiserslautern (Germany) on formal proofs for real-time systems. This collaboration will be concretized by an ANR-PRCI project called RT-PROOFS starting in 2018, which involves MPI-SWS, TU Braunschweig, Inria, and Onera.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

- Ismail Assayad (from U. Casablanca, Morocco) visited the team for one month in September 2018, to work on a two layer adaptive scheduling method.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events Organisation

9.1.1.1. General Chair, Scientific Chair

- Alain Girault is member of the steering committee of the International Federated Conference on Distributed Computing Techniques (DISCOTEC) and of the ACM International Conference on Embedded Software (EMSOFT).
- Gregor Gössler is member of the steering committee of the International Workshop on Causal Reasoning for Embedded and Safety-critical Systems Technologies (CREST).
- Jean-Bernard Stefani is the current chair of the steering committee of the IFIP FORTE international conference series, a member of the steering committee of the IFIP DISCOTEC conference series, and the current chair of the IFIP Working Group 6.1.

9.1.1.2. Member of the Organizing Committees

- Sophie Quinton was the co-organizer of a Dagstuhl seminar entitled “The Logical Execution Time Paradigm: New Perspectives for Multicore Systems”. https://www.dagstuhl.de/18092

9.1.2. Scientific Events Selection

9.1.2.1. Member of the Conference Program Committees

- Alain Girault served in the program committees of the Symposium on Industrial Embedded Systems (SIES’18), the Forum on specification and Design Languages (FDL’18), and the Conference on Applications of Concurrency to System Design (ACSD’18).

- Gregor Gössler served in the program committees of the 18th International Workshop on Automated Verification of Critical Systems (AVOCS 2018) and the 3rd international Workshop on Formal Reasoning about Causation, Responsibility, and Explanations in Science and Technology (CREST 2018).
- Sophie Quinton served in the program committees of the 30th Euromicro Conference on Real-Time Systems (ECRTS'18), the 9th International Workshop on Analysis Tools and Methodologies for Embedded and Real-time Systems (WATERS'18), the 39th IEEE Real-Time Systems Symposium (RTSS'18) and the 26th International Conference on Real-Time Networks and Systems (RTNS'18).

9.1.2.2. Reviewer

- Alain Girault reviewed papers for the ECRTS'18 conference.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Alain Girault is a member of the editorial board of the Journal on Embedded Systems.

9.1.3.2. Reviewer - Reviewing Activities

- Alain Girault reviewed articles for J. of Transportation Technologies (JTT) and IEEE Trans. Dependable and Secure Computing (TDSC).
- Gregor Gössler reviewed articles for IEEE Transactions on Automatic Control (TAC) and ACM Transactions on Embedded Computing Systems (TECS).
- Sophie Quinton reviewed an article for ACM Trans. on Embedded Computing Systems (TECS).

9.1.4. Research Administration

- Pascal Fradet is head of the committee for doctoral studies ("Responsable du comité des études doctorales") of the Inria Grenoble – Rhône-Alpes research center and local correspondent for the young researchers Inria mission ("Mission jeunes chercheurs").
- Alain Girault is vice-chair of the Inria Evaluation Committee.
- Xavier Nicollin is member of the committee for computing resources users ("Comité des Utilisateurs des Moyens Informatiques") of the Inria Grenoble – Rhône-Alpes research center.
- Jean-Bernard Stefani is head of science (délégué scientifique) of the Inria Grenoble – Rhône-Alpes research center and a member of the Inria Evaluation Committee.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence : Pascal Fradet, Théorie des Langages 1 & 2, 36 HeqTD, niveau L3, Grenoble INP (Ensimag), France

Licence : Pascal Fradet, Modèles de Calcul : λ-calcul, 12 HeqTD, niveau L3, Univ. Grenoble Alpes, France

Master : Pascal Fradet, Langages et Traducteurs, 16 HeqTD, niveau M1, Polytech Grenoble, Univ. Grenoble Alpes, France

Master : Xavier Nicollin, Sémantique et Analyse des Programmes, 45 HeqTD, niveau M1, Grenoble INP (Ensimag), France

Licence : Xavier Nicollin, Théorie des Langages 2, 36 HeqTD, niveau L3, Grenoble INP (Ensimag), France

Licence : Xavier Nicollin, Bases de la Programmation Impérative, 81 HeqTD (2017-2018), niveau L3, Grenoble INP (Ensimag), France

Licence : Sophie Quinton, Théorie des Langages 2, 18 HeqTD, niveau L3, Grenoble INP (Ensimag), France

Master : Sophie Quinton, Performance and Quantitative Properties, 6h, MOSIG, Univ. Grenoble Alpes, France

Master: Jean-Bernard Stefani, Formal Aspects of Component Software, 9h, MOSIG, Univ. Grenoble Alpes, France.

9.2.2. Supervision

- PhD in progress: Sihem Cherrared, "Fault Management in Multi-Tenant Programmable Networks", Univ. Rennes 1, since October 2016, co-advised by Eric Fabre and Gregor Gössler.
- PhD in progress: Christophe Prévot, "Early Performance assessment for evolving and variable Cyber-Physical Systems", Univ. Grenoble Alpes, since November 2015, co-advised by Alain Girault and Sophie Quinton.
- PhD in progress: Stephan Plassart, "On-line optimization in dynamic real-time systems", Univ. Grenoble Alpes, since September 2016, co-advised by Bruno Gaujal and Alain Girault.
- PhD in progress: Xiaojie Guo, "Formal Proofs for the Analysis of Real-Time Systems in COQ", Univ. Grenoble Alpes, since December 2016, co-advised by Pascal Fradet, Jean-François Monin, and Sophie Quinton.
- PhD in progress: Maxime Lesourd, "Generic Proofs for the Analysis of Real-Time Systems in COQ", Univ. Grenoble Alpes, since September 2017, co-advised by Pascal Fradet, Jean-François Monin, and Sophie Quinton.
- PhD in progress: Arash Shafiei, "Programming IoT and sofware defined radio with dynamic dataflow models of computation", Univ. Grenoble Alpes, since September 2017, co-advised by Pascal Fradet, Alain Girault, and Xavier Nicollin.
- PhD in progress: Martin Vassor, "Analysis and types for safe dynamic software reconfigurations", Univ. Grenoble Alpes, since November 2017, co-advised by Pascal Fradet and Jean-Bernard Stefani.
- M2 SIF in progress: T. Mari, "From diagnosis to causal analysis", U. Rennes, since November 2018, co-supervised by Gregor Gössler and Louise Travé-Massuyès (LAAS).
- PFE: Clément Arvis, "Génération automatique de musique", Grenoble INP/Ensimag, September 2018, supervized by Sophie Quinton.

9.2.3. Juries

- Alain Girault was referee for the PhD thesis of Colin Vidal, Université Côte d'Azur, and for the PhD thesis of Julien Hascoet, INSA Rennes. He was also vice-president of the Inria Senior Researcher jury (DR2) and of the Inria Junior Researcher national jury (CRCN).
- Gregor Gössler was examiner for the PhD jury of Vincent Wang (U. Pennsylvania).
- Jean-Bernard Stefani was examiner for the Habilitation (HDR) jury of Thomas Ledoux (U. Nantes).
- Sophie Quinton was member of the CRCN jury in Rennes.

9.3. Popularization

9.3.1. Interventions

Sophie Quinton gave a keynote at the MathC2+ event organized by Inria, entitled "Faire des preuves par ordinateur : Pourquoi et comment ?" (Computer-assisted proofs: Why and how?).

10. Bibliography

Major publications by the team in recent years

[1] S. ANDALAM, P. ROOP, A. GIRAULT, C. TRAULSEN.*A Predictable Framework for Safety-Critical Embedded Systems*, in "IEEE Trans. on Computers", July 2014, vol. 63, n^{o} 7, p. 1600–1612

[2] A. BOUAKAZ, P. FRADET, A. GIRAULT.*A Survey of Parametric Dataflow Models of Computation*, in "ACM Trans. Design Autom. Electr. Syst.", 2017, vol. 22, n^o^ 2, p. 38:1–38:25, https://doi.org/10.1145/2999539

[3] S. DJOKO DJOKO, R. DOUENCE, P. FRADET.*Aspects preserving properties*, in "Science of Computer Programming", 2012, vol. 77, n^o^ 3, p. 393-422

[4] G. FREHSE, A. HAMANN, S. QUINTON, M. WÖHRLE.*Formal Analysis of Timing Effects on Closed-loop Properties of Control Software*, in "35th IEEE Real-Time Systems Symposium 2014 (RTSS)", Rome, Italy, December 2014, https://hal.inria.fr/hal-01097622

[5] A. GIRARD, G. GÖSSLER, S. MOUELHI.*Safety Controller Synthesis for Incrementally Stable Switched Systems Using Multiscale Symbolic Models*, in "IEEE Transactions on Automatic Control", 2016, vol. 61, n^o^ 6, p. 1537-1549 [*DOI :* 10.1109/TAC.2015.2478131], https://hal.archives-ouvertes.fr/hal-01197426

[6] G. GÖSSLER, D. LE MÉTAYER.*A general framework for blaming in component-based systems*, in "Science of Computer Programming", 2015, vol. 113, Part 3 [*DOI :* 10.1016/J.SCICO.2015.06.010], https://hal.inria.fr/hal-01211484

[7] I. LANESE, C. A. MEZZINA, J.-B. STEFANI.*Reversibility in the higher-order π-calculus*, in "Theoretical Computer Science", 2016, vol. 625, p. 25-84 [*DOI :* 10.1016/J.TCS.2016.02.019], https://hal.inria.fr/hal-01303090

[8] S. QUINTON, M. HANKE, R. ERNST.*Formal analysis of sporadic overload in real-time systems*, in "2012 Design, Automation & Test in Europe Conference & Exhibition, DATE 2012, Dresden, Germany, March, 2012", 2012, p. 515–520, http://dx.doi.org/10.1109/DATE.2012.6176523

Publications of the year

Articles in International Peer-Reviewed Journal

[9] R. ERNST, S. KUNTZ, S. QUINTON, M. SIMONS.*The Logical Execution Time Paradigm: New Perspectives for Multicore Systems (Dagstuhl Seminar 18092)*, in "Dagstuhl Reports", 2018, vol. 8, p. 122 - 149 [*DOI :* 10.4230/DAGREP.8.2.122], https://hal.inria.fr/hal-01956964

[10] A. GIRAULT, C. PRÉVOT, S. QUINTON, R. HENIA, N. SORDON.*Improving and Estimating the Precision of Bounds on the Worst-Case Latency of Task Chains*, in "IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems", August 2018, vol. 37, n^o^ 11, p. 2578-2589 [*DOI :* 10.1109/TCAD.2018.2861016], https://hal.inria.fr/hal-01956931

Invited Conferences

[11] S. GRAF, S. QUINTON, A. GIRAULT, G. GÖSSLER.*Building Correct Cyber-Physical Systems: Why we need a Multiview Contract Theory*, in "FMICS 2018 - 23rd International Conference on Formal Methods for Industrial Critical Systems", Dublin, Ireland, LNCS, Springer, September 2018, vol. 11119, p. 19-31 [*DOI :* 10.1007/978-3-030-00244-2_2], https://hal.inria.fr/hal-01891146

[12] S. QUINTON.*Evaluation and Comparison of Real-Time Systems Analysis Methods and Tools*, in "FMICS 2018 - 23rd International Conference on Formal Methods for Industrial Critical Systems", Maynooth, Ireland, LNCS, Springer, September 2018, vol. 11119, p. 284-290 [*DOI :* 10.1007/978-3-030-00244-2_19], https://hal.inria.fr/hal-01903730

International Conferences with Proceedings

[13] L. Ahrendts, S. Quinton, T. Boroske, R. Ernst.*Verifying Weakly-Hard Real-Time Properties of Traffic Streams in Switched Networks*, in "ECRTS 2018 - 30th Euromicro Conference on Real-Time Systems", Barcelona, Spain, July 2018, p. 1-22 [*DOI :* 10.4230/LIPIcs.ECRTS.2018.15], https://hal.inria.fr/hal-01903759

[14] S. Cherrared, S. Imadali, E. Fabre, G. Gössler.*LUMEN: A Global Fault Management Framework For Network Virtualization Environments*, in "ICIN 2018 - 21st Conference on Innovation in Clouds, Internet and Networks and Workshops", Paris, France, IEEE, February 2018, p. 1-8 [*DOI :* 10.1109/ICIN.2018.8401622], https://hal.inria.fr/hal-01851610

[15] P. Fradet, A. Girault, R. Krishnaswamy, X. Nicollin, A. Shafiei.*RDF: Reconfigurable Dataflow*, in "2019 Design, Automation & Test in Europe Conference & Exhibition, DATE 2019", Florence, Italy, March 2019, https://hal.inria.fr/hal-01960788

[16] P. Fradet, X. Guo, J.-F. Monin, S. Quinton.*A Generalized Digraph Model for Expressing Dependencies*, in "RTNS '18 - 26th International Conference on Real-Time Networks and Systems", Chasseneuil-du-Poitou, France, October 2018, p. 1-11 [*DOI :* 10.1145/3273905.3273918], https://hal.inria.fr/hal-01878100

[17] P. Fradet, M. Lesourd, J.-F. Monin, S. Quinton.*A Generic Coq Proof of Typical Worst-Case Analysis*, in "RTSS 2018 - 39th IEEE Real-Time Systems Symposium", Nashville, United States, December 2018, p. 1-12, https://hal.inria.fr/hal-01903752

[18] A. Girault, G. Gössler, R. Guerraoui, J. Hamza, D.-A. Seredinschi.*Monotonic Prefix Consistency in Distributed Systems*, in "FORTE 2018 - 38th International Conference on Formal Techniques for Distributed Objects, Components, and Systems", Madrid, Spain, C. Baier, L. Caires (editors), Formal Techniques for Distributed Objects, Components, and Systems, Springer International Publishing, June 2018, vol. LNCS-10854, p. 41-57 [*DOI :* 10.1007/978-3-319-92612-4_3], https://hal.inria.fr/hal-01824817

[19] M. Vassor, J.-B. Stefani.*Checkpoint/rollback vs causally-consistent reversibility*, in "RC 2018 - 10th International Conference on Reversible Computation", Leicester, United Kingdom, Lecture Notes in Computer Science, Springer, September 2018, vol. 11106, p. 286-303 [*DOI :* 10.1007/978-3-319-99498-7_20], https://hal.inria.fr/hal-01953756

Research Reports

[20] A. Abdi, A. Girault, H. Zarandi.*ERPOT: A quad-criteria scheduling heuristic to optimize the execution time, failure rate, power consumption and temperature in multicores*, Inria ; 38, July 2018, n° RR-9196, p. 1-38, https://hal.inria.fr/hal-01848087

[21] R. Ernst, L. Ahrendts, K.-B. Gemlau, S. Quinton, H. Von Hasseln, J. Hennig.*System Level LET with Application to Automotive Design*, TU Braunschweig, 2018, p. 1-11, https://hal.inria.fr/hal-01962330

References in notes

[22] *Automotive Open System Architecture*, 2003, http://www.autosar.org

[23] *A Library for formally proven schedulability analysis*, http://prosa.mpi-sws.org/

[24] ARTEMIS JOINT UNDERTAKING.*ARTEMIS Strategic Research Agenda*, 2011

[25] S. ANDALAM, P. ROOP, A. GIRAULT.*Predictable Multithreading of Embedded Applications Using PRET-C*, in "International Conference on Formal Methods and Models for Codesign, MEMOCODE'10", Grenoble, France, IEEE, July 2010, p. 159–168

[26] I. ASSAYAD, A. GIRAULT.*Adaptive Mapping for Multiple Applications on Parallel Architectures*, in "Third International Symposium on Ubiquitous Networking, UNET'17", Casablanca, Morocco, May 2017, https://hal.inria.fr/hal-01672463

[27] P. AXER, R. ERNST, H. FALK, A. GIRAULT, D. GRUND, N. GUAN, B. JONSSON, P. MARWEDEL, J. REINEKE, C. ROCHANGE, M. SEBATIAN, R. VON HANXLEDEN, R. WILHELM, W. YI.*Building Timing Predictable Embedded Systems*, in "ACM Trans. Embedd. Comput. Syst.", 2014, To appear

[28] E. BAINOMUGISHA, A. CARRETON, T. VAN CUTSEM, S. MOSTINCKX, W. DE MEUTER.*A Survey on Reactive Programming*, in "ACM Computing Surveys", 2013, vol. 45, n^{o} 4

[29] N. BANSAL, T. KIMBREL, K. PRUHS.*Speed Scaling to Manage Energy and Temperature*, in "Journal of the ACM", 2007, vol. 54, n^{o} 1

[30] A. BASU, S. BENSALEM, M. BOZGA, J. COMBAZ, M. JABER, T.-H. NGUYEN, J. SIFAKIS.*Rigorous Component-Based System Design Using the BIP Framework*, in "IEEE Software", 2011, vol. 28, n^{o} 3

[31] V. BEBELIS, P. FRADET, A. GIRAULT, B. LAVIGUEUR.*BPDF: A Statically Analyzable Dataflow Model with Integer and Boolean Parameters*, in "International Conference on Embedded Software, EMSOFT'13", Montreal, Canada, ACM, September 2013

[32] A. BENVENISTE, P. CASPI, S. A. EDWARDS, N. HALBWACHS, P. LE GUERNIC, R. DE SIMONE.*The synchronous languages 12 years later*, in "Proceedings of the IEEE", 2003, vol. 91, n^{o} 1

[33] S. BORKAR.*Designing Reliable Systems from Unreliable Components: The Challenges of Transistor Variability and Degradation*, in "IEEE Micro", 2005, vol. 25, n^{o} 6

[34] A. BOUAKAZ, P. FRADET, A. GIRAULT.*A Survey of Parametric Dataflow Models of Computation*, in "ACM Transactions on Design Automation of Electronic Systems (TODAES)", January 2017, https://hal.inria.fr/hal-01417126

[35] A. BOUAKAZ, P. FRADET, A. GIRAULT.*Symbolic Analyses of Dataflow Graphs*, in "ACM Transactions on Design Automation of Electronic Systems (TODAES)", January 2017, https://hal.inria.fr/hal-01417146

[36] R. DAVIS, A. BURNS.*A Survey of Hard Real-Time Scheduling for Multiprocessor Systems*, in "ACM Computing Surveys", 2011, vol. 43, n^{o} 4

[37] S. A. EDWARDS, E. A. LEE.*The Case for the Precision Timed (PRET) Machine*, in "44th Design Automation Conference (DAC)", IEEE, 2007

[38] J. EKER, J. W. JANNECK, E. A. LEE, J. LIU, X. LIU, J. LUDVIG, S. NEUENDORFFER, S. SACHS, Y. XIONG.*Taming heterogeneity - the Ptolemy approach*, in "Proceedings of the IEEE", 2003, vol. 91, n^{o} 1

[39] P. FRADET, A. GIRAULT, L. JAMSHIDIAN, X. NICOLLIN, A. SHAFIEI.*Lossy channels in a dataflow model of computation*, in "Principles of Modeling, Festschrift in Honor of Edward A. Lee", Berkeley, United States, Lecture Notes in Computer Science, Springer, October 2017, https://hal.inria.fr/hal-01666568

[40] P. FRADET, A. GIRAULT, P. POLPAVKO.*SPDF: A schedulable parametric data-flow MoC*, in "Design, Automation and Test in Europe, DATE'12", IEEE, 2012

[41] B. GAUJAL, A. GIRAULT, S. PLASSART.*Dynamic Speed Scaling Minimizing Expected Energy Consumption for Real-Time Tasks*, UGA - Université Grenoble Alpes ; Inria Grenoble Rhône-Alpes ; Université de Grenoble, October 2017, n^{o} RR-9101, p. 1-35, https://hal.inria.fr/hal-01615835

[42] A. GIRARD, G. PAPPAS.*Approximation metrics for discrete and continuous systems*, in "IEEE Trans. on Automatic Control", 2007, vol. 52, n^{o} 5, p. 782–798

[43] D. GIZOPOULOS, M. PSARAKIS, S. V. ADVE, P. RAMACHANDRAN, S. K. S. HARI, D. SORIN, A. MEIXNER, A. BISWAS, X. VERA.*Architectures for Online Error Detection and Recovery in Multicore Processors*, in "Design Automation and Test in Europe (DATE)", 2011

[44] R. GUERRAOUI, M. PAVLOVIC, D.-A. SEREDINSCHI.*Trade-offs in replicated systems*, in "IEEE Data Engineering Bulletin", 2016, vol. 39, p. 14–26

[45] F. C. GÄRTNER.*Fundamentals of Fault-Tolerant Distributed Computing in Asynchronous Environments*, in "ACM Computing Surveys", 1999, vol. 31, n^{o} 1

[46] G. GÖSSLER, J.-B. STEFANI.*Fault Ascription in Concurrent Systems*, in "Proc. Trustworthy Global Computing - 10th International Symposium, TGC 2015", P. GANTY, M. LORETI (editors), LNCS, Springer, 2016, vol. 9533

[47] S. HAAR, E. FABRE.*Diagnosis with Petri Net Unfoldings*, in "Control of Discrete-Event Systems", Lecture Notes in Control and Information Sciences, Springer, 2013, vol. 433, chap. 15

[48] J. HALPERN, J. PEARL.*Causes and Explanations: A Structural-Model Approach. Part I: Causes*, in "British Journal for the Philosophy of Science", 2005, vol. 56, n^{o} 4, p. 843-887

[49] T. HENZINGER, J. SIFAKIS.*The Embedded Systems Design Challenge*, in "Formal Methods 2006", Lecture Notes in Computer Science, Springer, 2006, vol. 4085

[50] C. M. KIRSCH, A. SOKOLOVA.*The Logical Execution Time Paradigm*, in "Advances in Real-Time Systems (to Georg Färber on the occasion of his appointment as Professor Emeritus at TU München after leading the Lehrstuhl für Realzeit-Computersysteme for 34 illustrious years)", 2012, p. 103–120

[51] R. KÜSTERS, T. TRUDERUNG, A. VOGT.*Accountability: definition and relationship to verifiability*, in "ACM Conference on Computer and Communications Security", 2010, p. 526-535

[52] I. LANESE, C. A. MEZZINA, J.-B. STEFANI.*Reversing Higher-Order Pi*, in "21th International Conference on Concurrency Theory (CONCUR)", Lecture Notes in Computer Science, Springer, 2010, vol. 6269

[53] P. MENZIES.*Counterfactual Theories of Causation*, in "Stanford Encyclopedia of Philosophy", E. ZALTA (editor), Stanford University, 2009, http://plato.stanford.edu/entries/causation-counterfactual

[54] M. MOORE.*Causation and Responsibility*, Oxford, 1999

[55] J. PEARL.*Causal inference in statistics: An overview*, in "Statistics Surveys", 2009, vol. 3, p. 96-146

[56] P. RAMADGE, W. WONHAM.*Supervisory Control of a Class of Discrete Event Processes*, in "SIAM Journal on control and optimization", January 1987, vol. 25, n^o^ 1, p. 206–230

[57] J. RUSHBY.*Partitioning for Safety and Security: Requirements, Mechanisms, and Assurance*, NASA Langley Research Center, 1999, n^o^ CR-1999-209347

[58] J.-B. STEFANI.*Components as Location Graphs*, in "11th International Symposium on Formal Aspects of Component Software", Bertinoro, Italy, Lecture Notes in Computer Science, September 2014, vol. 8997, https://hal.inria.fr/hal-01094208

[59] M. STIGGE, P. EKBERG, N. GUAN, W. YI.*The Digraph Real-Time Task Model*, in "17th IEEE Real-Time and Embedded Technology and Applications Symposium, RTAS 2011, Chicago, Illinois, USA, 11-14 April 2011", 2011, p. 71–80, https://doi.org/10.1109/RTAS.2011.15

[60] P. TABUADA.*Verification and Control of Hybrid Systems - A Symbolic Approach*, Springer, 2009

[61] D. TERRY.*Replicated Data Consistency Explained Through Baseball*, Microsoft Research, October 2011, n^o^ MSR-TR-2011-137

[62] K. TINDELL.*Using offset information to analyse static priority pre-emptively scheduled task sets*, Technical report YCS 182, University of York, Department of Computer Science, 1992, https://books.google.fr/books?id=qARQHAAACAAJ

[63] J. WANG, P. S. ROOP, A. GIRAULT.*Energy and timing aware synchronous programming*, in "International Conference on Embedded Software, EMSOFT'16", Pittsburgh, United States, ACM, October 2016, 10 [*DOI :* 10.1145/2968478.2968500], https://hal.inria.fr/hal-01412100

[64] R. WILHELM, J. ENGBLOM, A. ERMEDAHL, N. HOLSTI, S. THESING, D. B. WHALLEY, G. BERNAT, C. FERDINAND, R. HECKMANN, T. MITRA, F. MUELLER, I. PUAUT, P. P. PUSCHNER, J. STASCHULAT, P. STENSTRÖM.*The Determination of Worst-Case Execution Times — Overview of the Methods and Survey of Tools*, in "ACM Trans. Embedd. Comput. Syst.", April 2008, vol. 7, n^o^ 3

[65] F. YAO, A. DEMERS, S. SHENKER.*A scheduling model for reduced CPU energy*, in "Proceedings of IEEE Annual Foundations of Computer Science", 1995, p. 374–382

Project-Team STEEP

Sustainability transition, environment, economy and local policy

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

IN PARTNERSHIP WITH:
CNRS

Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Earth, Environmental and Energy Sciences

Table of contents

Project-Team STEEP

Creation of the Team: 2010 January 01, updated into Project-Team: 2015 December 01

Keywords:

Computer Science and Digital Science:

A3.3.2. - Data mining
A6.1. - Methods in mathematical modeling
A9.6. - Decision support

Other Research Topics and Application Domains:

B3.1. - Sustainable development
B3.1.1. - Resource management
B3.4.3. - Pollution
B7. - Transport and logistics
B8.3. - Urbanism and urban planning
B8.5.1. - Participative democracy
B8.5.3. - Collaborative economy
B9.9. - Ethics

1. Team, Visitors, External Collaborators

Research Scientists

Emmanuel Prados [Team leader, Inria, Researcher, HDR]
Pierre-Yves Longaretti [CNRS, Researcher]
Peter Sturm [Inria, Senior Researcher, HDR]

Faculty Members

Denis Dupré [Univ Grenoble Alpes, Associate Professor]
Serge Fenet [Univ de Claude Bernard, Associate Professor]
Régis Perrier [Univ Grenoble Alpes, Associate Professor]
Guillaume Mandil [Univ Grenoble Alpes, Associate Professor, from Sep 2018]

Technical Staff

Christophe Boisseau [Inria]

PhD Students

Michela Bevione [Inria]
Luciano Gervasoni [Univ Grenoble Alpes, until Nov 2018]

Post-Doctoral Fellows

Thomas Capelle [Inria, until May 2018]
Jean-Yves Courtonne [Inria]
Fausto Lo Feudo [Univ Grenoble Alpes, from Mar 2018]

Administrative Assistant

Marie-Anne Dauphin-Rizzi [Inria]

2. Overall Objectives

2.1. Overview

STEEP started in January 2010, initially as an Inria “Action Exploratoire” (2010+2011). It is now an “Équipe Projet Inria” of Inria Grenoble - Rhône-Alpes and is also affiliated with the Jean Kuntzmann laboratory (LJK [0]).

STEEP is an interdisciplinary research team devoted to systemic modelling and simulation of the interactions between the environmental, economic and social factors in the context of a transition to sustainability at local (sub-national) scales. Our goal is to develop decision-making tools to support decision makers in the implementation of this transition by developing simulation and optimization programs. In other words, our objective is to set up some mathematical and computational tools which enable us to provide some parts of the answer to the challenges *how to operate the sustainable development at local scales? and which local governance for environmental public policies?*.

The work of STEEP follows several research directions, covering different application domains; these are described in “Scientific Foundations” and “Application Domains” respectively.

2.2. Sustainable development: issues and research opportunities

Environmental issues now pose a threat to human civilization worldwide. They range from falling water tables to eroding soils, expanding deserts, biodiversity loss, rising temperatures, *etc*. For example, half the world’s population lives in countries where water tables are falling as aquifers are being depleted. Roughly a third of the world’s cropland is losing topsoil at an excessive rate. Glaciers are melting in all of the world’s major mountains. The consequences on the present human societies are critical; they comprise for example a decreasing food security, important population movements (such as climate refugees) and explosive geopolitical tensions.

Sustainable development is often formulated in terms of a required balance between its environmental, economic and social dimensions, but in practice public policies addressing sustainability issues are dominantly oriented towards environment management in Western countries. This approach is problematic to some extent as environmental problems and sustainability issues result from socio-economic phenomena (for example the economic growth model which is strengthened by powerful and polluting technologies). Environmental problems have only recently been the object of media attention and public awareness. Most efforts bear on developing technological solutions. However, it is now clear that this will not be sufficient. We need to rethink our socio-economic and institutional models in order to leave room for a possible paradigm shift. In this perspective, we believe that crucial steps should be taken in research to help elaborating and implementing socio-economic alternatives.

The risks associated with delayed reaction and adaptation times make the situation urgent. Delayed reactions significantly increase the probability of overshoot of the planet carrying capacity followed by uncontrolled and irreversible evolution on a number of fronts . This systemic problem is amplified by two facts: the environment is degrading on all fronts at the same time, and at the global planetary scale, a first in human history.

Although environmental challenges are monitored worldwide, the search for appropriate lines of actions must nevertheless take place at all institutional levels, in particular at local scales.At such scales, the proximity and smaller number of stakeholders allows decision makers to reach a consensus much more easily than at national or international scales. The failure of the recent Copenhagen summit (and for that matter of all climate summits since the adoption of the Kyoto protocol in 1997) is a good illustration of the difficulties encountered in international negotiations. There are significant possibilities for operations at local scales, and the emergency of the situation gives the “think locally to act globally” logic an essential opportunity.

[0] http://ljk.imag.fr/

As of now, local decision levels have real political and economic leverage, and are more and more proactive on sustainability issues, either independently or in coordination through nationwide or European networks (we can refer for example to the European GMO-free Regions Network [0] or to the Network of European Regions for a Competitive and Sustainable TouRism [0]). Also, we think that two local scales are going to be increasingly dominant in the near future: urban areas (more exactly the employment areas of main cities) and "regions" (such as régions in France, Lander in Germany or Cantons in Switzerland). In particular, the sustainability of urban areas is one of the key issues of this century. As focal points of human activity, urban areas concentrate and amplify environmental pressures in a direct or indirect way.

Urbanization is a global and an ever-increasing trend process, with more than half the human population living in cities. Although urbanized areas still represent a very small fraction of the total terrestrial surface, urban resource consumption amounts to three-fourths of the annual total in energy, water, building materials, agricultural products etc., and pollution and waste management is a growing concern for urban planners worldwide. In France, for example, even if resource intensity (materials use divided by GDP [0]) has been reduced by half since the 70s, the actual material use (total and per inhabitant) has remained essentially constant, and household wastes have grown by 20% since 1995. Greenhouse gas (GHG) emissions have been reduced by a few percent since 1990, but the transportation share (a major issue on this front) has been steadily growing over the same period.

Furthermore, urban sprawl is a ubiquitous phenomenon showing no sign of slackening yet, even in countries where rural depopulation has long been stabilized. Urban sprawl in industrialized countries is largely driven by residential suburban growth. This phenomenon has both social and environmental consequences. First it implies an increase of daily mobility. In a context of high dependency on private cars and uncertainty on energy prices, this translates into an increased vulnerability of some population categories. It also induces an increase in greenhouse gas emissions, as well as an irreversible loss of cropland and a fragmentation of ecological habitat, with negative effects on biodiversity. The increasing concerns about climate change and upheaval in the market price of fossil fuels raise many questions about urban energy consumption while reviving the debate on the desirable urban structures and their determinants. Controlling urban sprawl is therefore a key sustainability issue.

Let us mention here that cities cannot be sustainable by themselves and that from this point of view, it does not make sense to focus on the municipality scale ("*communes*"). We think that it is very important to work at larger scales, typically, at employment catchment areas complemented by the adjacent agricultural and natural zones they are dependent on (that would correspond to the smallest scale for which a systemic analysis could make sense). Nevertheless, let us emphasize that because of resource imports and waste exports (e.g. GHG emissions), for any limited territory, the considered area will always depend on and impact other more or less distant territories. This is one of the key issues when trying to assess local sustainability.

Finally, let us note that the numerous and interrelated pressures exerted by human activities on the environment make the identification of sustainable development pathways arduous in a context of complex and sometimes conflicting stakeholders and socio-ecological interactions. This is why we also think that it is crucial to develop interdisciplinary and integrated approaches; consequently, our proposal tries to address the entire spectrum from scientific expertise to stakeholder decision-help.

STEEP, with its strong background in various areas of applied mathematics and modeling, can be a game changer in three connected key domains: urban economy, and related transportation and land use issues; material flow analysis and ecological accounting; and ecosystem services modeling. The group potential on these fronts relies on its capabilities to strongly improve existing integrated activity / land use / transportation models at the urban level on the one hand, and on the other, to build new and comprehensive decision-help tools for sustainability policies at the local and regional levels, in particular through the analysis of strategic social–environmental trade-offs between various policy options.

[0] http://www.gmo-free-regions.org

[0] http://www.necstour.eu

[0] Gross Domestic Product (GDP) is defined as an aggregate measure of production equal to the sum of the gross values added of all resident institutional units engaged in production.

3. Research Program

3.1. Development of numerical systemic models (economy / society /environment) at local scales

The problem we consider is intrinsically interdisciplinary: it draws on social sciences, ecology or science of the planet. The modeling of the considered phenomena must take into account many factors of different nature which interact with varied functional relationships. These heterogeneous dynamics are *a priori* nonlinear and complex: they may have saturation mechanisms, threshold effects, and may be density dependent. The difficulties are compounded by the strong interconnections of the system (presence of important feedback loops) and multi-scale spatial interactions. Environmental and social phenomena are indeed constrained by the geometry of the area in which they occur. Climate and urbanization are typical examples. These spatial processes involve proximity relationships and neighborhoods, like for example, between two adjacent parcels of land, or between several macroscopic levels of a social organization. The multi-scale issues are due to the simultaneous consideration in the modeling of actors of different types and that operate at specific scales (spatial and temporal). For example, to properly address biodiversity issues, the scale at which we must consider the evolution of rurality is probably very different from the one at which we model the biological phenomena.

In this context, to develop flexible integrated systemic models (upgradable, modular, ...) which are efficient, realistic and easy to use (for developers, modelers and end users) is a challenge in itself. What mathematical representations and what computational tools to use? Nowadays many tools are used: for example, cellular automata (e.g. in the LEAM model), agent models (e.g. URBANSIM [0]), system dynamics (e.g. World3), large systems of ordinary equations (e.g. equilibrium models such as TRANUS), and so on. Each of these tools has strengths and weaknesses. Is it necessary to invent other representations? What is the relevant level of modularity? How to get very modular models while keeping them very coherent and easy to calibrate? Is it preferable to use the same modeling tools for the whole system, or can we freely change the representation for each considered subsystem? How to easily and effectively manage different scales? (difficulty appearing in particular during the calibration process). How to get models which automatically adapt to the granularity of the data and which are always numerically stable? (this has also a direct link with the calibration processes and the propagation of uncertainties). How to develop models that can be calibrated with reasonable efforts, consistent with the (human and material) resources of the agencies and consulting firms that use them?

Before describing our research axes, we provide a brief overview of the types of models that we are or will be working with. As for LUTI (Land Use and Transportation Integrated) modeling, we have been using the TRANUS model since the start of our group. It is the most widely used LUTI model, has been developed since 1982 by the company Modelistica, and is distributed *via* Open Source software. TRANUS proceeds by solving a system of deterministic nonlinear equations and inequalities containing a number of economic parameters (e.g. demand elasticity parameters, location dispersion parameters, etc.). The solution of such a system represents an economic equilibrium between supply and demand.

On the other hand, the scientific domains related to ecosystem services and ecological accounting are much less mature than the one of urban economy from a modelling point of view (as a consequence of our more limited knowledge of the relevant complex processes and/or more limited available data). Nowadays, the community working on ecological accounting develops statistical models based on the enforcement of the mass conservation constraint for accounting for material fluxes through a territorial unit or a supply chain, relying on more or less simple data correlations when the relevant data is missing; the overall modelling makes heavy use of more or less sophisticated linear algebra and constrained optimization techniques. The ecosystem service community has been using statical models too, but is also developing more sophisticated models based for example on system dynamics, multi-agent type simulations or cellular models. In the ESNET project, STEEP has worked in particular on a land use/ land cover change (LUCC) modelling environments (Dinamica [0]) which belongs to the category of spatially explicit statistical models.

[0] http://www.urbansim.org

In the following, our two main research axes are described, from the point of view of applied mathematical development. The domains of application of this research effort is described in the application section, where some details about the context of each field is given.

3.2. Model calibration and validation

The overall calibration of the parameters that drive the equations implemented in the above models is a vital step. Theoretically, as the implemented equations describe e.g. socio-economic phenomena, some of these parameters should in principle be accurately estimated from past data using econometrics and statistical methods like regressions or maximum likelihood estimates, e.g. for the parameters of logit models describing the residential choices of households. However, this theoretical consideration is often not efficient in practice for at least two main reasons. First, the above models consist of several interacting modules. Currently, these modules are typically calibrated independently; this is clearly sub-optimal as results will differ from those obtained after a global calibration of the interaction system, which is the actual final objective of a calibration procedure. Second, the lack of data is an inherent problem.

As a consequence, models are usually calibrated by hand. The calibration can typically take up to 6 months for a medium size LUTI model (about 100 geographic zones, about 10 sectors including economic sectors, population and employment categories). This clearly emphasizes the need to further investigate and at least semi-automate the calibration process. Yet, in all domains STEEP considers, very few studies have addressed this central issue, not to mention calibration under uncertainty which has largely been ignored (with the exception of a few uncertainty propagation analyses reported in the literature).

Besides uncertainty analysis, another main aspect of calibration is numerical optimization. The general state-of-the-art on optimization procedures is extremely large and mature, covering many different types of optimization problems, in terms of size (number of parameters and data) and type of cost function(s) and constraints. Depending on the characteristics of the considered models in terms of dimension, data availability and quality, deterministic or stochastic methods will be implemented. For the former, due to the presence of non-differentiability, it is likely, depending on their severity, that derivative free control methods will have to be preferred. For the latter, particle-based filtering techniques and/or metamodel-based optimization techniques (also called response surfaces or surrogate models) are good candidates.

These methods will be validated, by performing a series of tests to verify that the optimization algorithms are efficient in the sense that 1) they converge after an acceptable computing time, 2) they are robust and 3) that the algorithms do what they are actually meant to. For the latter, the procedure for this algorithmic validation phase will be to measure the quality of the results obtained after the calibration, i.e. we have to analyze if the calibrated model fits sufficiently well the data according to predetermined criteria.

To summarize, the overall goal of this research axis is to address two major issues related to calibration and validation of models: (a) defining a calibration methodology and developing relevant and efficient algorithms to facilitate the parameter estimation of considered models; (b) defining a validation methodology and developing the related algorithms (this is complemented by sensitivity analysis, see the following section). In both cases, analyzing the uncertainty that may arise either from the data or the underlying equations, and quantifying how these uncertainties propagate in the model, are of major importance. We will work on all those issues for the models of all the applied domains covered by STEEP.

3.3. Sensitivity analysis

A sensitivity analysis (SA) consists, in a nutshell, in studying how the uncertainty in the output of a model can be apportioned to different sources of uncertainty in the model inputs. It is complementary to an uncertainty analysis, which focuses on quantifying uncertainty in model output. SA's can be useful for several purposes, such as guiding model development and identifying the most influential model parameters and critical data items. Identifying influential model parameters may help in divising metamodels (or, surrogate models) that approximate an original model and may be simulated, calibrated, or analyzed more efficiently. As for detecting

[0] http://www.csr.ufmg.br/dinamica/

critical data items, this may indicate for which type of data more effort must be spent in the data collection process in order to eventually improve the model's reliability. Finally, SA can be used as one means for validating models, together with validation based on historical data (or, put simply, using training and test data) and validation of model parameters and outputs by experts in the respective application area.

The first two applications of SA are linked to model calibration, discussed in the previous section. Indeed, prior to the development of the calibration tools, one important step is to select the significant or sensitive parameters and to evaluate the robustness of the calibration results with respect to data noise (stability studies). This may be performed through a global sensitivity analysis, e.g. by computation of Sobol's indices. Many problems had been to be circumvented e.g. difficulties arising from dependencies of input variables, variables that obey a spatial organization, or switch inputs. We take up on current work in the statistics community on SA for these difficult cases.

As for the third application of SA, model validation, a preliminary task bears on the propagation of uncertainties. Identifying the sources of uncertainties and their nature is crucial to propagate them via Monte Carlo techniques. To make a Monte Carlo approach computationally feasible, it is necessary to develop specific metamodels. Both the identification of the uncertainties and their propagation require a detailed knowledge of the data collection process; these are mandatory steps before a validation procedure based on SA can be implemented. First, we focus on validating LUTI models, starting with the CITiES ANR project: here, an SA consists in defining various land use policies and transportation scenarios and in using these scenarios to test the integrated land use and transportation model. Current approaches for validation by SA consider several scenarios and propose various indicators to measure the simulated changes. We work towards using sensitivity indices based on functional analysis of variance, which allow us to compare the influence of various inputs on the indicators. For example it allow the comparison of the influences of transportation and land use policies on several indicators.

4. Highlights of the Year

4.1. Highlights of the Year

The STEEP research team has initiated in 2016 a series of conferences-debates entitled "Understanding & Acting" (« Comprendre et agir ») that examines sustainability issues in order to help researchers and citizens to increase their awareness of the various issues at stake in order to initiate relevant individual and collective actions. The presentations are captured on video and then made directly accessible on the YouTube Channel "Comprendre et Agir". At the end of 2018 the YouTube channel reached more than **150,000 views with a rate of integral viewings remaining at above 25%.** This rate is quite important since the YouTube videos of the conferences last between 35 and 45 minutes. Our Youtube channel now has more than 2000 subscribers.

5. New Software and Platforms

5.1. Software tools for the TRANUS LUTI Model

KEYWORDS: Urban planning - Transport model - LUTI

FUNCTIONAL DESCRIPTION: This year, we have consolidated and extended our software tools for the TRANUS LUTI model, thanks to support by Inria allowing to hire an engineer for one year Emna Jribi (ADT TRACAV project). Various tasks have been accomplished, concerning three types of functionality, these are as follows. First, calibration of TRANUS. The software implementation of our methods for calibrating the TRANUS land-use component has been cleaned up. It has been encapsulated such as to be seamlessly integrated within the TRANUS workflow (consisting of a sequence of executables, exchanging data through binary and other files). Second, graphical user interfaces to facilitate the repeated execution of TRANUS executables, for generating reports on results or for the exploration of the space of some critical model parameters. Third, we have continued to work on the embedding of TRANUS within the open source QGIS platform (a widely used Geographic Information System).

- Participants: Emna Jribi, Thomas Capelle and Peter Sturm
- Contact: Peter Sturm
- URL: https://gitlab.inria.fr/tranus

5.2. USAT

Urban Sprawl Analysis Toolkit
KEYWORDS: Urban sprawl - Urban planning
FUNCTIONAL DESCRIPTION: This software allows to calculate and analyse indices of urban sprawl from open data (OpenStreetMap), aimed to be used by urban scientists and urban planners. A spatialized version of indices measuring the accessibility, dispersion and land use mix is calculated. The implemented methods are described in [9].

- Participants: Luciano Gervasoni, Serge Fenet and Peter Sturm
- Partner: LIRIS
- Contact: Peter Sturm
- URL: https://github.com/lgervasoni/urbansprawl

5.3. USAT WEB

Urban Sprawl Analysis Toolkit Web-service
KEYWORDS: Urban planning - Urban sprawl
FUNCTIONAL DESCRIPTION: This is a web-service on top of the software USAT described above. The web-service will allow any user to select a region of interest and to launch the calculation and display of sprawl indices using USAT. It is in the process of being hosted on the HPC platform of IN2P3, after which it will be made open to the public. The source code for this web-service is already available at the below site.

The web-service is described in [10].

- Participants: Lucas Rezakhanlou, Peter Sturm, Luciano Gervasoni and Serge Fenet
- Contact: Peter Sturm
- Publication: USAT (Urban Sprawl Analysis Toolkit) : une plateforme web d'analyse de l'étalement urbain à partir de données massives ouvertes
- URL: https://gitlab.inria.fr/lrezakha/usat-web

5.4. InterfacesTRANUS

FUNCTIONAL DESCRIPTION: This software contains two interfaces dedicated to facilitating the usage of the TRANUS integrated land use and transport model+software. The first interface is dedicated to enabling the execution of the TRANUS binary programs without the need to use the console or the TRANUS GUI. The second interface provides an aid for calibrating a TRANUS model, by interactively exploring ranges of different parameters of a TRANUS model and visualising model outputs across these ranges.

- Participants: Julien Armand, Peter Sturm and Thomas Capelle
- Contact: Peter Sturm
- URL: https://gitlab.inria.fr/tranus/TRANUS_Interfaces

5.5. LUM_OSM

Land Use Mix calculation from OpenStreepMap data

FUNCTIONAL DESCRIPTION: The software uses Mapzen Metro Extracts to retrieve the OpenStreetMap data of a given region in the PostgreSQL format. Afterwards, a continuous representation of residential and activity land uses is created. Finally, a GIS output containing the degree of land use mixture is calculated by means of using the land uses maps. The implemented approach is documented in the paper "A framework for evaluating urban land use mix from crowd-sourcing data", http://hal.inria.fr/hal-01396792

- Participants: Luciano Gervasoni, Marti Bosch Padros, Peter Sturm and Serge Fenet
- Partners: EPFL - Ecole Polytechnique Fédérale de Lausanne - LIRIS
- Contact: Peter Sturm
- URL: http://github.com/martibosch/landusemix

5.6. QGIS_Tranus_Reports

FUNCTIONAL DESCRIPTION: This software allows to graphically visualise data output by the TRANUS LUTI model (and possibly, of any other data of the same structure). In particular, this concerns any data items defined per zone of a modelled territory (productions, indicators, etc.). The software is designed as a plugin for the geographical information system platform QGIS and can be run interactively as well as by the command line or by a call from within another software. The interactive mode (within QGIS) allows the user to define graphical outputs to be generated from TRANUS output files (type of graphs to be generated – 2D or 3D – color coding to be used, choice of data to be displayed, etc.). Visualisation of data is done in the form of 2D graphs or 3D models defined using java-script.

- Participants: Fausto Lo Feudo, Huu Phuoc Nguyen, Patricio Inzaghi, Peter Sturm and Thomas Capelle
- Contact: Peter Sturm
- URL: https://gitlab.inria.fr/tranus/QGIS_Tranus_Reports

5.7. Comptabilité Ecologique

FUNCTIONAL DESCRIPTION: Databases, database handling tools and data visualization tools (on the website). Databases include socio-economic and environmental datasets. Visualization tools include interactive piecharts, maps and Sankey diagrams.

- Participants: Jean-Yves Courtonne and Pierre-Yves Longaretti
- Contact: Jean-Yves Courtonne
- URL: http://www.eco-data.fr

6. New Results

6.1. Calibration of the Tranus Land Use Module: Optimisation-Based Algorithms, their Validation, and Parameter Selection by Statistical Model Selection

Instantiating land use and transport integrated models (LUTI modelling) is a complicated task, requiring substantial data collection, parameter estimation and expert analysis. In this work, we present a partial effort towards the automation of the calibration of Tranus, one of the most popular LUTI models. First, we give a detailed mathematical description of the activity module and the usual calibration approach. Secondly, we reformulate the estimation of the endogenous parameters called shadow prices as an optimisation problem. We also propose an optimisation algorithm for the calibration of the substitution submodel, setting a base for future fully integrated calibration. We analyse the case of transportable and non-transportable economic sectors and propose a detailed mathematical scheme for each case. We also discuss how to validate calibration results and propose to use synthetic data generated from real world problems in order to assess convergence properties and accuracy of calibration methods. Results of this methodology are presented for realistic scenarios. Finally, we propose a model selection scheme to reduce the number of shadow prices that need to be calibrated, with the aim of reducing the risk of overfitting to data. This work is published in [2].

6.2. Convolutional neural networks for disaggregated population mapping using open data

High resolution population count data are vital for numerous applications such as urban planning, transportation model calibration, and population growth impact measurements, among others. In this work, we present and evaluate an end-to-end framework for computing disaggregated population mapping employing convolutional neural networks (CNNs). Using urban data extracted from the OpenStreetMap database, a set of urban features are generated which are used to guide population density estimates at a higher resolution. A population density grid at a 200 by 200 meter spatial resolution is estimated, using as input gridded population data of 1 by 1 kilometer. Our approach relies solely on open data with a wide geographical coverage, ensuring replicability and potential applicability to a great number of cities in the world. Fine-grained gridded population data is used for 15 French cities in order to train and validate our model. A stand-alone city is kept out for the validation procedure. The results demonstrate that the neural network approach using massive OpenStreetMap data outperforms other approaches proposed in related works. This work is published in [5].

6.3. Uncertainties of Domestic Road Freight Statistics: Insights for Regional Material Flow Studies

Freight statistics are at the core of many studies in the field of industrial ecology because they depict the physical inter-dependencies of territories and allow to link worldwide productions and consumptions. Recent studies have been increasingly focusing on subnational scales, often relying on domestic freight data. In this perspective, this article analyses the uncertainties of the French domestic road freight survey, road being by far the most common mode of transport in the country. Based on a statistical analysis of the survey, we propose a model to estimate the uncertainty of any given domestic road transport flow. We also assess uncertainty reduction when averaging the flows over several years, and obtain for instance a 30% reduction for a 3-year average. We then study the impact of the uncertainties on regional material flow studies such as the Economy-Wide Material Flow Analysis of the Bourgogne region. Overall the case studies advocate for a systematic assessment of freight uncertainties, as neither the disaggregation level nor the quantities traded are good enough predictors. This justifies the need for an easy-to-implement estimation model. Finally, basic comparison with the German and Swedish surveys tend to indicate that the main conclusions presented in this article are likely to be valid in other European countries. This work is published in [3].

6.4. A method for downscaling open population data

To extend our ongoing work on urban sprawl indicators (see above), we have developed a method to perform dissagregated population estimations at building level using open data. Our goal is to estimate the number of people living at the fine level of individual households by using open urban data and coarse-scaled population data. First, a fine scale description of residential land use per building is built using OpenStreetMap. Then, using coarse-scale gridded population data, we perform the down-scaling for each household given their containing area for residential usage. We rely solely on open data in order to ensure replicability, and to be able to apply our method to any city in the world, as long as sufficient data exists. The evaluation is carried out using fine-grained census block data for cities in France as ground-truth.

This work is published in [6] and the associated software implementation is made available as open source code at https://github.com/lgervasoni/urbansprawl.

6.5. Modelling the relationships between urban morphology, pollutant generation and concentration in the air using PLS path modelling

We have simultaneously modelled the factors that contribute to shaping the urban environment in terms of population density and activities and the level of land use mix on the one hand, and the mechanisms through which this urban morphology is linked to the emission of pollutants and their concentration in the air in the municipalities of the Auvergne-Rhône-Alpes region. To do this, we used the PLS path modelling approach, which is a method of estimating structural equations to model the relationships between latent variables obtained by extracting the information contained in the multidimensional data used to measure them. This work was carried out as part of Diop Samba's internship [8].

6.6. Implementation of the World3 model in Python for parametric exploration

The 'World3' model is a digital tool for simulating long-term interactions between population, industrial growth, food production and the boundaries of terrestrial ecosystems. This model was developed in the 1970s. We have ported this model to a modern infrastructure (Python3 + related libraries), in order to be able to apply parameter learning, data analysis and sensitivity study techniques to it. This work was carried out as part of the internship of Aina Rasoldier.

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. QAMECS / MOBIL'AIR : ATMOSPHERIC POLLUTION: Characterization of novel exposure markers, of biological, health, economic and societal impacts and evaluation of public policies

Project funded by ADEME, Grenoble metropolis, IDEX Université Grenoble Alpes
Duration: 2016 – 2022
Project coordinator : Remy Slama (INSERM) and Sandrine Mathy (GAEL, CNRS). Inria Coordinator: Emmanuel Prados
Other partners: Air Rhône-Alpes, CNRS, Sciences Po Grenoble, Inserm, IAB, Université Grenoble-Alpes
Abstract: Urban atmospheric pollution is one of the main threats to human health that can be to some extent controlled by public action. In Europe, many cities have implemented various types of low emission zones (LEZ, focused on traffic and heating emissions), France being a notable exception. Although fine particulate matter (PM2.5) is usually assessed through its mass concentration, other metrics, such as PM chemical speciation as well as the so far little considered oxidative potential (OP) of PM, are worth considering, both in terms of associations with human health and in the context of monitoring of the efficiency of LEZ. QAMECS covers all dimensions from atmospheric emissions, impact of meteorological conditions on air pollution human behaviours related to transportation, environmental levels, health, associated economic costs and societal awareness. The project relies on environmental measurements, modelling, repeated observational (representative) population studies, an existing mother-child cohort, a controlled human experiment, health impact and related economic assessment. It is conducted by a consortium of specialists of chemistry and physics of air pollution, economics, sociology, epidemiology, geography, in relation with local authorities. It will bring results important for urban planning, public health, and more fundamental research on the measurement of PM and assessment of their biological and health impact.

7.2. National Initiatives

7.2.1. AF Filières : Analyse des Flux des Filières biomasse pour des stratégies régionales de bioéconomie

Project funded by ADEME
Duration: 2017-2019
Coordinator: Jean-Yves COURTONNE (Equipe STEEP, Inria) [Emmanuel Prados (STEEP/Inria) for Inria partner]

Other partners: Equipe STEEP, Inria, Grenoble Rhônalpénergie-Environnement (RAEE), Lyon Laboratoire d'Economie Forestière (LEF), INRA / AgroParisTech Nancy.
Keywords: Environmental assessment, Ecological accounting, Material Flow Analysis, Sustainable supply chains, Multicriteria analysis.
Abstract: Flow analyses of biomass supply chains for regional bioeconomy policies. The goals of the project are the following:

- Improve knowledge on the material flows of the forest-wood and agri-food supply chains in France at national and regional levels,
- Provide a holistic vision of the situation by associating environmental and socio-economic indicators to material flows,
- Provide a more precise assessments (quantitatively and qualitatively) in the case of the Auvergne-Rhône-Alpes region.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. General Chair, Scientific Chair

- Denis Dupré, Pierre-Yves Longaretti and Emmanuel Prados are the Scientific Chair of the series of conferences entitled "Understanding and Acting"; see section 8.3.1.

8.1.1.2. Member of the Organizing Committees

- P. Sturm was member of the Organization and the Program Committees of the 2018 edition of the Inria Science Days (*Journées Scientifiques Inria*)

8.1.2. Scientific Events Selection

8.1.2.1. Member of the Conference Program Committees

- Emmanuel Prados was member of the Program Committees of the Workshop on "Information and decision-making systems and democracy" at the INFORSID 2018 Congress , Nante, France 28-31 May 2018.
- Emmanuel Prados was member of the Program Committees of the Workshop on "EIAH and sustainability" at RJC-EIAH 2018, Besançon, France 3-5 April 2018.

8.1.3. Journal

8.1.3.1. Reviewer - Reviewing Activities

- P. Sturm was reviewer for Environment and Planning B: Urban Analytics and City Science, International Journal of Computer Vision, and Journal of Mathematical Imaging and Vision.

8.1.4. Invited Talks

- E. Prados give an invited talk to the **Journées scientitifiques de l'Inria** "", Bordeaux, France, june 27-29, 2018. Organizers: Inria (http://journees-scientifiques2018.inria.fr/).
- E. Prados give an invited talk to the **LIED Thematic Seminar** "Effondrement, déclin ou catastrophologie ? : théories sur les risques et l'avenir des sociétés" (Laboratoire interdisciplinaire sur les énergies de demain) with Pierre-Yves Longaretti and Grégoire Chambaz, Paris, France, June 19, 2018. Organizers: Petros Chatzimpiros.
- E. Prados give an invited talk to the **round table** on military nuclear [links with the risk of collapse of modern society] at the **first anniversary of the "Nuclear Knowledges" program**, Science Po, Paris, France, May 31, 2018. Organizers: Benoît Pelopidas.

- E. Prados give an invited talk to the **Conference for the Société Francaise de Physique** : “Vers un effondrement de notre société moderne ?”, Grenoble, France, June 7, 2018 ; Conférence with Pierre-Yves Longaretti. Organizer: Thomas Thuillier.

8.1.5. Research Administration

- P. Sturm is, since 2015, Deputy Scientific Director of Inria, in charge of the domain Perception–Cognition–Interaction
- E. Prados was in charge of the working group “Digital tools for the progress of society” for the Strategic Plan of Inria, 2017-2018. His work resulted in the writing of the Scientific Challenges entitled “Digital world, society, and complexity”.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

As teacher-researchers, Denis Dupré and Serge Fenet completed their full teaching service in 2018 (full-time) in their field of study. Regis Perrier did a half service.

8.2.2. Supervision

PhD : Luciano Gervasoni, *Contributions to the formalization and implementation of spatial urban indices using open data: application to urban sprawl studies*, University Grenoble Alpes, November 19th, 2018, Peter Sturm (supervisor) and Serge Fenet (co-supervisor).

PhD in progress : Michela Bevione, *Socio-ecological transition, wealth creation and territorial metabolism: the case of Maurienne Valley*, October 2016, Pierre-Yves Longaretti (co-supervisor) and Nicolas Buclet (supervisor). Phd thesis in collaboration with the PACTE Laboratory. This interdisciplinary research on territorial ecology aims at bridging the gap between quantitative modeling and socio-economic and qualitative approaches to better understand the interactions between human society and the environment.

8.2.3. Juries

- P. Sturm chaired the Habilitation committee of Rémi Boutteau (Normandie Université, president), was reviewer of the PhD theses of Laura Fernàndez Julià (Université Paris-Est) and Mathias Gallardo (Université Clermont Auvergne) and chaired the PhD committee of Fatima Aziz (Université de Limoges).

8.3. Popularization

8.3.1. Conferences “Understanding and Acting” :

Saving Civilization is not a spectator sport
Lester Brown

Following a dynamics of exponential growth in a finite world, humanity today faces a number of unprecedented and tightly interlinked challenges. With a growing number of environmental limits being largely and irreversibly exceeded (GHG concentrations in the atmosphere, biodiversity loss, soil erosion, freshwater shortages...), social, economic, geopolitical, humanitarian (etc.) consequences are becoming more urgent than ever to address, while the threat of an uncontrolled global collapse is now more than a prospect. It is urgent to initiate deep, structural, socioeconomic changes on virtually all aspects of our increasingly global societies (economics, industrial and agricultural production, consumption, education, all requiring major new local and global policies).

In view of these facts, the STEEP research team has initiated in 2016 a series of conferences-debates entitled "Understanding & Acting" ("Comprendre et agir") that examines these issues in order to help researchers and citizens to increase their awareness of the various issues at stake in order to initiate relevant individual and collective actions. From now on, the scientific community at large must realize that its duty also lies in helping citizens to better understand these issues. If the fraction of people in society whose privilege is to be paid to think about society's problems do not seize this opportunity in the critical times we face, who will? Researchers must become more involved in the search of socioeconomic alternatives and help citizens to implement them. The interactions between researchers and citizens must also to be reinvented.

The presentations of this series of conferences typically last between 30 to 45 minutes; they are followed by a 45 minute public debate with the audience. The presentations are captured on video and then made directly accessible on the YouTube Channel "Comprendre et Agir". At the end of 2017 the YouTube channel reached almost 45,000 views with a rate of integral viewings re¬maining at above 25%.

Examples of conferences given in 2016-2017:

- "Why and how to recognize money as a common good?" by Jean-Michel Servet, Graduate institute of international and development studies (IHEID) at Genève.
- "What alternative to the crisis of representation?" by Loic Blondiaux, University of Paris I Panthéon-Sorbonne.
- "Reinventing agriculture and food in the 21st century?" by Gilles Billen, CNRS.
- "Can the commons movement re-activate democracy? by Christian Laval, University of Paris Ouest Nanterre La Défense.
- "Limiting climate change: Why do we not do it?" by Denis Dupré, University of Grenoble Alpes.
- "Understanding society collapses. What future for ours? " by Emmanuel Prados, Inria.
- "Environment, economy, collapse: squaring the circle?" by Pierre-Yves Longaretti, CNRS.

Link to the web page of the series [program, abstracts, dates, complements etc.]:
https://team.inria.fr/steep/les-conferences-debats-comprendre-et-agir/
Link to the YouTube channel :
https://www.youtube.com/channel/UCJbcXCcOA63M8VMysAbmt_A

8.3.2. *Other popularization activities :*

- Emmanuel Prados was guest speaker at
  - the **public debate "defying disaster"** that held at Grenoble City Hall., Grenoble, France, 13 December 2018. Organizers: City of Grenoble. "Climat, défier la catastrophe : vers un nouvel humanisme"
    http://www.grenoble.fr/uploads/Externe/a1/757_069_Conference-Climat-defier-la-catastrophe-vers-un-nouvel-humanisme.pdf
    Video of the conference on YouTube : https://www.youtube.com/watch?v=IfvFHx0kYa0
  - the **exchange day on travel modelling in Grenoble** "", Grenoble, France, 5 October 2018. Organizers: Agence d'urbanisme de Grenoble.
  - the **Controversation** entitled "Digital technology in the face of environmental impasse: an exit door or a way forward?", Aix en Provence, France, 17th of February 2018. Organizers: association "Tous Chercheurs" in collaboration with the students of the School of Journalism and Communication of Aix-Marseille ; Amina Mokrane.
  - the **Opening Conference at the Alterre Day** «Voyages en transitions » : "De l'urgence de tout changer face à l'effondrement qui vient", Dijon, France, March 21th 2018. Organizers: Alterre Bourgogne-Franche-Comté (Agence régionale pour l'environnement et le développement soutenable) ; Fabienne Lapiche-Jaouen.
    https://www.alterrebourgognefranchecomte.org/c/agenda/10604/journee-alterre-sur-le-theme-des-transitions?

- Emmanuel Prados participated **Fictionnalisation de controverse** in the research-action project FORCCAST. 22 April, 2018, Paris.
- Emmanuel Prados gave a conference at **STMicroelectronics Grenoble** : ‘Collapse: understanding and accepting the game over of our modern society”, Grenoble, France, March 8 2018. Organizers: Philippe Vincent and Violaine Pasini.

9. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] R. BOUTTEAU, P. STURM, P. VASSEUR, C. DEMONCEAUX.*Circular Laser/Camera-based attitude and altitude estimation: minimal and robust solutions*, in "Journal of Mathematical Imaging and Vision", March 2018, vol. 60, n^{o} 3, p. 382-400 [*DOI :* 10.1007/S10851-017-0764-Y], https://hal.inria.fr/hal-01684031

[2] T. CAPELLE, P. STURM, A. VIDARD, B. MORTON.*Calibration of the Tranus Land Use Module: Optimisation-Based Algorithms, their Validation, and Parameter Selection by Statistical Model Selection*, in "Computers, Environment and Urban Systems", 2018 [*DOI :* 10.1016/J.COMPENVURBSYS.2017.04.009], https://hal.inria.fr/hal-01519654

[3] J.-Y. COURTONNE, P.-Y. LONGARETTI, D. DUPRÉ.*Uncertainties of Domestic Road Freight Statistics: Insights for Regional Material Flow Studies*, in "Journal of Industrial Ecology", 2018, p. 1-36 [*DOI :* 10.1111/JIEC.12651], https://hal.archives-ouvertes.fr/hal-01610261

International Conferences with Proceedings

[4] T. BIRDAL, B. BUSAM, N. NAVAB, S. ILIC, P. STURM.*A Minimalist Approach to Type-Agnostic Detection of Quadrics in Point Clouds*, in "CVPR 2018 - IEEE Conference on Computer Vision and Pattern Recognition", Salt Lake City, United States, June 2018, p. 3530-3540, https://hal.inria.fr/hal-01855977

[5] L. GERVASONI, S. FENET, R. PERRIER, P. STURM.*Convolutional neural networks for disaggregated population mapping using open data*, in "DSAA 2018 - 5th IEEE International Conference on Data Science and Advanced Analytics", Turin, Italy, October 2018, p. 1-10, https://hal.inria.fr/hal-01852585

[6] L. GERVASONI, S. FENET, P. STURM.*Une méthode pour l’estimation désagrégée de données de population à l’aide de données ouvertes*, in "EGC 2018 - 18ème Conférence Internationale sur l’Extraction et la Gestion des Connaissances", Paris, France, January 2018, 1, https://hal.inria.fr/hal-01667975

Conferences without Proceedings

[7] S. CARRARA, M. BEVIONE, H. S. DE BOER, D. GERNAAT, S. MIMA, R. C. PIETZCKER, M. TAVONI.*Exploring pathways of solar PV learning-by-doing in Integrated Assessment Models*, in "3rd AIEE Energy Symposium on Current and future challenges to energy security-the energy transition", Milan, Italy, Associazione Italia Economisti dell’ Energia, December 2018, p. 1-18, https://hal.archives-ouvertes.fr/hal-01960855

Other Publications

[8] S. F. DIOP.*Modélisation des liens entre la morphologie urbaine, la génération de polluants et leur concentration dans l'air par l'approche PLS path modelling*, Université Grenoble Alpes, September 2018, https://hal.inria.fr/hal-01880194

References in notes

[9] L. GERVASONI, M. BOSCH, S. FENET, P. STURM.*LUM-OSM : une plateforme pour l'évaluation de la mixité urbaine à partir de données participatives*, in "Atelier Gestion et Analyse des données Spatiales et Temporelles", Grenoble, France, Cyril De Runz and Eric Kergosien and Thomas Guyet and Christian Sallaberry, January 2017, https://hal.inria.fr/hal-01548341

[10] L. REZAKHANLOU, S. FENET, L. GERVASONI, P. STURM.*USAT (Urban Sprawl Analysis Toolkit) : une plateforme web d'analyse de l'étalement urbain à partir de données massives ouvertes*, in "Atelier Démo de la conférence SAGEO (Spatial Analysis and Geomatics)", Rouen, France, November 2017, https://hal.inria.fr/hal-01610738

Project-Team THOTH

Learning visual models from large-scale data

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Vision, perception and multimedia interpretation

Table of contents

Project-Team THOTH

Creation of the Team: 2016 January 01, updated into Project-Team: 2016 March 01

Keywords:

Computer Science and Digital Science:

A3.4. - Machine learning and statistics
A5.3. - Image processing and analysis
A5.4. - Computer vision
A5.9. - Signal processing
A6.2.6. - Optimization
A8.2. - Optimization
A9.2. - Machine learning
A9.3. - Signal analysis
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B5.6. - Robotic systems
B8.4. - Security and personal assistance
B8.5. - Smart society
B9.5.1. - Computer science
B9.5.6. - Data science

1. Team, Visitors, External Collaborators

Research Scientists

Julien Mairal [Team leader, Inria, Researcher, HDR]
Karteek Alahari [Inria, Researcher]
Grégory Rogez [Inria, Starting Research Position]
Cordelia Schmid [Team leader, Inria, Senior Researcher, HDR]
Jakob Verbeek [Inria, Senior Researcher, HDR]

Technical Staff

Ghislain Durif [Inria]
François Gindraud [Inria, from Oct 2018]
Xavier Martin [Inria]

PhD Students

Minttu Alakuijala [Inria, from Nov 2018]
Alberto Bietti [Inria]
Mathilde Caron [Inria, from Apr until Sep 2018, Facebook CIFRE, from Oct 2018]
Dexiong Chen [Univ. Paris-Saclay]
Guilhem Cheron [Inria, until Dec 2018]
Mikita Dvornik [Inria]
Maha Elbayad [Univ. Grenoble Alpes]
Valentin Gabeur [Inria, from Oct 2018]
Yana Hasson [Inria]
Ekaterina Iakovleva [Inria, from Oct 2018]
Roman Klokov [Inria]

Andrei Kulunchakov [Inria]
Pauline Luc [Facebook CIFRE]
Thomas Lucas [Univ. Grenoble Alpes]
Grégoire Mialon [Inria, from Oct 2018]
Alexander Pashevich [Inria]
Alexandre Sablayrolles [Facebook CIFRE]
Konstantin Shmelkov [Inria]
Vladyslav Sydorov [Inria]
Pavel Tokmakov [Inria, until Jan 2018]
Gul Varol [Inria]
Nitika Verma [Univ. Grenoble Alpes]
Daan Wynen [Inria]

Post-Doctoral Fellows

Henrique Morimitsu [Inria, until Feb 2018]
Adria Ruiz Ovejero [Inria, from Jan 2018]

Visiting Scientist

Pia Bideau [Univ. Massachusetts Amherst, from Sep 2018 until Dec 2018]

Administrative Assistant

Nathalie Gillot [Inria]

2. Overall Objectives

2.1. Overall Objectives

In 2021, it is expected that nearly 82% of the Internet traffic will be due to videos, and that it would take an individual over 5 million years to watch the amount of video that will cross global IP networks each month by then. Thus, there is a pressing and in fact increasing demand to annotate and index this visual content for home and professional users alike. The available text and speech-transcript metadata is typically not sufficient by itself for answering most queries, and visual data must come into play. On the other hand, it is not imaginable to learn the models of visual content required to answer these queries by manually and precisely annotating every relevant concept, object, scene, or action category in a representative sample of everyday conditions—if only because it may be difficult, or even impossible to decide a priori what are the relevant categories and the proper granularity level. This suggests reverting back to the original metadata as source of annotation, despite the fact that the information it provides is typically sparse (e.g., the location and overall topic of newscasts in a video archive) and noisy (e.g., a movie script may tell us that two persons kiss in some scene, but not when, and the kiss may occur off screen or not have survived the final cut). On the other hand, this weak form of "embedded annotation" is rich and diverse, and mining the corresponding visual data from the web, TV or film archives guarantees that it is representative of the many different scene settings depicted in situations typical of on-line content. Thus, leveraging this largely untapped source of information, rather than attempting to hand label all possibly relevant visual data, is a key to the future use of on-line imagery.

Today's object recognition and scene understanding technology operates in a very different setting; it mostly relies on fully supervised classification engines, and visual models are essentially (piecewise) rigid templates learned from hand labeled images. The sheer scale of on-line data and the nature of the embedded annotation call for a departure from this fully supervised scenario. The main idea of the Thoth project-team is to develop a new framework for learning the structure and parameters of visual models by actively exploring large digital image and video sources (off-line archives as well as growing on-line content, with millions of images and thousands of hours of video), and exploiting the weak supervisory signal provided by the accompanying metadata. This huge volume of visual training data will allow us to learn complex non-linear models with a large number of parameters, such as deep convolutional networks and higher-order graphical models. This is an ambitious goal, given the sheer volume and intrinsic variability of the visual data available on-line, and the lack of a universally accepted formalism for modeling it. Yet, the potential payoff is a breakthrough in visual object recognition and scene understanding capabilities. Further, recent advances at a smaller scale suggest that this is realistic. For example, it is already possible to determine the identity of multiple people from news images and their captions, or to learn human action models from video scripts. There has also been recent progress in adapting supervised machine learning technology to large-scale settings, where the training data is very large and potentially infinite, and some of it may not be labeled. Methods that adapt the structure of visual models to the data are also emerging, and the growing computational power and storage capacity of modern computers are enabling factors that should of course not be neglected.

One of the main objective of Thoth is to transform massive visual data into trustworthy knowledge libraries. For that, it addresses several challenges.

- designing and learning structured models capable of representing complex visual information.
- learning visual models from minimal supervision or unstructured meta-data.
- large-scale learning and optimization.

3. Research Program

3.1. Designing and learning structured models

The task of understanding image and video content has been interpreted in several ways over the past few decades, namely image classification, detecting objects in a scene, recognizing objects and their spatial extents in an image, estimating human poses, recovering scene geometry, recognizing activities performed by humans. However, addressing all these problems individually provides us with a partial understanding of the scene at best, leaving much of the visual data unexplained.

One of the main goals of this research axis is to go beyond the initial attempts that consider only a subset of tasks jointly, by developing novel models for a more complete understanding of scenes to address all the component tasks. We propose to incorporate the structure in image and video data explicitly into the models. In other words, our models aim to satisfy the complex sets of constraints that exist in natural images and videos. Examples of such constraints include: (i) relations between objects, like signs for shops indicate the presence of buildings, people on a road are usually walking or standing, (ii) higher-level semantic relations involving the type of scene, geographic location, and the plausible actions as a global constraint, e.g., an image taken at a swimming pool is unlikely to contain cars, (iii) relating objects occluded in some of the video frames to content in other frames, where they are more clearly visible as the camera or the object itself move, with the use of long-term trajectories and video object proposals.

This research axis will focus on three topics. The first is developing deep features for video. This involves designing rich features available in the form of long-range temporal interactions among pixels in a video sequence to learn a representation that is truly spatio-temporal in nature. The focus of the second topic is the challenging problem of modeling human activities in video, starting from human activity descriptors to building intermediate spatio-temporal representations of videos, and then learning the interactions among humans, objects and scenes temporally. The last topic is aimed at learning models that capture the relationships among several objects and regions in a single image scene, and additionally, among scenes in the case of an image collection or a video. The main scientific challenges in this topic stem from learning the structure of the probabilistic graphical model as well as the parameters of the cost functions quantifying the relationships among its entities. In the following we will present work related to all these three topics and then elaborate on our research directions.

- **Deep features for vision.** Deep learning models provide a rich representation of complex objects but in return have a large number of parameters. Thus, to work well on difficult tasks, a large amount of data is required. In this context, video presents several advantages: objects are observed from a large range of viewpoints, motion information allows the extraction of moving objects and parts, and objects can be differentiated by their motion patterns. We initially plan to develop deep features for videos that incorporate temporal information at multiple scales. We then plan to further exploit the rich content in video by incorporating additional cues, such as the detection of people and their body-joint locations in video, minimal prior knowledge of the object of interest, with the goal of learning a representation that is more appropriate for video understanding. In other words, a representation that is learned from video data and targeted at specific applications. For the application of recognizing human activities, this involves learning deep features for humans and their body-parts with all their spatiotemporal variations, either directly from raw video data or "pre-processed" videos containing human detections. For the application of object tracking, this task amounts to learning object-specific deep representations, further exploiting the limited annotation provided to identify the object.

- **Modeling human activities in videos.** Humans and their activities are not only one of the most frequent and interesting subjects in videos but also one of the hardest to analyze owing to the complexity of the human form, clothing and movements. As part of this task, the Thoth project-team plans to build on state-of-the-art approaches for spatio-temporal representation of videos. This will involve using the dominant motion in the scene as well as the local motion of individual parts undergoing a rigid motion. Such motion information also helps in reasoning occlusion relationships among people and objects, and the state of the object. This novel spatio-temporal representation ultimately provides the equivalent of object proposals for videos, and is an important component for learning algorithms using minimal supervision. To take this representation even further, we aim to integrate the proposals and the occlusion relationships with methods for estimating human pose in videos, thus leveraging the interplay among body-joint locations, objects in the scene, and the activity being performed. For example, the locations of shoulder, elbow and wrist of a person drinking coffee are constrained to move in a certain way, which is completely different from the movement observed when a person is typing. In essence, this step will model human activities by dynamics in terms of both low-level movements of body-joint locations and global high-level motion in the scene.
- **Structured models.** The interactions among various elements in a scene, such as, the objects and regions in it, the motion of object parts or entire objects themselves, form a key element for understanding image or video content. These rich cues define the structure of visual data and how it evolves spatio-temporally. We plan to develop a novel graphical model to exploit this structure. The main components in this graphical model are spatio-temporal regions (in the case of video or simply image regions), which can represent object parts or entire objects themselves, and the interactions among several entities. The dependencies among the scene entities are defined with a higher order or a global cost function. A higher order constraint is a generalization of the pairwise interaction term, and is a cost function involving more than two components in the scene, e.g., several regions, whereas a global constraint imposes a cost term over the entire image or video, e.g., a prior on the number of people expected in the scene. The constraints we plan to include generalize several existing methods, which are limited to pairwise interactions or a small restrictive set of higher-order costs. In addition to learning the parameters of these novel functions, we will focus on learning the structure of the graph itself—a challenging problem that is seldom addressed in current approaches. This provides an elegant way to go beyond state-of-the-art deep learning methods, which are limited to learning the high-level interaction among parts of an object, by learning the relationships among objects.

3.2. Learning of visual models from minimal supervision

Today's approaches to visual recognition learn models for a limited and fixed set of visual categories with fully supervised classification techniques. This paradigm has been adopted in the early 2000's, and within it enormous progress has been made over the last decade.

The scale and diversity in today's large and growing image and video collections (such as, e.g., broadcast archives, and personal image/video collections) call for a departure from the current paradigm. This is the case because to answer queries about such data, it is unfeasible to learn the models of visual content by manually and precisely annotating every relevant concept, object, scene, or action category in a representative sample of everyday conditions. For one, it will be difficult, or even impossible to decide a-priori what are the relevant categories and the proper granularity level. Moreover, the cost of such annotations would be prohibitive in most application scenarios. One of the main goals of the Thoth project-team is to develop a new framework for learning visual recognition models by actively exploring large digital image and video sources (off-line archives as well as growing on-line content), and exploiting the weak supervisory signal provided by the accompanying metadata (such as captions, keywords, tags, subtitles, or scripts) and audio signal (from which we can for example extract speech transcripts, or exploit speaker recognition models).

Textual metadata has traditionally been used to index and search for visual content. The information in metadata is, however, typically sparse (e.g., the location and overall topic of newscasts in a video archive [0]) and noisy (e.g., a movie script may tell us that two persons kiss in some scene, but not when, and the kiss may occur off screen or not have survived the final cut). For this reason, metadata search should be complemented by visual content based search, where visual recognition models are used to localize content of interest that is not mentioned in the metadata, to increase the usability and value of image/video archives. *The key insight that we build on in this research axis is that while the metadata for a single image or video is too sparse and noisy to rely on for search, the metadata associated with large video and image databases collectively provide an extremely versatile source of information to learn visual recognition models.* This form of "embedded annotation" is rich, diverse and abundantly available. Mining these correspondences from the web, TV and film archives, and online consumer generated content sites such as Flickr, Facebook, or YouTube, guarantees that the learned models are representative for many different situations, unlike models learned from manually collected fully supervised training data sets which are often biased.

The approach we propose to address the limitations of the fully supervised learning paradigm aligns with "Big Data" approaches developed in other areas: we rely on the orders-of-magnitude-larger training sets that have recently become available with metadata to compensate for less explicit forms of supervision. This will form a sustainable approach to learn visual recognition models for a much larger set of categories with little or no manual intervention. Reducing and ultimately removing the dependency on manual annotations will dramatically reduce the cost of learning visual recognition models. This in turn will allow such models to be used in many more applications, and enable new applications based on visual recognition beyond a fixed set of categories, such as natural language based querying for visual content. This is an ambitious goal, given the sheer volume and intrinsic variability of the every day visual content available on-line, and the lack of a universally accepted formalism for modeling it. Yet, the potential payoff is a breakthrough in visual object recognition and scene understanding capabilities.

This research axis is organized into the following three sub-tasks:

- **Weakly supervised learning.** For object localization we will go beyond current methods that learn one category model at a time and develop methods that learn models for different categories concurrently. This allows "explaining away" effects to be leveraged, i.e., if a certain region in an image has been identified as an instance of one category, it cannot be an instance of another category at the same time. For weakly supervised detection in video we will consider detection proposal methods. While these are effective for still images, recent approaches for the spatio-temporal domain need further improvements to be similarly effective. Furthermore, we will exploit appearance and motion information jointly over a set of videos. In the video domain we will also continue to work on learning recognition models from subtitle and script information. The basis of leveraging the script data which does not have a temporal alignment with the video, is to use matches in the narrative in the script and the subtitles (which do have a temporal alignment with the video). We will go beyond simple correspondences between names and verbs relating to self-motion, and match more complex sentences related to interaction with objects and other people. To deal with the limited amount of occurrences of such actions in a single movie, we will consider approaches that learn action models across a collection of movies.
- **Online learning of visual models.** As a larger number of visual category models is being learned,

[0] For example at the Dutch national broadcast archive Netherlands Institute of Sound and Vision, with whom we collaborated in the EU FP7 project AXES, typically one or two sentences are used in the metadata to describe a one hour long TV program.

online learning methods become important, since new training data and categories will arrive over time. We will develop online learning methods that can incorporate new examples for existing category models, and learn new category models from few examples by leveraging similarity to related categories using multi-task learning methods. Here we will develop new distance-based classifiers and attribute and label embedding techniques, and explore the use of NLP techniques such as skipgram models to automatically determine between which classes transfer should occur. Moreover, NLP will be useful in the context of learning models for many categories to identify synonyms, and to determine cases of polysemy (e.g. jaguar car brand v.s. jaguar animal), and merge or refine categories accordingly. Ultimately this will result in methods that are able to learn an“encyclopedia” of visual models.

- **Visual search from unstructured textual queries.** We will build on recent approaches that learn recognition models on-the-fly (as the query is issued) from generic image search engines such as Google Images. While it is feasible to learn models in this manner in a matter of seconds, it is challenging to use the model to retrieve relevant content in real-time from large video archives of more than a few thousand hours. To achieve this requires feature compression techniques to store visual representations in memory, and cascaded search techniques to avoid exhaustive search. This approach, however, leaves untouched the core problem of how to associate visual material with the textual query in the first place. The second approach we will explore is based on image annotation models. In particular we will go beyond image-text retrieval methods by using recurrent neural networks such as Elman networks or long short-term memory (LSTM) networks to generate natural language sentences to describe images.

3.3. Large-scale learning and optimization

We have entered an era of massive data acquisition, leading to the revival of an old scientific utopia: it should be possible to better understand the world by automatically converting data into knowledge. It is also leading to a new economic paradigm, where data is a valuable asset and a source of activity. Therefore, developing scalable technology to make sense of massive data has become a strategic issue. Computer vision has already started to adapt to these changes.

In particular, very high dimensional models such as deep networks are becoming highly popular and successful for visual recognition. This change is closely related to the advent of big data. On the one hand, these models involve a huge number of parameters and are rich enough to represent well complex objects such as natural images or text corpora. On the other hand, they are prone to overfitting (fitting too closely to training data without being able to generalize to new unseen data) despite regularization; to work well on difficult tasks, they require a large amount of labelled data that has been available only recently. Other cues may explain their success: the deep learning community has made significant engineering efforts, making it possible to learn in a day on a GPU large models that would have required weeks of computations on a traditional CPU, and it has accumulated enough empirical experience to find good hyper-parameters for its networks.

To learn the huge number of parameters of deep hierarchical models requires scalable optimization techniques and large amounts of data to prevent overfitting. This immediately raises two major challenges: how to learn without large amounts of labeled data, or with weakly supervised annotations? How to efficiently learn such huge-dimensional models? To answer the above challenges, we will concentrate on the design and theoretical justifications of deep architectures including our recently proposed deep kernel machines, with a focus on weakly supervised and unsupervised learning, and develop continuous and discrete optimization techniques that push the state of the art in terms of speed and scalability.

This research axis will be developed into three sub-tasks:

- **Deep kernel machines for structured data.** Deep kernel machines combine advantages of kernel methods and deep learning. Both approaches rely on high-dimensional models. Kernels implicitly operate in a space of possibly infinite dimension, whereas deep networks explicitly construct high-dimensional nonlinear data representations. Yet, these approaches are complementary: Kernels can be built with deep learning principles such as hierarchies and convolutions, and approximated by multilayer neural networks. Furthermore, kernels work with structured data and have well understood theoretical principles. Thus, a goal of the Thoth project-team is to design and optimize the training of such deep kernel machines.

- **Large-scale parallel optimization.** Deep kernel machines produce nonlinear representations of input data points. After encoding these data points, a learning task is often formulated as a *large-scale convex optimization problem*; for example, this is the case for linear support vector machines, logistic regression classifiers, or more generally many empirical risk minimization formulations. We intend to pursue recent efforts for making convex optimization techniques that are dedicated to machine learning more scalable. Most existing approaches address scalability issues either in model size (meaning that the function to minimize is defined on a domain of very high dimension), or in the amount of training data (typically, the objective is a large sum of elementary functions). There is thus a large room for improvements for techniques that jointly take these two criteria into account.
- **Large-scale graphical models.** To represent structured data, we will also investigate graphical models and their optimization. The challenge here is two-fold: designing an adequate cost function and minimizing it. While several cost functions are possible, their utility will be largely determined by the efficiency and the effectiveness of the optimization algorithms for solving them. It is a combinatorial optimization problem involving billions of variables and is NP-hard in general, requiring us to go beyond the classical approximate inference techniques. The main challenges in minimizing cost functions stem from the large number of variables to be inferred, the inherent structure of the graph induced by the interaction terms (e.g., pairwise terms), and the high-arity terms which constrain multiple entities in a graph.

3.4. Datasets and evaluation

Standard benchmarks with associated evaluation measures are becoming increasingly important in computer vision, as they enable an objective comparison of state-of-the-art approaches. Such datasets need to be relevant for real-world application scenarios; challenging for state-of-the-art algorithms; and large enough to produce statistically significant results.

A decade ago, small datasets were used to evaluate relatively simple tasks, such as for example interest point matching and detection. Since then, the size of the datasets and the complexity of the tasks gradually evolved. An example is the Pascal Visual Object Challenge with 20 classes and approximately 10,000 images, which evaluates object classification and detection. Another example is the ImageNet challenge, including thousands of classes and millions of images. In the context of video classification, the TrecVid Multimedia Event Detection challenges, organized by NIST, evaluate activity classification on a dataset of over 200,000 video clips, representing more than 8,000 hours of video, which amounts to 11 months of continuous video.

Almost all of the existing image and video datasets are annotated by hand; it is the case for all of the above cited examples. In some cases, they present limited and unrealistic viewing conditions. For example, many images of the ImageNet dataset depict upright objects with virtually no background clutter, and they may not capture particularly relevant visual concepts: most people would not know the majority of subcategories of snakes cataloged in ImageNet. This holds true for video datasets as well, where in addition a taxonomy of action and event categories is missing.

Our effort on data collection and evaluation will focus on two directions. First, we will design and assemble video datasets, in particular for action and activity recognition. This includes defining relevant taxonomies of actions and activities. Second, we will provide data and define evaluation protocols for weakly supervised learning methods. This does not mean of course that we will forsake human supervision altogether: some amount of ground-truth labeling is necessary for experimental validation and comparison to the state of the art. Particular attention will be payed to the design of efficient annotation tools.

Not only do we plan to collect datasets, but also to provide them to the community, together with accompanying evaluation protocols and software, to enable a comparison of competing approaches for action recognition and large-scale weakly supervised learning. Furthermore, we plan to set up evaluation servers together with leader-boards, to establish an unbiased state of the art on held out test data for which the ground-truth annotations are not distributed. This is crucial to avoid tuning the parameters for a specific dataset and to guarantee a fair evaluation.

- **Action recognition.** We will develop datasets for recognizing human actions and human-object interactions (including multiple persons) with a significant number of actions. Almost all of today's action recognition datasets evaluate classification of short video clips into a number of predefined categories, in many cases a number of different sports, which are relatively easy to identify by their characteristic motion and context. However, in many real-world applications the goal is to identify and localize actions in entire videos, such as movies or surveillance videos of several hours. The actions targeted here are "real-world" and will be defined by compositions of atomic actions into higher-level activities. One essential component is the definition of relevant taxonomies of actions and activities. We think that such a definition needs to rely on a decomposition of actions into poses, objects and scenes, as determining all possible actions without such a decomposition is not feasible. We plan to provide annotations for spatio-temporal localization of humans as well as relevant objects and scene parts for a large number of actions and videos.
- **Weakly supervised learning.** We will collect weakly labeled images and videos for training. The collection process will be semi-automatic. We will use image or video search engines such as Google Image Search, Flickr or YouTube to find visual data corresponding to the labels. Initial datasets will be obtained by manually correcting whole-image/video labels, i.e., the approach will evaluate how well the object model can be learned if the entire image or video is labeled, but the object model has to be extracted automatically. Subsequent datasets will features noisy and incorrect labels. Testing will be performed on PASCAL VOC'07 and ImageNet, but also on more realistic datasets similar to those used for training, which we develop and manually annotate for evaluation. Our dataset will include both images and videos, the categories represented will include objects, scenes as well as human activities, and the data will be presented in realistic conditions.
- **Joint learning from visual information and text.** Initially, we will use a selection from the large number of movies and TV series for which scripts are available on-line, see for example http://www.dailyscript.com and http://www.weeklyscript.com. These scripts can easily be aligned with the videos by establishing correspondences between script words and (timestamped) spoken ones obtained from the subtitles or audio track. The goal is to jointly learn from visual content and text. To measure the quality of such a joint learning, we will manually annotate some of the videos. Annotations will include the space-time locations of the actions as well as correct parsing of the sentence. While DVDs will, initially, receive most attention, we will also investigate the use of data obtained from web pages, for example images with captions, or images and videos surrounded by text. This data is by nature more noisy than scripts.

4. Application Domains

4.1. Visual applications

Any solution to automatically understanding images and videos on a semantic level will have an immediate impact on a wide range of applications. For example:

- Semantic-level image and video access is highly relevant for visual search on the Web, in professional archives and personal collections.
- Visual data organization is applicable to organizing family photo and video albums as well as to large-scale information retrieval.
- Visual object recognition has potential applications ranging from surveillance, service robotics for assistance in day-to-day activities as well as the medical domain.
- Action recognition is highly relevant to visual surveillance, assisted driving and video access.
- Real-time scene understanding is relevant for human interaction through devices such as HoloLens, Oculus Rift.

4.2. Pluri-disciplinary research

Machine learning is intrinsically pluri-disciplinary. By developing large-scale machine learning models and algorithms for processing data, the Thoth team became naturally involved in pluri-disciplinary collaborations that go beyond visual modelling. In particular,

- extensions of unsupervised learning techniques originally developed for modelling the statistics of natural images have been deployed in neuro-imaging for fMRI data with the collaboration of the Parietal team from Inria.
- similarly, deep convolutional data representations, also originally developed for visual data, have been successfully extended to the processing of biological sequences, with collaborators from bio-informatics.
- Thoth also collaborates with experts in natural language and text processing, for applications where visual modalities need to be combined with text data.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Alberto Bietti received the Jean-Claude Dodu 2018 prize at Journees SMAI-MODE, Autrans.
- Pauline Luc was one of the top-200 reviewers at NeurIPS 2018.
- Grégory Rogez and Cordelia Schmid received an Amazon Academic Research Award.
- Cordelia Schmid received the Koenderink prize for fundamental contributions in computer vision that have withstood the test of time at ECCV 2018.

5.1.2. Dissemination

- The team co-organized PAISS 2018, an international AI summer school in Grenoble. This event brought together 200 participants representing 44 different nationalities. The participants were selected from 700 applications, with 60% students, 15% academics, and 25% industrials. 25% of these participants were women.

6. New Software and Platforms

6.1. LCR-Net

Localization-Classification-Regression Network for Human Pose
KEYWORDS: Object detection - Recognition of human movement
FUNCTIONAL DESCRIPTION: We propose an end-to-end architecture for joint 2D and 3D human pose estimation in natural images. Key to our approach is the generation and scoring of a number of pose proposals per image, which allows us to predict 2D and 3D pose of multiple people simultaneously. Our architecture contains 3 main components: 1) the pose proposal generator that suggests potential poses at different locations in the image, 2) a classifier that scores the different pose proposals , and 3) a regressor that refines pose proposals both in 2D and 3D.

- Participants: Grégory Rogez, Philippe Weinzaepfel and Cordelia Schmid
- Contact: Grégory Rogez
- Publication: LCR-Net: Localization-Classification-Regression for Human Pose
- URL: https://thoth.inrialpes.fr/src/LCR-Net/

6.2. CKN-seq

Convolutional Kernel Networks for Biological Sequences
KEYWORD: Bioinformatics
SCIENTIFIC DESCRIPTION: The growing amount of biological sequences available makes it possible to learn genotype-phenotype relationships from data with increasingly high accuracy. By exploiting large sets of sequences with known phenotypes, machine learning methods can be used to build functions that predict the phenotype of new, unannotated sequences. In particular, deep neural networks have recently obtained good performances on such prediction tasks, but are notoriously difficult to analyze or interpret. Here, we introduce a hybrid approach between kernel methods and convolutional neural networks for sequences, which retains the ability of neural networks to learn good representations for a learning problem at hand, while defining a well characterized Hilbert space to describe prediction functions. Our method outperforms state-of-the-art convolutional neural networks on a transcription factor binding prediction task while being much faster to train and yielding more stable and interpretable results.
FUNCTIONAL DESCRIPTION: CKN-Seq is a software package for predicting transcription factor binding sites. It was shipped with the BiorXiv preprint

D. Chen, L. Jacob, and J. Mairal. Predicting Transcription Factor Binding Sites with Convolutional Kernel Networks. 2017.

The software is implemented in PyTorch.

- Participants: Laurent Jacob, Dexiong Chen and Julien Mairal
- Partners: CNRS - UGA
- Contact: Julien Mairal
- Publication: Biological Sequence Modeling with Convolutional Kernel Networks
- URL: https://gitlab.inria.fr/dchen/CKN-seq

6.3. Loter

Loter: A software package to infer local ancestry for a wide range of species
KEYWORDS: Local Ancestry Inference - Bioinformatics
SCIENTIFIC DESCRIPTION: Admixture between populations provides opportunity to study biological adaptation and phenotypic variation. Admixture studies can rely on local ancestry inference for admixed individuals, which consists of computing at each locus the number of copies that originate from ancestral source populations. Loter is a software package that does not require any biological parameter besides haplotype data in order to make local ancestry inference available for a wide range of species.
FUNCTIONAL DESCRIPTION: Loter is a Python package for haplotype phasing and local ancestry inference.
NEWS OF THE YEAR: The software package was shipped with the biorxiv preprint T. Dias-Alves, J. Mairal, and M. Blum. Loter: A Software Package to Infer Local Ancestry for a Wide Range of Species. preprint BiorXiv. 2017

- Participants: Thomas Dias-Alves, Michael Blum and Julien Mairal
- Partners: UGA - CNRS
- Contact: Julien Mairal
- Publication: Loter: A software package to infer local ancestry for a wide range of species
- URL: https://github.com/bcm-uga/Loter

6.4. SPAMS

SPArse Modeling Software
KEYWORDS: Signal processing - Machine learning
FUNCTIONAL DESCRIPTION: SPAMS is an open-source software package for sparse estimation

NEWS OF THE YEAR: The version 2.6.1 of the software package is now compatible with Python v3, R v3, comes with pre-compiled Matlab packages, and is now available on the conda and PyPi package managers.

- Participants: Ghislain Durif and Julien Mairal
- Contact: Julien Mairal
- URL: http://spams-devel.gforge.inria.fr/

6.5. LVO

Learning Video Object Segmentation with Visual Memory
KEYWORD: Video analysis
FUNCTIONAL DESCRIPTION: This is a public implementation of the method described in the following paper: Learning Video Object Segmentation with Visual Memory [ICCV 2017] (https://hal.archives-ouvertes.fr/hal-01511145v2/document).

This paper addresses the task of segmenting moving objects in unconstrained videos. We introduce a novel two-stream neural network with an explicit memory module to achieve this. The two streams of the network encode spatial and temporal features in a video sequence respectively, while the memory module captures the evolution of objects over time. The module to build a "visual memory" in video, i.e., a joint representation of all the video frames, is realized with a convolutional recurrent unit learned from a small number of training video sequences. Given a video frame as input, our approach assigns each pixel an object or background label based on the learned spatio-temporal features as well as the "visual memory" specific to the video, acquired automatically without any manually-annotated frames. The visual memory is implemented with convolutional gated recurrent units, which allows to propagate spatial information over time. We evaluate our method extensively on two benchmarks, DAVIS and Freiburg-Berkeley motion segmentation datasets, and show state-of-the-art results. For example, our approach outperforms the top method on the DAVIS dataset by nearly 6

- Participants: Karteek Alahari, Cordelia Schmid and Pavel Tokmakov
- Contact: Pavel Tokmakov
- Publication: Learning Video Object Segmentation with Visual Memory
- URL: http://lear.inrialpes.fr/research/lvo/

6.6. SURREAL

Learning from Synthetic Humans
KEYWORDS: Synthetic human - Segmentation - Neural networks
FUNCTIONAL DESCRIPTION: The SURREAL dataset consisting of synthetic videos of humans, and models trained on this dataset are released in this package. The code for rendering synthetic images of people and for training models is also included in the release.

- Participants: Gül Varol Simsekli, Xavier Martin, Ivan Laptev and Cordelia Schmid
- Contact: Gül Varol Simsekli
- Publication: Learning from Synthetic Humans
- URL: http://www.di.ens.fr/willow/research/surreal/

6.7. attn2d

Pervasive Attention
KEYWORDS: NLP - Deep learning - Machine translation
SCIENTIFIC DESCRIPTION: Pervasive attention : 2D Convolutional Networks for Sequence-to-Sequence Prediction

FUNCTIONAL DESCRIPTION: An open source PyTorch implementation of the pervasive attention model described in: Maha Elbayad, Laurent Besacier, and Jakob Verbeek. 2018. Pervasive Attention: 2D Convolutional Networks for Sequence-to-Sequence Prediction. In Proceedings of the 22nd Conference on Computational Natural Language Learning (CoNLL 2018)

- Participants: Maha Elbayad and Jakob Verbeek
- Contact: Maha Elbayad
- Publication: Pervasive Attention: 2D Convolutional Neural Networks for Sequence-to-Sequence Prediction
- URL: https://github.com/elbayadm/attn2d

7. New Results

7.1. Visual Recognition in Images and Videos

7.1.1. Actor and Observer: Joint Modeling of First and Third-Person Videos

Participants: Gunnar Sigurdsson [CMU], Abhinav Gupta [CMU], Cordelia Schmid, Ali Farhadi [AI2, Univ. Washington], Karteek Alahari.

Several theories in cognitive neuroscience suggest that when people interact with the world, or simulate interactions, they do so from a first-person egocentric perspective, and seamlessly transfer knowledge between third-person (observer) and first-person (actor). Despite this, learning such models for human action recognition has not been achievable due to the lack of data. Our work in [33] takes a step in this direction, with the introduction of Charades-Ego, a large-scale dataset of paired first-person and third-person videos, involving 112 people, with 4000 paired videos. This enables learning the link between the two, actor and observer perspectives. Thereby, we address one of the biggest bottlenecks facing egocentric vision research, providing a link from first-person to the abundant third-person data on the web. We use this data to learn a joint representation of first and third-person videos, with only weak supervision, and show its effectiveness for transferring knowledge from the third-person to the first-person domain.

7.1.2. Learning to Segment Moving Objects

Participants: Pavel Tokmakov, Cordelia Schmid, Karteek Alahari.

We study the problem of segmenting moving objects in unconstrained videos [14]. Given a video, the task is to segment all the objects that exhibit independent motion in at least one frame. We formulate this as a learning problem and design our framework with three cues: (i) independent object motion between a pair of frames, which complements object recognition, (ii) object appearance, which helps to correct errors in motion estimation, and (iii) temporal consistency, which imposes additional constraints on the segmentation. The framework is a two-stream neural network with an explicit memory module. The two streams encode appearance and motion cues in a video sequence respectively , while the memory module captures the evolution of objects over time, exploiting the temporal consistency. The motion stream is a convolutional neural network trained on synthetic videos to segment independently moving objects in the optical flow field. The module to build a visual memory in video, i.e., a joint representation of all the video frames, is realized with a convolutional recurrent unit learned from a small number of training video sequences. For every pixel in a frame of a test video, our approach assigns an object or background label based on the learned spatio-temporal features as well as the 'visual memory' specific to the video. We evaluate our method extensively on three benchmarks, DAVIS, Freiburg-Berkeley motion seg-mentation dataset and SegTrack. In addition, we provide an extensive ablation study to investigate both the choice of the training data and the influence of each component in the proposed framework. An overview of our model is shown in Figure 1.

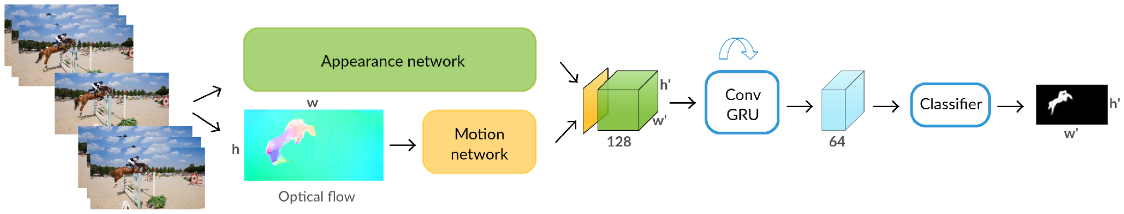


Figure 1. Overview of our segmentation approach [14]. Each video frame is processed by the appearance (green) and the motion (yellow) networks to produce an intermediate two-stream representation. The ConvGRU module combines this with the learned visual memory to compute the final segmentation result. The width (w') and height (h') of the feature map and the output are $w/8$ and $h/8$ respectively.

7.1.3. Unsupervised Learning of Artistic Styles with Archetypal Style Analysis

Participants: Daan Wynen, Cordelia Schmid, Julien Mairal.

In [36], we introduce an unsupervised learning approach to automatically discover, summarize, and manipulate artistic styles from large collections of paintings. Our method (summarized in Figure 2) is based on archetypal analysis, which is an unsupervised learning technique akin to sparse coding with a geometric interpretation. When applied to neural style representations from a collection of artworks, it learns a dictionary of archetypal styles, which can be easily visualized. After training the model, the style of a new image, which is characterized by local statistics of deep visual features, is approximated by a sparse convex combination of archetypes. This enables us to interpret which archetypal styles are present in the input image, and in which proportion. Finally, our approach allows us to manipulate the coefficients of the latent archetypal decomposition, and achieve various special effects such as style enhancement, transfer, and interpolation between multiple archetypes.

7.1.4. Learning from Web Videos for Event Classification

Participants: Nicolas Chesneau, Karteek Alahari, Cordelia Schmid.

Traditional approaches for classifying event videos rely on a manually curated training dataset. While this paradigm has achieved excellent results on benchmarks such as TrecVid multimedia event detection (MED) challenge datasets, it is restricted by the effort involved in careful annotation. Recent approaches have attempted to address the need for annotation by automatically extracting images from the web, or generating queries to retrieve videos. In the former case, they fail to exploit additional cues provided by video data, while in the latter, they still require some manual annotation to generate relevant queries. We take an alternate approach in [4], leveraging the synergy between visual video data and the associated textual metadata, to learn event classifiers without manually annotating any videos. Specifically, we first collect a video dataset with queries constructed automatically from textual description of events, prune irrelevant videos with text and video data, and then learn the corresponding event classifiers. We evaluate this approach in the challenging setting where no manually annotated training set is available, i.e., EK0 in the TrecVid challenge, and show state-of-the-art results on MED 2011 and 2013 datasets.

7.1.5. How good is my GAN?

Participants: Konstantin Shmelkov, Cordelia Schmid, Karteek Alahari.

Generative adversarial networks (GANs) are one of the most popular methods for generating images today. While impressive results have been validated by visual inspection, a number of quantitative criteria have emerged only recently. We argue here that the existing ones are insufficient and need to be in adequation with the task at hand. In [32] introduce two measures based on image classification—GAN-train and GAN-test (illustrated in Figure 3), which approximate the recall (diversity) and precision (quality of the image) of GANs

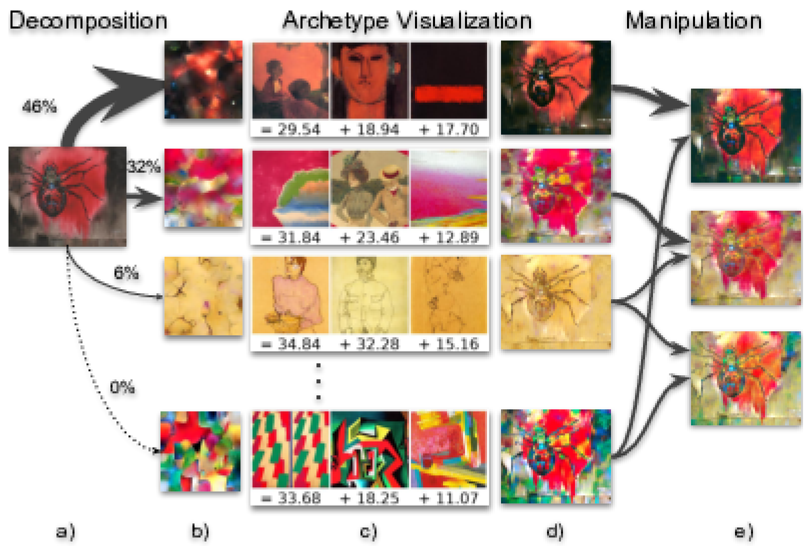


Figure 2. Using deep archetypal style analysis, we can represent the style of an artwork (a) as a convex combination of archetypes. The archetypes can be visualized as synthesized textures (b), as a convex combination of artworks (c) or, when analyzing a specific artwork, as stylized versions of that artwork itself (d). Free recombination of the archetypal styles then allows for novel stylizations of the input.

respectively. We evaluate a number of recent GAN approaches based on these two measures and demonstrate a clear difference in performance. Furthermore, we observe that the increasing difficulty of the dataset, from CIFAR10 over CIFAR100 to ImageNet, shows an inverse correlation with the quality of the GANs, as clearly evident from our measures.

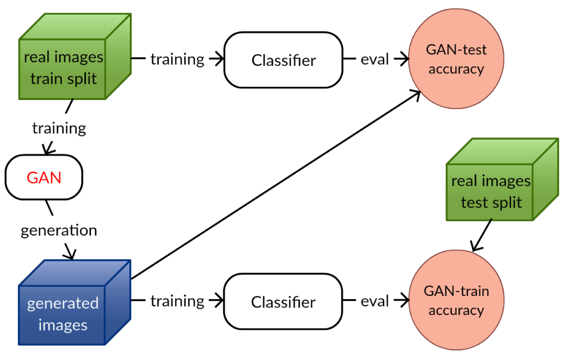


Figure 3. Illustration of GAN-train and GAN-test. GAN-train learns a classifier on GAN generated images and measures the performance on real test images. This evaluates the diversity and realism of GAN images. GAN-test learns a classifier on real images and evaluates it on GAN images. This measures how realistic GAN images are.

7.1.6. Modeling Visual Context is Key to Augmenting Object Detection Datasets

Participants: Nikita Dvornik, Julien Mairal, Cordelia Schmid.

Performing data augmentation for learning deep neural networks is well known to be important for training visual recognition systems. By artificially increasing the number of training examples, it helps reducing overfitting and improves generalization. For object detection, classical approaches for data augmentation consist of generating images obtained by basic geometrical transformations and color changes of original training images. In [23], we go one step further and leverage segmentation annotations to increase the number of object instances present on training data. For this approach to be successful, we show that modeling appropriately the visual context surrounding objects is crucial to place them in the right environment. Otherwise, we show that the previous strategy actually hurts. Clear difference between the two approaches can is presented in Figure 4. With our context model, we achieve significant mean average precision improvements when few labeled examples are available on the VOC'12 benchmark.

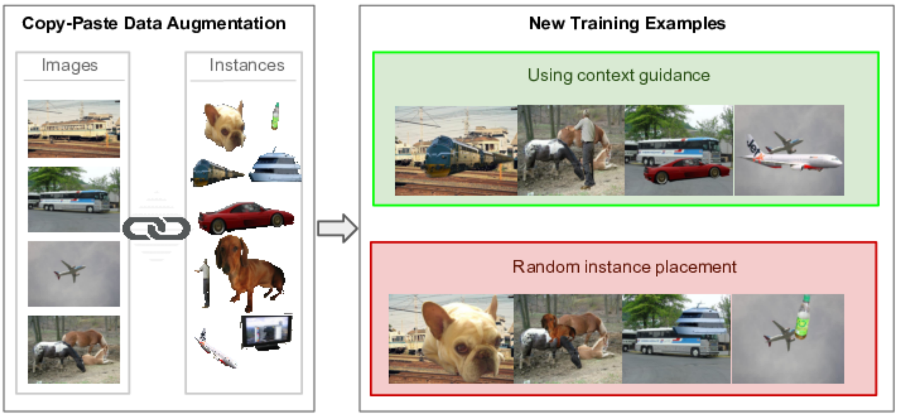


Figure 4. ***Examples of data-augmented training examples produced by our approach.*** *Images and objects are taken from the VOC'12 dataset that contains segmentation annotations. We compare the output obtained by pasting the objects with our context model vs. those obtained with random placements. Even though the results are not perfectly photorealistic and display blending artefacts, the visual context surrounding objects is more often correct with the explicit context model.*

7.1.7. On the Importance of Visual Context for Data Augmentation in Scene Understanding

Participants: Nikita Dvornik, Julien Mairal, Cordelia Schmid.

Performing data augmentation for learning deep neural networks is known to be important for training visual recognition systems. By artificially increasing the number of training examples, it helps reducing overfitting and improves generalization. While simple image transformations such as changing color intensity or adding random noise can already improve predictive performance in most vision tasks, larger gains can be obtained by leveraging task-specific prior knowledge. In [42], we consider object detection and semantic segmentation and augment the training images by blending objects in existing scenes, using instance segmentation annotations. We observe that randomly pasting objects on images hurts the performance, unless the object is placed in the right context. To resolve this issue, we propose an explicit context model by using a convolutional neural network, which predicts whether an image region is suitable for placing a given object or not. In our experiments, we show that by using copy-paste data augmentation with context guidance we are able to improve detection and segmentation on the PASCAL VOC12 and COCO datasets, with significant gains when few labeled examples are available. The way to augment for different tasks and annotations is presented in Figure 5. We also show that the method is not limited to datasets that come with expensive pixel-wise instance annotations and can be used when only bounding box annotations are available, by employing weakly-supervised learning for instance masks approximation.

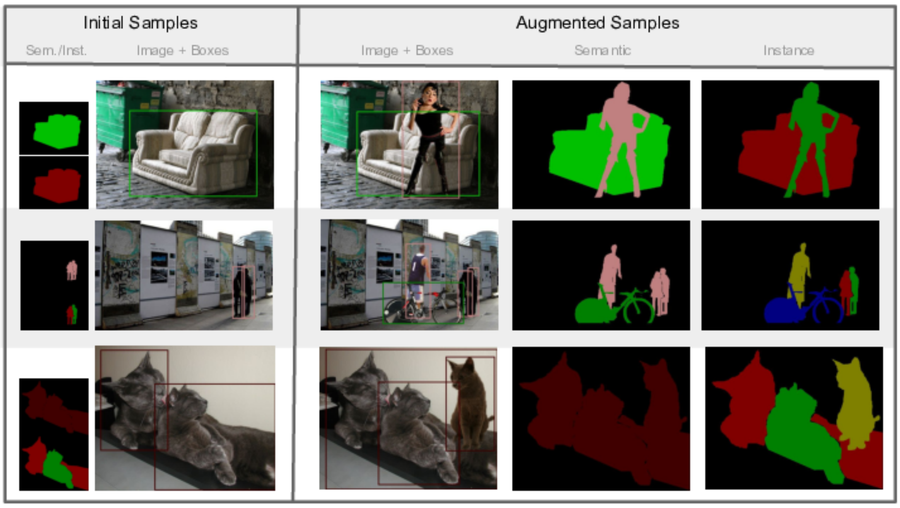


*Figure 5. **Data augmentation for different types of annotations.** The first column contains samples from the training dataset with corresponding semantic/instance segmentation and bounding box annotations. Columns 2-4 present the result of applying context-driven augmentation to the initial sample with corresponding annotations.*

7.1.8. Predicting future instance segmentation by forecasting convolutional features

Participants: Pauline Luc, Camille Couprie [Facebook AI Research], Yann Lecun [Facebook AI Research], Jakob Verbeek.

Anticipating future events is an important prerequisite towards intelligent behavior. Video forecasting has been studied as a proxy task towards this goal. Recent work has shown that to predict semantic segmentation of future frames, forecasting at the semantic level is more effective than forecasting RGB frames and then segmenting these. In [28], we consider the more challenging problem of future instance segmentation, which additionally segments out individual objects. To deal with a varying number of output labels per image, we develop a predictive model in the space of fixed-sized convolutional features of the Mask R-CNN instance segmentation model. We apply the "detection head" of Mask R-CNN on the predicted features to produce the instance segmentation of future frames. Experiments show that this approach significantly improves over strong baselines based on optical flow and repurposed instance segmentation architectures. We show an overview of the proposed method in Figure 6.

7.1.9. Joint Future Semantic and Instance Segmentation Prediction

Participants: Camille Couprie [Facebook AI Research], Pauline Luc, Jakob Verbeek.

The ability to predict what will happen next from observing the past is a key component of intelligence. Methods that forecast future frames were recently introduced towards better machine intelligence. However, predicting directly in the image color space seems an overly complex task, and predicting higher level representations using semantic or instance segmentation approaches were shown to be more accurate. In

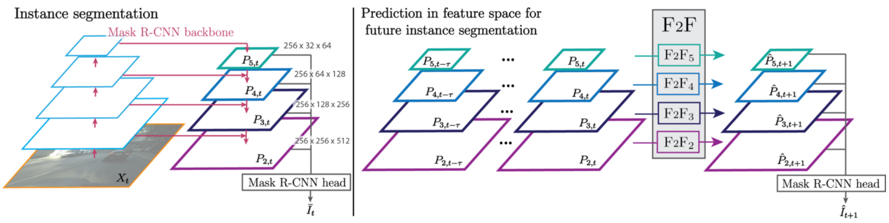


Figure 6. For future instance segmentation, we extract a pyramid of features from frames $t - \tau$ to t, and use them to predict the pyramid features for frame $t + 1$. We learn separate feature-to-feature prediction models for each level of the pyramid. The predicted features are then given as input to a downstream network to produce future instance segmentation.

[20], we introduce a novel prediction approach that encodes instance and semantic segmentation information in a single representation based on distance maps. Our graph-based modeling of the instance segmentation prediction problem allows us to obtain temporal tracks of the objects as an optimal solution to a watershed algorithm. Our experimental results on the Cityscapes dataset present state-of-the-art semantic segmentation predictions, and instance segmentation results outperforming a strong baseline based on optical flow. We show an overview of the proposed method in Figure 7.

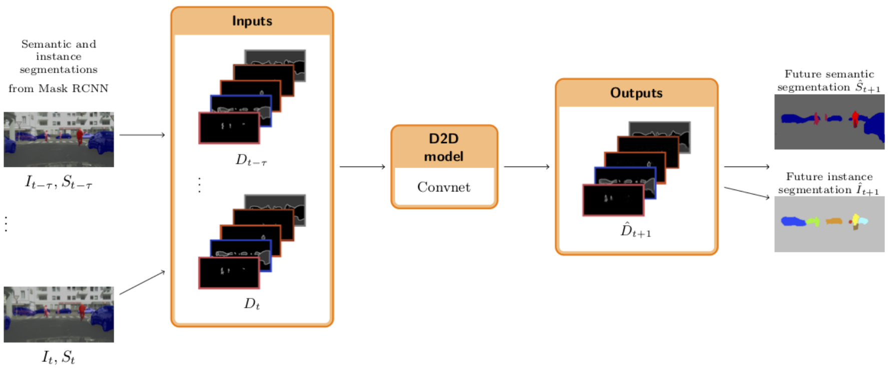


Figure 7. Our representation enables both future semantic and instance segmentation prediction. It relies on distance maps from the different objects contours: For each channel of an input segmentation, corresponding to a specific class, the segmentation is decomposed into zeros for background, ones for objects and high values for contours. Then a convnet is trained to predict the future representation. Taking its argmax lets us recover the future semantic segmentation, and computing a watershed from it leads to the future instance segmentation.

7.1.10. Depth-based Hand Pose Estimation: Methods, Data, and Challenges

Participants: James S. Supancic [UC Irvine], Grégory Rogez, Yi Yang [Baidu Research], Jamie Shotton [Microsoft Research], Deva Ramanan [Carnegie Mellon University].

Hand pose estimation has matured rapidly in recent years. The introduction of commodity depth sensors and a multitude of practical applications have spurred new advances. In [13], we provide an extensive analysis of the state-of-the-art, focusing on hand pose estimation from a single depth frame. We summarize important conclusions here: (1) Pose estimation appears roughly solved for scenes with isolated hands. However, methods still struggle to analyze cluttered scenes where hands may be interacting with nearby objects and surfaces. To spur further progress we introduce a challenging new dataset with diverse, cluttered scenes. (2) Many methods evaluate themselves with disparate criteria , making comparisons difficult. We define a consistent evaluation criteria, rigorously motivated by human experiments. (3) We introduce a simple nearest-neighbor baseline that outperforms most existing systems (see results in Fig. 8). This implies that most systems do not generalize beyond their training sets. This also reinforces the under-appreciated point that training data is as important as the model itself. We conclude with directions for future progress.

Figure 8. We evaluate a broad collection of hand pose estimation algorithms on different training and testsets under consistent criteria. Test sets which contained limited variety, in pose and range, or which lacked complex backgrounds were notably easier. To aid our analysis, we introduce a simple 3D exemplar (nearest-neighbor) baseline that both detects and estimates pose suprisingly well, outperforming most existing systems. We show the best-matching detection window in (middle) and the best-matching exemplar in (bottom). We use our baseline to rank dataset difficulty, compare algorithms, and show the importance of training set design.

7.1.11. Image-based Synthesis for Deep 3D Human Pose Estimation

Participants: Grégory Rogez, Cordelia Schmid.

In [11], we address the problem of 3D human pose estimation in the wild. A significant challenge is the lack of training data, i.e., 2D images of humans annotated with 3D poses. Such data is necessary to train state-of-the-art CNN architectures. Here, we propose a solution to generate a large set of photorealistic synthetic images of humans with 3D pose annotations. We introduce an image-based synthesis engine that artificially augments a dataset of real images with 2D human pose annotations using 3D Motion Capture (MoCap) data. Given a candidate 3D pose our algorithm selects for each joint an image whose 2D pose locally matches the projected 3D pose. The selected images are then combined to generate a new synthetic image by stitching local image patches in a kinematically constrained manner. See examples in Figure 9. The resulting images are used to train an end-to-end CNN for full-body 3D pose estimation. We cluster the training data into a large number of pose classes and tackle pose estimation as a K-way classification problem. Such an approach is viable only with large training sets such as ours. Our method outperforms the state of the art in terms of 3D pose estimation in controlled environments (Human3.6M) and shows promising results for in-the-wild images (LSP). This demonstrates that CNNs trained on artificial images generalize well to real images. Compared to data generated from more classical rendering engines, our synthetic images do not require any domain adaptation or fine-tuning stage.

Figure 9. Given a candidate 3D pose, our algorithm selects for each joint an image whose annotated 2D pose locally matches the projected 3D pose. The selected images are then combined to generate a new synthetic image by stitching local image patches in a kinematically constrained manner. We show 6 examples corresponding to the same 3D pose observed from 6 different camera viewpoints.

7.1.12. LCR-Net++: Multi-person 2D and 3D Pose Detection in Natural Images

Participants: Grégory Rogez, Philippe Weinzaepfel [Naver Labs Europe], Cordelia Schmid.

In [12], we propose an end-to-end architecture for joint 2D and 3D human pose estimation in natural images. Key to our approach is the generation and scoring of a number of pose proposals per image, which allows us to predict 2D and 3D pose of multiple people simultaneously. See example in Figure 10. Hence, our approach does not require an approximate localization of the humans for initialization. Our architecture, named LCR-Net, contains 3 main components: 1) the pose proposal generator that suggests potential poses at different locations in the image; 2) a classifier that scores the different pose proposals ; and 3) a regressor that refines pose proposals both in 2D and 3D. All three stages share the convolutional feature layers and are trained jointly. The final pose estimation is obtained by integrating over neighboring pose hypotheses , which is shown to improve over a standard non maximum suppression algorithm. Our approach significantly outperforms the state of the art in 3D pose estimation on Human3.6M, a controlled environment. Moreover, it shows promising results on real images for both single and multi-person subsets of the MPII 2D pose benchmark and demonstrates satisfying 3D pose results even for multi-person images.

7.1.13. Link and code: Fast indexing with graphs and compact regression codes

Participants: Matthijs Douze [Facebook AI Research], Alexandre Sablayrolles, Hervé Jégou [Facebook AI Research].

Similarity search approaches based on graph walks have recently attained outstanding speed-accuracy trade-offs, taking aside the memory requirements. In [21], we revisit these approaches by considering, additionally, the memory constraint required to index billions of images on a single server. This leads us to propose a method based both on graph traversal and compact representations. We encode the indexed vectors using quantization and exploit the graph structure to refine the similarity estimation, see Figure 11. In essence, our method takes the best of these two worlds: the search strategy is based on nested graphs, thereby providing high precision with a relatively small set of comparisons. At the same time it offers a significant memory compression. As a

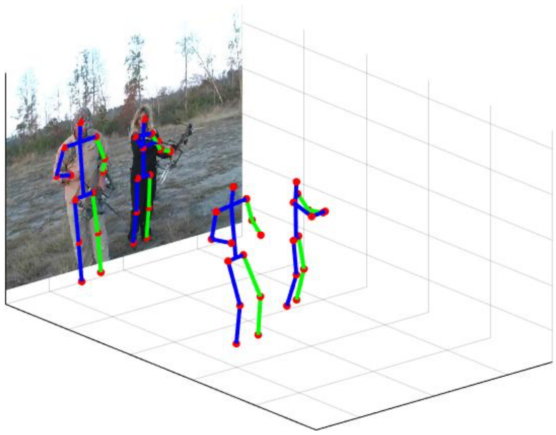


Figure 10. Examples of joint 2D-3D pose detections in a natural image. Even in case of occlusion or truncation, we estimate the joint locations by reasoning in term of full-body 2D-3D poses.

result, our approach outperforms the state of the art on operating points considering 64–128 bytes per vector, as demonstrated by our results on two billion-scale public benchmarks.

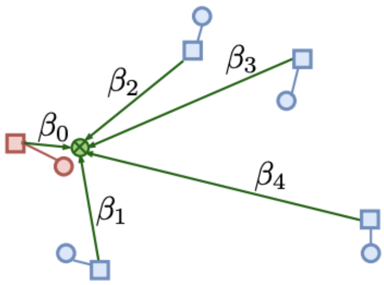


Figure 11. Illustration of our approach: we adopt a graph traversal strategy that maintains a connectivity between all database points. We further improve the estimate by regressing each database vector from its encoded neighbors.

7.1.14. Sparse weakly supervised models for object localization in road environment

Participants: Valentina Zadrija [Univ. Zagreb], Josip Krapac [Univ. Zagreb], Sinisa Segvic [Univ. Zagreb], Jakob Verbeek.

In [16] we propose a novel weakly supervised object localization method based on Fisher-embedding of low-level features (CNN, SIFT), and model sparsity at the component level. Fisher-embedding provides an interesting alternative to raw low-level features, since it allows fast and accurate scoring of image subwindows with a model trained on entire images. Model sparsity reduces overfitting and enables fast evaluation. We also propose two new techniques for improving performance when our method is combined with nonlinear normalizations of the aggregated Fisher representation of the image. These techniques are i) intra-component metric normalization and ii) first-order approximation to the score of a normalized image representation. We evaluate our weakly supervised localization method on real traffic scenes acquired from driver's perspective. The method dramatically improves the localization AP over the dense non-normalized Fisher vector baseline (16 percentage points for zebra crossings, 21 percentage points for traffic signs) and leads to a huge gain in execution speed (91× for zebra crossings, 74× for traffic signs). See Figure 12 for several example outputs.

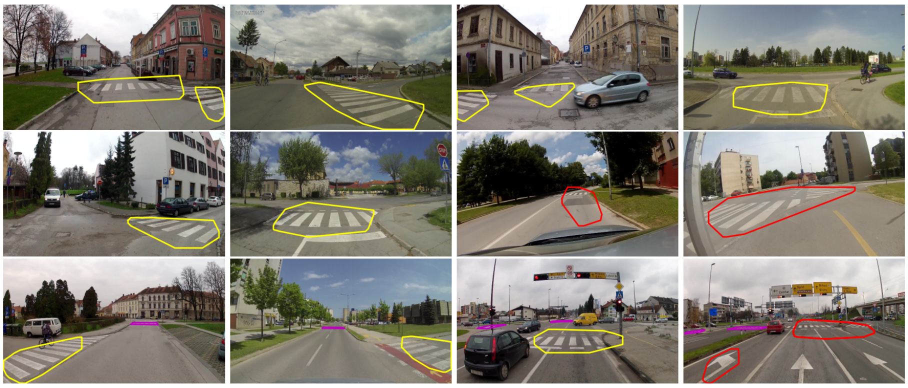


Figure 12. Localization results on test images: correct localization polygons (yellow), false postive responses (red), and ground-truth polygons for false negatives (magenta).

7.1.15. Scene Coordinate Regression with Angle-Based Reprojection Loss for Camera Relocalization

Participants: Xiaotian Li [Aalto Univ.], Juha Ylioinas [Aalto Univ.], Jakob Verbeek, Juho Kannala [Univ. Oulu].

Image-based camera relocalization is an important problem in computer vision and robotics. Recent works utilize convolutional neural networks (CNNs) to regress for pixels in a query image their corresponding 3D world coordinates in the scene. The final pose is then solved via a RANSAC-based optimization scheme using the predicted coordinates, see Figure 13. Usually, the CNN is trained with ground truth scene coordinates, but it has also been shown that the network can discover 3D scene geometry automatically by minimizing single-view reprojection loss. However, due to the deficiencies of reprojection loss, the network needs to be carefully initialized. In [27], we present a new angle-based reprojection loss which resolves the issues of the original reprojection loss. With this new loss function, the network can be trained without careful initialization, and the system achieves more accurate results. The new loss also enables us to utilize available multi-view constraints, which further improve performance.

7.1.16. FeaStNet: Feature-Steered Graph Convolutions for 3D Shape Analysis

Participants: Nitika Verma, Edmond Boyer [Inria, MORPHEO], Jakob Verbeek.

Convolutional neural networks (CNNs) have massively impacted visual recognition in 2D images, and are now ubiquitous in state-of-the-art approaches. While CNNs naturally extend to other domains, such as audio and video, where data is also organized in rectangular grids, they do not easily generalize to other types of data such as 3D shape meshes, social network graphs or molecular graphs. In our recent paper [35], we propose a novel graph-convolutional network architecture to handle such data. The architecture builds on a generic formulation that relaxes the 1-to-1 correspondence between filter weights and data elements around the center of the convolution, see Figure 14 for an illustration. The main novelty of our architecture is that the shape of the filter is a function of the features in the previous network layer, which is learned as an integral

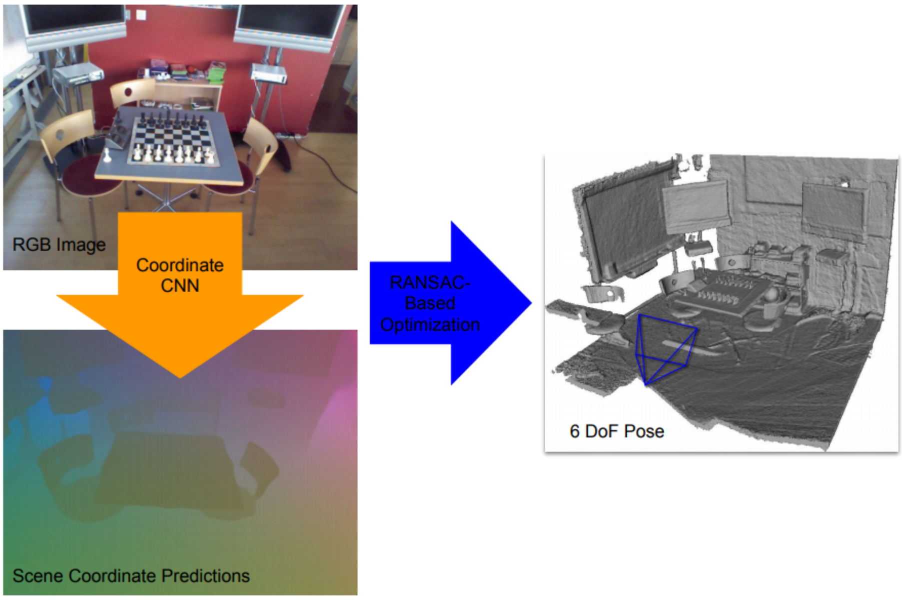


Figure 13. Localization pipeline. In this two-stage pipeline, a coordinate CNN first produces scene coordinate predictions from an RGB image, and then the predicted correspondences are fed into a RANSAC-based solver to determine the camera pose.

part of the neural network. Experimental evaluations on digit recognition and 3D shape correspondence yield state-of-the-art results, significantly improving over previous work for shape correspondence.

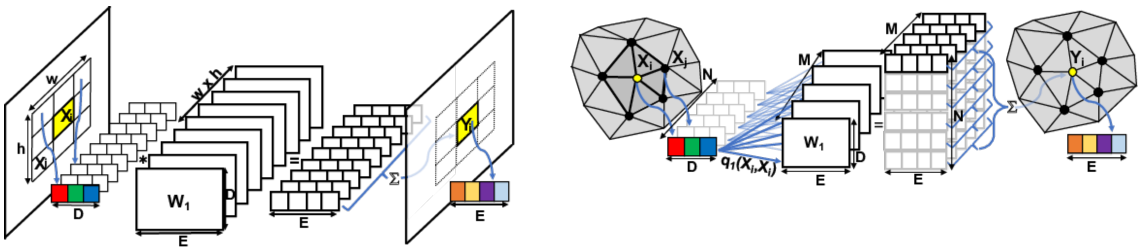


Figure 14. Left: Illustration of a standard CNN, representing the parameters as a set of $M = w \times h$ weight matrices, each of size $D \times E$. Each weight matrix is associated with a single relative position in the input patch. Right: Our graph convolutional network, where each node in the input patch is associated in a soft manner to each of the M weight matrices based on its features using the weight $q_m(\mathbf{x}_i, \mathbf{x}_j)$.

7.2. Statistical Machine Learning

7.2.1. Modulated Policy Hierarchies

Participants: Alexander Pashevich, Danijar Hafner [Google Brain], James Davidson [Vernalis (R&D) Ltd.], Rahul Sukthankar [Google], Cordelia Schmid.

Solving tasks with sparse rewards is a main challenge in reinforcement learning. While hierarchical controllers are an intuitive approach to this problem, current methods often require manual reward shaping, alternating training phases, or manually defined sub tasks. In [45], we introduce modulated policy hierarchies (MPH), that can learn end-to-end to solve tasks from sparse rewards. To achieve this, we study different modulation signals and exploration for hierarchical controllers. Specifically, we find that communicating via bit-vectors is more efficient than selecting one out of multiple skills, as it enables mixing between them (see Figure 15). To facilitate exploration, MPH uses its different time scales for temporally extended intrinsic motivation at each level of the hierarchy. We evaluate MPH on the robotics tasks of pushing and sparse block stacking, where it outperforms recent baselines.

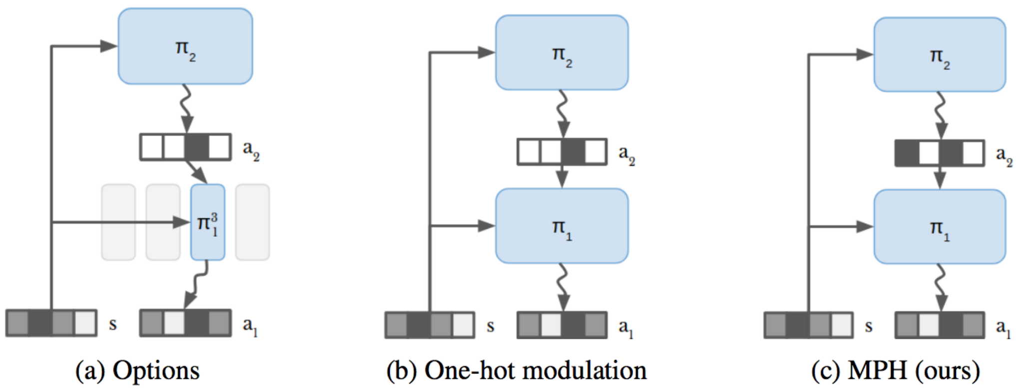


Figure 15. Overview of hierarchical policies. (a) The options agent selects between separate skill networks using a categorical master policy. (b) The one-hot agent combines the skills into a single network and is modulated by a 1-hot signal. (c) Our modulated policy hierarchy sends a binary vector, allowing for richer communication and mixing of skills.

7.2.2. Group Invariance, Stability to Deformations, and Complexity of Deep Convolutional Representations

Participants: Alberto Bietti, Julien Mairal.

The success of deep convolutional architectures is often attributed in part to their ability to learn multiscale and invariant representations of natural signals. However, a precise study of these properties and how they affect learning guarantees is still missing. In [38], we consider deep convolutional representations of signals; we study their invariance to translations and to more general groups of transformations, their stability to the action of diffeomorphisms, and their ability to preserve signal information. This analysis is carried by introducing a multilayer kernel based on convolutional kernel networks and by studying the geometry induced by the kernel mapping. We then characterize the corresponding reproducing kernel Hilbert space (RKHS), showing that it contains a large class of convolutional neural networks with homogeneous activation functions. This analysis allows us to separate data representation from learning, and to provide a canonical measure of model complexity, the RKHS norm, which controls both stability and generalization of any learned model. In addition to models in the constructed RKHS, our stability analysis also applies to convolutional networks with generic activations such as rectified linear units, and we discuss its relationship with recent generalization bounds based on spectral norms.

7.2.3. A Contextual Bandit Bake-off

Participants: Alberto Bietti, Alekh Agarwal [Microsoft Research], John Langford [Microsoft Research].

Contextual bandit algorithms are essential for solving many real-world interactive machine learning problems. Despite multiple recent successes on statistically and computationally efficient methods, the practical behavior of these algorithms is still poorly understood. In [37], we leverage the availability of large numbers of supervised learning datasets to compare and empirically optimize contextual bandit algorithms, focusing on practical methods that learn by relying on optimization oracles from supervised learning. We find that a recent method using optimism under uncertainty works the best overall. A surprisingly close second is a simple greedy baseline that only explores implicitly through the diversity of contexts, followed by a variant of Online Cover which tends to be more conservative but robust to problem specification by design. Along the way, we also evaluate and improve several internal components of contextual bandit algorithm design. Overall, this is a thorough study and review of contextual bandit methodology.

7.2.4. Learning Disentangled Representations with Reference-Based Variational Autoencoders

Participants: Adria Ruiz, Oriol Martinez [Universitat Pompeu Fabra, Barcelona], Xavier Binefa [Universitat Pompeu Fabra, Barcelona], Jakob Verbeek.

Learning disentangled representations from visual data, where different high-level generative factors are independently encoded, is of importance for many computer vision tasks. Supervised approaches, however, require a significant annotation effort in order to label the factors of interest in a training set. To alleviate the annotation cost, in [47] we introduce a learning setting which we refer to as "reference-based disentangling". Given a pool of unlabelled images, the goal is to learn a representation where a set of target factors are disentangled from others. The only supervision comes from an auxiliary "reference set" that contains images where the factors of interest are constant. See Fig. 16 for illustrative examples. In order to address this problem, we propose reference-based variational autoencoders, a novel deep generative model designed to exploit the weak supervisory signal provided by the reference set. During training, we use the variational inference framework where adversarial learning is used to minimize the objective function. By addressing tasks such as feature learning, conditional image generation or attribute transfer, we validate the ability of the proposed model to learn disentangled representations from minimal supervision.

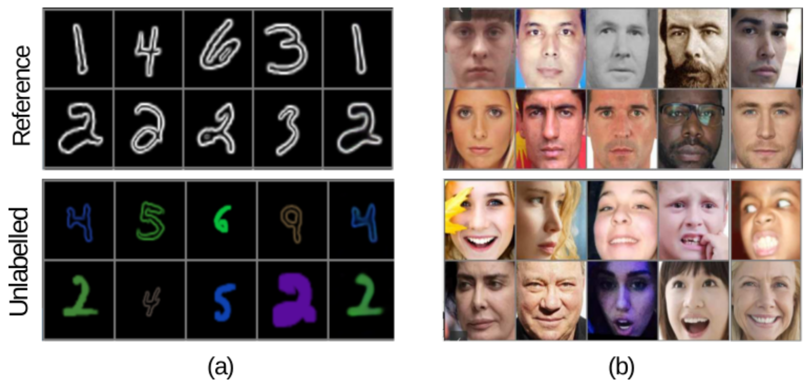


Figure 16. Illustration of different reference-based disentangling problems. (a) Disentangling style from digits. The reference distribution is composed by numbers with a fixed style (b) Disentangling factors of variations related with facial expressions. Reference images correspond to neutral faces. Note that pairing information between unlabelled and reference images is not available during training.

7.2.5. On Regularization and Robustness of Deep Neural Networks

Participants: Alberto Bietti, Grégoire Mialon, Julien Mairal.

For many supervised learning tasks, deep neural networks are known to work well when large amounts of annotated data are available. Yet, Despite their success, deep neural networks suffer from several drawbacks: they lack robustness to small changes of input data known as "adversarial examples" and training them with small amounts of annotated data is challenging. In [39], we study the connection between regularization and robustness of deep neural networks by viewing them as elements of a reproducing kernel Hilbert space (RKHS) of functions and by regularizing them using the RKHS norm. Even though this norm cannot be computed, we consider various approximations based on upper and lower bounds. These approximations lead to new strategies for regularization, but also to existing ones such as spectral norm penalties or constraints, gradient penalties, or adversarial training. Besides, the kernel framework allows us to obtain margin-based bounds on adversarial generalization. We show that our new algorithms lead to empirical benefits for learning on small datasets and learning adversarially robust models. We also discuss implications of our regularization framework for learning implicit generative models.

7.2.6. Mixed batches and symmetric discriminators for GAN training

Participants: Thomas Lucas, Corentin Tallec [Inria, TAU], Jakob Verbeek, Yann Ollivier [Facebook AI Research].

Generative adversarial networks (GANs) are powerful generative models based on providing feedback to a generative network via a discriminator network. However, the discriminator usually assesses individual samples. This prevents the discriminator from accessing global distributional statistics of generated samples, and often leads to *mode dropping*: the generator models only part of the target distribution. In [29] we propose to feed the discriminator with *mixed batches* of true and fake samples, and train it to predict the ratio of true samples in the batch. The latter score does not depend on the order of samples in a batch. Rather than learning this invariance, we introduce a generic permutation-invariant discriminator architecture, which is illustrated in Figure 17. This architecture is provably a universal approximator of all symmetric functions. Experimentally, our approach reduces mode collapse in GANs on two synthetic datasets, and obtains good results on the CIFAR10 and CelebA datasets, both qualitatively and quantitatively.

7.2.7. Auxiliary Guided Autoregressive Variational Autoencoders

Participants: Thomas Lucas, Jakob Verbeek.

Generative modeling of high-dimensional data is a key problem in machine learning. Successful approaches include latent variable models and autoregressive models. The complementary strengths of these approaches, to model global and local image statistics respectively, suggest hybrid models combining the strengths of both. Our contribution in [30] is to train such hybrid models using an auxiliary loss function that controls which information is captured by the latent variables and what is left to the autoregressive decoder, as illustrated in Figure 18. In contrast, prior work on such hybrid models needed to limit the capacity of the autoregressive decoder to prevent degenerate models that ignore the latent variables and only rely on autoregressive modeling. Our approach results in models with meaningful latent variable representations, and which rely on powerful autoregressive decoders to model image details. Our model generates qualitatively convincing samples, and yields state-of-the-art quantitative results.

7.2.8. End-to-End Incremental Learning

Participants: Francisco Castro [Univ. Malaga], Manuel Marin-Jimenez [Univ. Cordoba], Nicolas Guil [Univ. Malaga], Cordelia Schmid, Karteek Alahari.

Although deep learning approaches have stood out in recent years due to their state-of-the-art results, they continue to suffer from catastrophic forgetting, a dramatic decrease in overall performance when training with new classes added incrementally. This is due to current neural network architectures requiring the entire dataset, consisting of all the samples from the old as well as the new classes, to update the model—a requirement that becomes easily unsustainable as the number of classes grows. We address this issue with

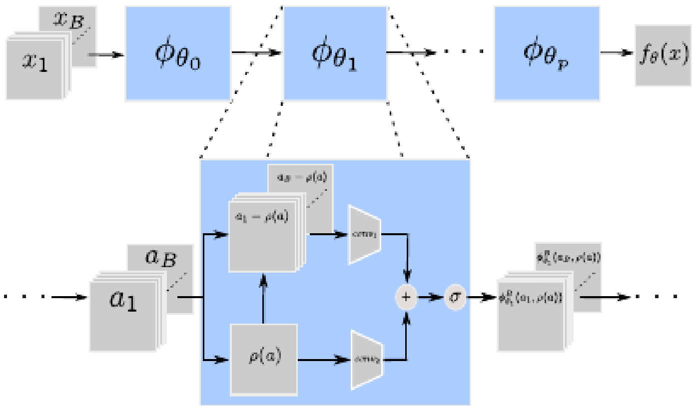


Figure 17. Graphical representation of our discriminator architecture. Each convolutional layer of an otherwise classical CNN architecture is modified to include permutation invariant batch statistics, denoted $\rho(x)$. This is repeated at every layer so that the network gradually builds up more complex statistics.

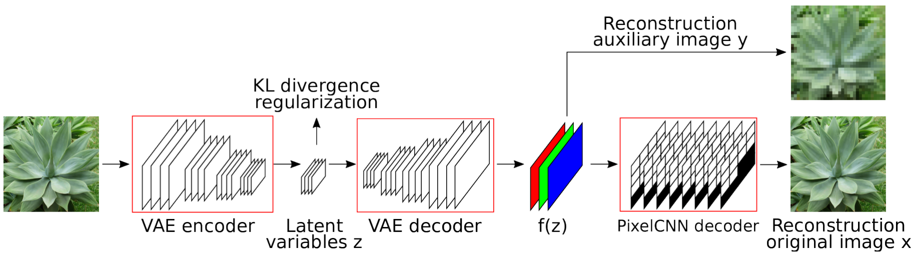


Figure 18. Schematic illustration of our auxiliary guided autoregressive variational autoencoder (AGAVE). An input image is encoded into a latent representation and decoded back into an image. This first reconstruction is guided by an auxiliary maximum likelihood loss and regularized with a Kullback-Liebler divergence. An autoregressive model is then conditionned on the auxiliary reconstruction and also trained with maximum likelihood.

our approach [17] to learn deep neural networks incrementally, using new data and only a small exemplar set corresponding to samples from the old classes. This is based on a loss composed of a distillation measure to retain the knowledge acquired from the old classes, and a cross-entropy loss to learn the new classes. Our incremental training is achieved while keeping the entire framework end-to-end, i.e., learning the data representation and the classifier jointly, unlike recent methods with no such guarantees. We evaluate our method extensively on the CIFAR-100 and ImageNet (ILSVRC 2012) image classification datasets, and show state-of-the-art performance.

7.3. Large-scale Optimization for Machine Learning

7.3.1. Stochastic Subsampling for Factorizing Huge Matrices

Participants: Julien Mairal, Arthur Mensch [Inria, Parietal], Gael Varoquaux [Inria, Parietal], Bertrand Thirion [Inria, Parietal].

In [10], we present a matrix-factorization algorithm that scales to input matrices with both huge number of rows and columns. Learned factors may be sparse or dense and/or non-negative, which makes our algorithm suitable for dictionary learning, sparse component analysis, and non-negative matrix factorization. Our algorithm streams matrix columns while subsampling them to iteratively learn the matrix factors. At each iteration, the row dimension of a new sample is reduced by subsampling, resulting in lower time complexity compared to a simple streaming algorithm. Our method comes with convergence guarantees to reach a stationary point of the matrix-factorization problem. We demonstrate its efficiency on massive functional Magnetic Resonance Imaging data (2 TB), and on patches extracted from hyperspectral images (103 GB). For both problems, which involve different penalties on rows and columns, we obtain significant speed-ups compared to state-of-the-art algorithms. The main principle of the method is illustrated in Figure 19.

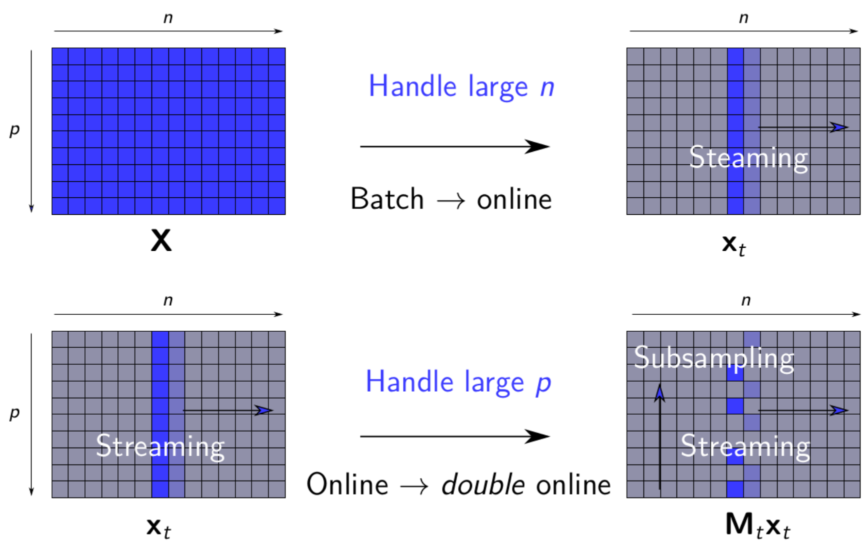


Figure 19. Illustration of the matrix factorization algorithm, which streams columns in one dimension while subsampling them.

7.3.2. An Inexact Variable Metric Proximal Point Algorithm for Generic Quasi-Newton Acceleration

Participants: Hongzhou Lin, Julien Mairal, Zaid Harchaoui [Univ. Washington].

In [43], we propose a generic approach to accelerate gradient-based optimization algorithms with quasi-Newton principles. The proposed scheme, called QuickeNing, can be applied to incremental first-order methods such as stochastic variance-reduced gradient (SVRG) or incremental surrogate optimization (MISO). It is also compatible with composite objectives, meaning that it has the ability to provide exactly sparse solutions when the objective involves a sparsity-inducing regularization. QuickeNing relies on limited-memory BFGS rules, making it appropriate for solving high-dimensional optimization problems. Besides, it enjoys a worst-case linear convergence rate for strongly convex problems. We present experimental results where QuickeNing gives significant improvements over competing methods for solving large-scale high-dimensional machine learning problems, see Figure 20 for example.

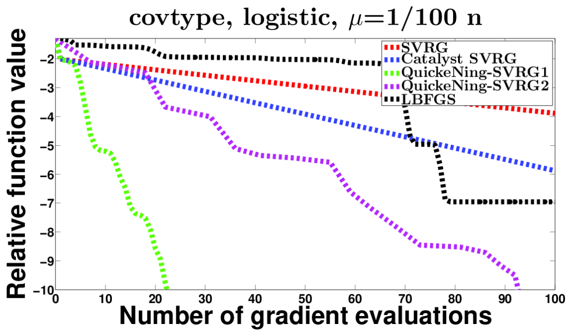


Figure 20. An illustration of the minimization of logistic regression. Significant improvement is observed by applying QuickeNing.

7.3.3. Catalyst Acceleration for First-order Convex Optimization: from Theory to Practice

Participants: Hongzhou Lin, Julien Mairal, Zaid Harchaoui [Univ. Washington].

In [9], we introduce a generic scheme for accelerating gradient-based optimization methods in the sense of Nesterov. The approach, called Catalyst, builds upon the inexact accelerated proximal point algorithm for minimizing a convex objective function, and consists of approximately solving a sequence of well-chosen auxiliary problems, leading to faster convergence. One of the key to achieve acceleration in theory and in practice is to solve these sub-problems with appropriate accuracy by using the right stopping criterion and the right warm-start strategy. In this work, we give practical guidelines to use Catalyst and present a comprehensive theoretical analysis of its global complexity. We show that Catalyst applies to a large class of algorithms, including gradient descent, block coordinate descent, incremental algorithms such as SAG, SAGA, SDCA, SVRG, Finito/MISO, and their proximal variants. For all of these methods, we provide acceleration and explicit support for non-strongly convex objectives. We conclude with extensive experiments showing that acceleration is useful in practice, especially for ill-conditioned problems.

7.3.4. Catalyst Acceleration for Gradient-Based Non-Convex Optimization

Participants: Courtney Paquette [Univ. Washington], Hongzhou Lin, Dmitriy Drusvyatskiy [Univ. Washington], Julien Mairal, Zaid Harchaoui [Univ. Washington].

In [31], we introduce a generic scheme to solve nonconvex optimization problems using gradient-based algorithms originally designed for minimizing convex functions. When the objective is convex, the proposed approach enjoys the same properties as the Catalyst approach of Lin et al, 2015. When the objective is nonconvex, it achieves the best known convergence rate to stationary points for first-order methods. Specifically, the proposed algorithm does not require knowledge about the convexity of the objective; yet, it obtains an overall worst-case efficiency of $O(\epsilon^{-2})$ and, if the function is convex, the complexity reduces to the near-optimal rate $O(\epsilon^{-2/3})$. We conclude the paper by showing promising experimental results obtained by applying the proposed approach to SVRG and SAGA for sparse matrix factorization and for learning neural networks (see Figure 21).

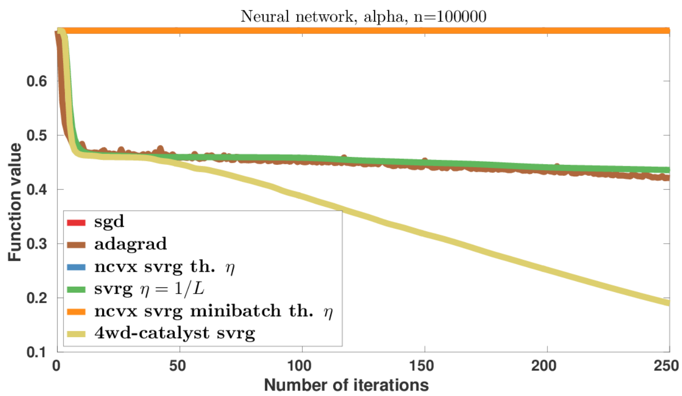


Figure 21. Comparison of different algorithms for the minimization of a two-layer neural network. Applying our method provides a clear acceleration in terms of function value.

7.4. Pluri-disciplinary Research

7.4.1. Biological Sequence Modeling with Convolutional Kernel Networks

Participants: Dexiong Chen, Laurent Jacob [CNRS, LBBE Laboratory], Julien Mairal.

The growing number of annotated biological sequences available makes it possible to learn genotype-phenotype relationships from data with increasingly high accuracy. When large quan- tities of labeled samples are available for training a model, convolutional neural networks can be used to predict the phenotype of unannotated sequences with good accuracy. Unfortunately, their performance with medium- or small-scale datasets is mitigated, which requires inventing new data-efficient approaches. In [40], we introduce a hybrid approach between convolutional neural networks and kernel methods to model biological sequences. Our method 22 enjoys the ability of convolutional neural networks to learn data representations that are adapted to a specific task, while the kernel point of view yields algorithms that perform significantly better when the amount of training data is small. We illustrate these advantages for transcription factor binding prediction and protein homology detection, and we demonstrate that our model is also simple to interpret, which is crucial for discovering predictive motifs in sequences. The source code is freely available at https://gitlab.inria.fr/dchen/CKN-seq.

7.4.2. Token-level and sequence-level loss smoothing for RNN language models

Participants: Maha Elbayad, Laurent Besacier [LIG], Jakob Verbeek.

In [25] we investigate the limitations of the maximum likelihood estimation (MLE) used when training recurrent neural network language models. First, the MLE treats all sentences that do not match the ground truth as equally poor, ignoring the structure of the output space. Second, it suffers from "exposure bias": during training tokens are predicted given ground-truth sequences, while at test time prediction is conditioned

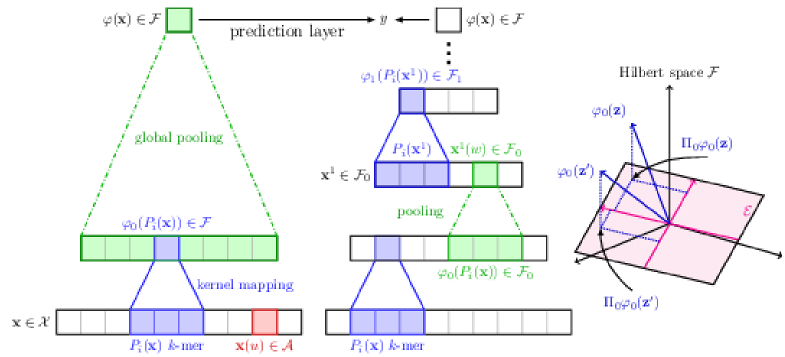


Figure 22. Construction of single-layer (left) and multilayer (middle) CKN-seq and the approximation of one layer (right). For a single-layer model, each k-mer $P_i(\mathbf{x})$ is mapped to $\varphi_0(P_i(\mathbf{x}))$ in $\mathcal{F}$ and projected to $\Pi\varphi_0(P_i(\mathbf{x}))$ parametrized by $\psi_0(P_i(\mathbf{x}))$. Then, the final finite-dimensional sequence is obtained by the global pooling, $\psi(\mathbf{x}) = \frac{1}{m}\sum_{i=0}^{m}\psi_0(P_i(\mathbf{x}))$. The multilayer construction is similar, but relies on intermediate maps, obtained by local pooling.

on generated output sequences. To overcome these limitations we build upon the recent reward augmented maximum likelihood approach i.e., sequence-level smoothing that encourages the model to predict sentences close to the ground truth according to a given performance metric. We extend this approach to token-level loss smoothing, and propose improvements to the sequence-level smoothing approach. Our experiments on two different tasks, image captioning (see Fig. 23) and machine translation, show that token-level and sequence-level loss smoothing are complementary, and significantly improve results.

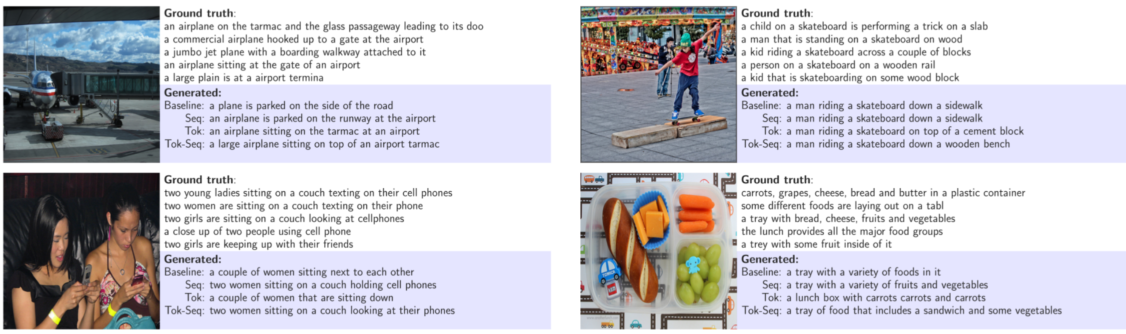


Figure 23. Examples of generated captions with the baseline MLE and our models with attention.

7.4.3. *Pervasive Attention: 2D Convolutional Neural Networks for Sequence-to-Sequence Prediction*

Participants: Maha Elbayad, Laurent Besacier [LIG], Jakob Verbeek.

Current state-of-the-art machine translation systems are based on encoder-decoder architectures, that first encode the input sequence, and then generate an output sequence based on the input encoding. Both are interfaced with an attention mechanism that recombines a fixed encoding of the source tokens based on the decoder state. In [24], we propose an alternative approach which instead relies on a single 2D convolutional neural network across both sequences as illustrated in Figure 24. Each layer of our network re-codes source tokens on the basis of the output sequence produced so far. Attention-like properties are therefore pervasive throughout the network. Our model yields excellent results, outperforming state-of-the-art encoder-decoder systems, while being conceptually simpler and having fewer parameters.

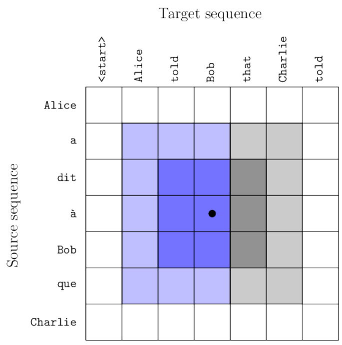


Figure 24. Convolutional layers in our model use masked 3×3 filters so that features are only computed from previous output symbols. Illustration of the receptive fields after one (dark blue) and two layers (light blue), together with the masked part of the field of view of a normal 3×3 filter (gray)

7.4.4. Probabilistic Count Matrix Factorization for Single Cell Expression Data Analysis

Participant: Ghislain Durif.

The development of high-throughput biology technologies now allows the investigation of the genome-wide diversity of transcription in single cells. This diversity has shown two faces: the expression dynamics (gene to gene variability) can be quantified more accurately, thanks to the measurement of lowly-expressed genes. Second, the cell-to-cell variability is high, with a low proportion of cells expressing the same gene at the same time/level. Those emerging patterns appear to be very challenging from the statistical point of view, especially to represent and to provide a summarized view of single-cell expression data. PCA is one of the most powerful framework to provide a suitable representation of high dimensional datasets, by searching for latent directions catching the most variability in the data. Unfortunately, classical PCA is based on Euclidean distances and projections that work poorly in presence of over-dispersed counts that show drop-out events (zero-inflation) like single-cell expression data. In [22], we propose a probabilistic Count Matrix Factorization (pCMF) approach for single-cell expression data analysis, that relies on a sparse Gamma-Poisson factor model. This hierarchical model is inferred using a variational EM algorithm. We show how this probabilistic framework induces a geometry that is suitable for single-cell data visualization, and produces a compression of the data that is very powerful for clustering purposes. Our method is competed to other standard representation methods like t-SNE, and we illustrate its performance for the representation of zero-inflated over-dispersed count data. We also illustrate our work with results on a publicly available data set, being single-cell expression profile of neural stem cells. Our work is implemented in the pCMF R-package.

7.4.5. Extracting Universal Representations of Cognition across Brain-Imaging Studies

Participants: Arthur Mensch [Inria, Parietal], Julien Mairal, Bertrand Thirion [Inria, Parietal], Gael Varoquaux [Inria, Parietal].

We show in [44] how to extract shared brain representations that predict mental processes across many cognitive neuroimaging studies. Focused cognitive-neuroimaging experiments study precise mental processes with carefully-designed cognitive paradigms; however the cost of imaging limits their statistical power. On the other hand, large-scale databasing efforts increase considerably the sample sizes, but cannot ask precise cognitive questions. To address this tension, we develop new methods that turn the heterogeneous cognitive information held in different task-fMRI studies into common-universal-cognitive models. Our approach does not assume any prior knowledge of the commonalities shared by the studies in the corpus; those are inferred during model training. The method uses deep-learning techniques to extract representations - task-optimized networks - that form a set of basis cognitive dimensions relevant to the psychological manipulations, as illustrated in Figure 25. In this sense, it forms a novel kind of functional atlas, optimized to capture mental state across many functional-imaging experiments. As it bridges information on the neural support of mental processes, this representation improves decoding performance for 80% of the 35 widely-different functional imaging studies that we consider. Our approach opens new ways of extracting information from brain maps, increasing statistical power even for focused cognitive neuroimaging studies, in particular for those with few subjects.

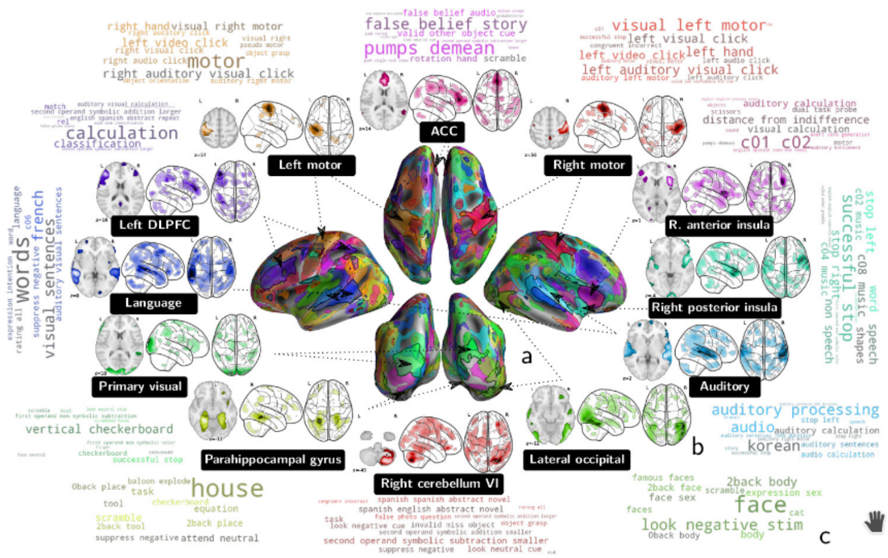


Figure 25. Visualization of some of task-optimized networks. Our approach allows to learn networks that are important for inter-subject decoding across studies. These networks, individually focal and collectively well spread across the cortex, are readily associated with the cognitive tasks that they contribute to predict. We display a selection of these networks, named with the salient anatomical brain region they recruit, along with a word-cloud representation of the stimuli whose likelihood increases with the network activation.

7.4.6. Loter: Inferring local ancestry for a wide range of species

Participants: Thomas Dias-Alves, Julien Mairal, Michael Blum [CNRS, TIMC Laboratory].

Admixture between populations provides opportunity to study biological adaptation and phenotypic variation. Admixture studies can rely on local ancestry inference for admixed individuals, which consists of computing at each locus the number of copies that originate from ancestral source populations, as illustrated in Figure 26. Existing software packages for local ancestry inference are tuned to provide accurate results on human data and recent admixture events. In [5], we introduce Loter, an open-source software package that does not require any biological parameter besides haplotype data in order to make local ancestry inference available for a wide range of species. Using simulations, we compare the performance of Loter to HAPMIX, LAMP-LD, and

RFMix. HAPMIX is the only software severely impacted by imperfect haplotype reconstruction. Loter is the less impacted software by increasing admixture time when considering simulated and admixed human genotypes. LAMP-LD and RFMIX are the most accurate method when admixture took place 20 generations ago or less; Loter accuracy is comparable or better than RFMix accuracy when admixture took place of 50 or more generations; and its accuracy is the largest when admixture is more ancient than 150 generations. For simulations of admixed Populus genotypes, Loter and LAMP-LD are robust to increasing admixture times by contrast to RFMix. When comparing length of reconstructed and true ancestry tracts, Loter and LAMP-LD provide results whose accuracy is again more robust than RFMix to increasing admixture times. We apply Loter to admixed Populus individuals and lengths of ancestry tracts indicate that admixture took place around 100 generations ago. The Loter software package and its source code are available at https://github.com/bcm-uga/Loter.

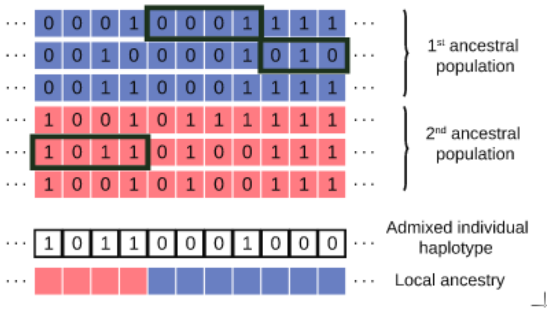


Figure 26. Graphical description of Local Ancestry Inference as implemented in the software Loter. Given a collection of parental haplotypes from the source populations depicted in blue and red, Loter assumes that an haplotype of an admixed individuals is modeled as a mosaic of existing parental haplotypes.

8. Bilateral Contracts and Grants with Industry

8.1. MSR-Inria joint lab: structured large-scale machine learning

Participants: Julien Mairal, Alberto Bietti.

Machine learning is now ubiquitous in industry, science, engineering, and personal life. While early successes were obtained by applying off-the-shelf techniques, there are two main challeges faced by machine learning in the “ big data ” era : structure and scale. The project proposes to explore three axes, from theoretical, algorithmic and practical perspectives: (1) large-scale convex optimization, (2) large-scale combinatorial optimization and (3) sequential decision making for structured data. The project involves two Inria sites and four MSR sites and started at the end of 2013. Alberto Bietti visited MSR New York in 2018.

8.2. Amazon

Participants: Grégory Rogez, Cordelia Schmid.

We received an Amazon Faculty Research Award in 2018. The objective is 3D human action recognition from monocular RGB videos. The idea is to extend our recent work on human 3D pose estimation to videos and to develop an approach for action recognition based on temporal pose based on appropriate 3D features.

8.3. Intel

Participants: Cordelia Schmid, Karteek Alahari.

The Intel Network on Intelligent Systems in Europe brings together leading researchers in robotics, computer vision, motor control, and machine learning. We are part of this network and have participated in the annual retreat in 2018. Funding will be provided on an annual basis, every year, as long as we are part of the network.

8.4. Facebook

Participants: Cordelia Schmid, Jakob Verbeek, Julien Mairal, Karteek Alahari, Pauline Luc, Alexandre Sablayrolles, Mathilde Caron.

The collaboration started in 2016. The topics include image retrieval with CNN based descriptors, weakly supervised object detection and semantic segmentation, and learning structured models for action recognition in videos. In 2016, Pauline Luc started her PhD funded by a CIFRE grant, jointly supervised by Jakob Verbeek (Inria) and Camille Couprie (Facebook AI Research). THOTH has been selected in 2016 as a recipient for the Facebook GPU Partnership program. In this context Facebook has donated two state-of-the-art servers with 8 GPUs. In 2017, Alexandre Sablayrolles started his CIFRE grant, jointly supervised by Cordelia Schmid, and Herve Jegou and Matthijs Douze at Facebook AI Research. In 2018, Mathilde Caron started as a CIFRE PhD student, jointly supervised by Julien Mairal, and Armand Joulin and Piotr Bojanowski at Facebook AI Research.

8.5. NAVER LABS Europe

Participants: Cordelia Schmid, Karteek Alahari, Julien Mairal, Jakob Verbeek, Vasileios Choutas, Nieves Crasto.

This collaboration started when NAVER LABS Europe was Xerox Research Centre Europe, and has been on-going since October 2009 with two co-supervised CIFRE scholarships (2009–2012, 2011-2014). Starting June 2014 we signed a third collaborative agreement for a duration of three years. The goal is to develop approaches for deep learning based image description and pose estimation in videos. Jakob Verbeek and Diane Larlus (XRCE) jointly supervise a PhD-level intern for a period of 6 months in 2016-2017. XRCE then became Naver in 2017. A one-year research contract on action recognition in videos started in Sep 2017. The approach developed by Vasileios Choutas implements pose-based motion features, which are shown to be complementary to state-of-the-art I3D features. Nieves Crasto's internship in 2018 was jointly supervised by Philippe Weinzaepfel (NAVER LABS), Karteek Alahari and Cordelia Schmid.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. DeCore (Deep Convolutional and Recurrent networks for image, speech, and text)

Participants: Jakob Verbeek, Maha Elbayad.

DeCore is a project-team funded by the Persyval Lab for 3.5 years (september 2016 - February 2020), coordinated by Jakob Verbeek. It unites experts from Grenoble's applied-math and computer science labs LJK, GIPSA-LAB and LIG in the areas of computer vision, machine learning, speech, natural language processing, and information retrieval. The purpose of DeCore is to stimulate collaborative interdisciplinary research on deep learning in the Grenoble area, which is likely to underpin future advances in machine perception (vision, speech, text) over the next decade. It provides funding for two full PhD students. Maha Elbayad is one of them, supervised by Jakob Verbeek and Laurant Besacier (LIG, UGA).

9.1.2. PEPS AMIES AuMalis POLLEN

Participant: Karteek Alahari.

This is a collaborative project with POLLEN, a startup in the Grenoble area, which develops POLLEN Metrology, a software editor specialized in signal processing, hybrid metrology and machine learning for the automatic processing of heterogeneous data. This funding supports a postdoc to accelerate the introduction of artificial intelligence, and in particular computer vision, techniques, into the manufacture of new generation of microprocessors. Karteek Alahari and Valerie Perrier (LJK, UGA) jointly supervise a postdoc as part of this collaboration.

9.2. National Initiatives

9.2.1. ANR Project Macaron

Participants: Julien Mairal, Zaid Harchaoui [Univ. Washington], Laurent Jacob [CNRS, LBBE Laboratory], Michael Blum [CNRS, TIMC Laboratory], Joseph Salmon [Telecom ParisTech], Mikita Dvornik, Thomas Dias-Alves, Daan Wynen.

The project MACARON is an endeavor to develop new mathematical and algorithmic tools for making machine learning more scalable. Our ultimate goal is to use data for solving scientific problems and automatically converting data into scientific knowledge by using machine learning techniques. Therefore, our project has two different axes, a methodological one, and an applied one driven by explicit problems. The methodological axis addresses the limitations of current machine learning for simultaneously dealing with large-scale data and huge models. The second axis addresses open scientific problems in bioinformatics, computer vision, image processing, and neuroscience, where a massive amount of data is currently produced, and where huge-dimensional models yield similar computational problems.

This is a 4 years and half project, funded by ANR under the program "Jeunes chercheurs, jeunes chercheuses", which started in October 2014. The principal investigator is Julien Mairal.

9.2.2. ANR Project DeepInFrance

Participants: Jakob Verbeek, Adria Ruiz Ovejero.

DeepInFrance (Machine learning with deep neural networks) project also aims at bringing together complementary machine learning, computer vision and machine listening research groups working on deep learning with GPUs in order to provide the community with the knowledge, the visibility and the tools that brings France among the key players in deep learning. The long-term vision of Deep in France is to open new frontiers and foster research towards algorithms capable of discovering sense in data in an automatic manner, a stepping stone before the more ambitious far-end goal of machine reasoning. The project partners are: INSA Rouen, Univ. Caen, Inria, UPMC, Aix-Marseille Univ., Univ. Nice Sophia Antipolis.

9.2.3. ANR Project AVENUE

Participant: Karteek Alahari.

This new ANR project (started in October 2018) aims to address the perception gap between human and artificial visual systems through a visual memory network for human-like interpretation of scenes. To this end, we address three scientific challenges. The first is to learn a network representation of image, video and text data collections, to leverage their inherent diverse cues. The second is to depart from supervised learning paradigms, without compromising on the performance. The third one is to perform inference with the learnt network, e.g., to estimate physical and functional properties of objects, or give cautionary advice for navigating a scene. The principal investigator is Karteek Alahari, and the project involves participants from CentraleSupelec and Ecole des Ponts in Paris.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. ERC Advanced grant Allegro

Participants: Cordelia Schmid, Pavel Tokmakov, Konstantin Shmelkov, Vladyslav Sydorov, Daan Wynen, Mikita Dvornik, Xavier Martin.

The ERC advanced grant ALLEGRO started in April 2013 and will end in April 2019. The aim of ALLEGRO is to automatically learn from large quantities of data with weak labels. A massive and ever growing amount of digital image and video content is available today. It often comes with additional information, such as text, audio or other meta-data, that forms a rather sparse and noisy, yet rich and diverse source of annotation, ideally suited to emerging weakly supervised and active machine learning technology. The ALLEGRO project will take visual recognition to the next level by using this largely untapped source of data to automatically learn visual models. We will develop approaches capable of autonomously exploring evolving data collections, selecting the relevant information, and determining the visual models most appropriate for different object, scene, and activity categories. An emphasis will be put on learning visual models from video, a particularly rich source of information, and on the representation of human activities, one of today's most challenging problems in computer vision.

9.3.1.2. ERC Starting grant Solaris

Participants: Julien Mairal, Ghislain Durif, Andrei Kulunchakov, Alberto Bietti, Dexiong Chen, Gregoire Mialon.

The project SOLARIS started in March 2017 for a duration of five years. The goal of the project is to set up methodological and theoretical foundations of deep learning models, in the context of large-scale data processing. The main applications of the tools developed in this project are for processing visual data, such as videos, but also structured data produced in experimental sciences, such as biological sequences.

The main paradigm used in the project is that of kernel methods and consist of building functional spaces where deep learning models live. By doing so, we want to derive theoretical properties of deep learning models that may explain their success, and also obtain new tools with better stability properties. Another work package of the project is focused on large-scale optimization, which is a key to obtain fast learning algorithms.

9.4. International Initiatives

9.4.1. Inria International Labs

Inria@EastCoast

Associate Team involved in the International Lab:

9.4.1.1. GAYA

Title: Semantic and Geometric Models for Video Interpretation

International Partner (Institution - Laboratory - Researcher):

Carnegie Mellon University (United States) - Robotics Institute - Deva Ramanan

Start year: 2016

See also: https://team.inria.fr/gaya/

We propose to form an associate team GAYA, with the primary goal of interpreting videos in terms of recognizing actions, understanding the human-human and human-object interactions. Despite several years of research, it is yet unclear what is an efficient and robust video representation to attack this challenge. In order to address this, GAYA will focus on building semantic models, wherein we learn the video feature representation with limited supervision, and also geometric models, where we study the geometric properties of object shapes to better recognize them. The team consists of researchers from two Inria project-teams (LEAR and WILLOW) and a US university (Carnegie Mellon University [CMU]). It will allow the three teams to effectively combine their respective strengths in areas such as inference and machine learning approaches for vision tasks, feature representation, large-scale learning, geometric reasoning. The main expected outcomes of this collaboration are: effective learnt representations of video content, new machine learning algorithms for handling minimally annotated data, large-scale public datasets for benchmarking, theoretical analysis of objects shapes and contours.

9.4.2. Inria International Partners

9.4.2.1. Informal International Partners

- **MPI Tübingen:** Cordelia Schmid collaborates with Michael Black, a research director at MPI, starting in 2013. End of 2015 she was award a Humbolt research award funding a long-term research project with colleagues at MPI. She spent one month at MPI in April 2018. In 2018, the project resulted in the development of an approach for object interaction.
- **University of Washington:** Julien Mairal collaborates with Zaid Harchaoui, former member of the team, on the topic of large-scale optimization.

9.4.3. Participation in Other International Programs

- **Indo-French project EVEREST** with IIIT Hyderabad, India, funded by CEFIPRA (Centre Franco-Indien pour la Promotion de la Recherche Avancee). The aim of this project between Cordelia Schmid, Karteek Alahari and C. V. Jawahar (IIIT Hyderabad) is to enable the use of rich, complex models that are required to address the challenges of high-level computer vision. The work plan for the project will follow three directions. First, we will develop a learning framework that can handle weak annotations. Second, we will build formulations to solve the non-convex optimization problem resulting from the learning framework. Third, we will develop efficient and accurate energy minimization algorithms, in order to make the optimization computationally feasible.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

9.5.1.1. Internships

- Pia Bideau (PhD student, Univ. Massachusetts Amherst) was an intern in the team from Sep to Dec 2018.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- C. Schmid is one of the general chairs for ECCV 2020.

10.1.1.2. Member of the Organizing Committees

- Several permanent members of the team co-organized the international summer school PAISS 2018.
- J. Mairal is a member of the organizing committee for the international conference SIAM Imaging Science 2020.
- J. Mairal co-organized the Journées SMAI-MODE, which will took place in March 2018.
- J. Mairal is a co-organizer of the workshop OSL'19 at Les Houches.
- G. Rogez was one of the organizers of the CVPR workshop on Human Pose, Motion, Activities and Shape in 3D (3D HUMANS 2018).

- C. Schmid was one of the organizers of Workshop on Artificial Intelligence, Horizon Maths, Paris, 2018.

10.1.2. Scientific Events Selection

10.1.2.1. Member of the Conference Program Committees

- K. Alahari: Area chair for CVPR 2018, ICCV 2019.
- K. Alahari: Senior program committee member for IJCAI 2018, IJCAI 2019.
- J. Mairal: Area chair for ICML and NeurIPS 2018.
- C. Schmid: Area chair for NeurIPS 2018, ICML 2018, ECCV 2018, ICCV 2019.
- J. Verbeek: Area chair for ECCV 2018, ICCV 2019.

10.1.2.2. Reviewer

The permanent members of the team reviewed numerous papers for numerous international conferences in computer vision and machine learning, including CVPR, ECCV, NeurIPS, ICML.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- K. Alahari: Associate editor for Computer Vision and Image Understanding journal, since 2018.
- J. Mairal: Associate editor of the International Journal of Computer Vision, since 2015.
- J. Mairal: Associate editor of Journal of Mathematical Imaging and Vision, since 2015.
- J. Mairal: Associate editor of the SIAM Journal of Imaging Science, since 2018.
- C. Schmid: Editor in Chief of the International Journal of Computer Vision, since 2013.
- C. Schmid: Associate editor for Foundations and Trends in Computer Graphics and Vision, since 2005.
- J. Verbeek: Associate editor for Image and Vision Computing Journal, 2011-2018.
- J. Verbeek: Associate editor for International Journal on Computer Vision, since 2014.
- J. Verbeek: Associate editor for IEEE Transactions Pattern Analysis and Machine Intelligence, since 2018.

10.1.3.2. Reviewer - Reviewing Activities

The permanent members of the team reviewed numerous papers for numerous international journals in computer vision (IJCV, PAMI, CVIU), machine learning (JMLR, Machine Learning). Some of them also review for journals in optimization (SIAM Journal on Optimization, Mathematical Programming), image processing (SIAM Imaging Science).

10.1.4. Invited Talks

- K. Alahari: Invited talk at IISc Bangalore, India, Dec 2018.
- K. Alahari: Invited talk at Hyderabad AI Symposium, India, Dec 2018.
- K. Alahari: Invited seminar at Simon Fraser Univ., Vancouver, Canada, Nov 2018.
- K. Alahari: Invited seminar at POSTECH, Pohang, South Korea, Oct 2018.
- K. Alahari: COVIEW workshop, ACM Multimedia, Seoul, South Korea, Oct 2018.
- K. Alahari: Le Futur des Images workshop, IXXI, Lyon, Oct 2018.
- K. Alahari: Invited seminar at Universidad de Malaga, Apr 2018.
- K. Alahari: CVPR AC Meeting, Toronto, Feb 2018.
- A. Bietti: Séminaire de statistiques et machine learning, Telecom ParisTech.
- A. Bietti: Laplace reading group, ENS Paris.
- A. Bietti: Grenoble Optimization Days, LJK, UGA.

- G. Durif: Seminar at LJK, UGA, May 2018.
- G. Durif: Seminar at Institut de Recherche Mathematique Avancée (IRMA), Strasbourg Univ., June 2018.
- M. Elbayad: Seminar at NAVER LABS Europe, Grenoble, Sep 2018.
- P. Luc: Invited talk in the class “Deep Learning for Image Analysis” at Ecole des Mines.
- J. Mairal: Seminaire Parisien d’Optimisation, Paris.
- J. Mairal: Workshop on the Future of Random Projections, Paris.
- J. Mairal: talk in mini-symposium at ISMP, Bordeaux.
- J. Mairal: keynote in Theory of Deep Learning Workshop, ICML 2018, Stockholm.
- J. Mairal: keynote in CEFRL workshop, ECCV 2018, Munich.
- J. Mairal: invited speaker at AI and ML workshop, Telecom ParisTech.
- J. Mairal: Seminar, Gatsby Unit, UCL, London.
- J. Mairal: Seminar, College de France, Paris.
- J. Mairal: Seminar, University of Geneva.
- J. Mairal: Seminar, Université d’Avignon.
- J. Mairal: Seminar, Université de Marseille.
- G. Rogez: Invited speaker, Apple Seminar, Salt Lake City, June 2018.
- G. Rogez: Seminar, NAVER LABS Europe, Grenoble, December 2018.
- C. Schmid: Invited speaker at the Google Multimodal Machine Perception Workshop, San Francisco, October 2018.
- C. Schmid: Invited speaker, “What is optical flow for?” workshop at ECCV, September 2018.
- C. Schmid: Invited speaker, 3rd Intl. Workshop on Video Segmentation at ECCV, September 2018.
- C. Schmid: Keynote speaker at Deep Learning Conference, Rennes, September 2018.
- C. Schmid: Keynote speaker at ActivityNet workshop, in conjunction with CVPR, June 2018.
- C. Schmid: Invited talk at CVPR Good Citizen of CVPR event, in conjunction with CVPR, June 2018.
- C. Schmid: Keynote speaker at 3D Humans workkshop, in conjunction with CVPR, June 2018.
- C. Schmid: Invited speaker at Symposium on AI, Académie des sciences, Paris, February 2018.
- C. Schmid: Presentation at Prairie/industry meeting, Paris, December 2018.
- C. Schmid: Presentation at Google workshop on 3D Deep Learning, October 2018.
- C. Schmid: Seminar for AI residents, Google Mountain View, June 2018.
- C. Schmid: Seminar at MPI Tübingen, April 2018.
- C. Schmid: Seminars at Google Zürich (April), Mountain View and Paris (March).
- C. Schmid: Seminar at Leopoldina section meeting, Ulm, February 2018.
- J. Verbeek: Seminar Facebook AI Research, Paris, December 2018.
- J. Verbeek: Seminar NAVER LABS Europe, Grenoble, December 2018.
- J. Verbeek: Seminar Univ. Amsterdam, July 2018.
- J. Verbeek: Workshop Mathematics and Deep Learning at Univ. Aix-Marseille, November 2018.
- J. Verbeek: ARC6 Deep Learning & Deep Reinforcement Learning Workshop, Lyon.

10.1.5. Scientific Expertise

- J. Mairal: Reviewer for ERC (Consolidator and Starting).
- J. Mairal: Panel member for ANR.

- J. Mairal: Judge for the IBM Watson AI Xprize.
- G. Rogez: reviewer for ANR.

10.1.6. Research Administration

- J. Mairal: Participation in the setting up of the 3IA institute in Grenoble.
- C. Schmid: Member, board of directors of the Computer Vision Foundation (CVF), since 2016.
- C. Schmid: Member, PAMI-TC awards committee and the PAMI-TC executive commitee.
- J. Verbeek: Member, 2018 recruitment committee for an Assistant Professor position at Univ. Grenoble Alpes, Laboratoire Jean Kuntzmann (LJK).
- J. Verbeek: Member, steering committee MinaLogic, innovation cluster for digital technologies based in France's Auvergne-Rhône-Alpes region, since 2018.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Doctorat: K. Alahari, Lecturer at the CVIT summer school on machine learning, 4h eqTD, IIIT Hyderabad, India.

Doctorat: A. Bietti, Mini-course "Optimization for large-scale machine learning", 10h eqTD, conference SMAI-MODE, Autrans, France.

Doctorat: J. Mairal, Mini-course "Optimization for large-scale machine learning", 10h eqTD, conference SMAI-MODE, Autrans, France.

Doctorat: J. Mairal, Mini-course "Optimization for large-scale machine learning", 4.5h eqTD, conference Mascot-Num, Nantes, France.

Doctorat: J. Mairal, Lecturer at the YSU - ISTC Joint Summer School on Machine Learning, 12h eqTD, Yerevan, Armenia.

Doctorat: J. Mairal, Lecturer at the PAISS summer school, 2h eqTD, Grenoble, France.

Doctorat: C. Schmid, Course on action recognition, 2h eqTD, PRAIRIE Artifical Intelligence Summer School, July 2018, Grenoble, France.

Master: K. Alahari, Understanding Big Visual Data, 13.5h eqTD, M2, Grenoble INP, France.

Master: K. Alahari, Graphical Models Inference and Learning, 18h eqTD, M2, CentraleSupelec, Paris, France.

Master: K. Alahari, Introduction to computer vision, 9h eqTD, M1, ENS Paris, France.

Master: J. Mairal, Kernel methods for statistical learning, 15h eqTD, M2, Ecole Normale Supérieure, Cachan, France.

Master: C. Schmid, Object recognition and computer vision, 9h eqTD, M2, ENS Paris, France

Master: J. Verbeek, K. Alahari, C. Schmid, Machine Learning and Object Recognition, 27h eqTD, M2, Grenoble University, France

Master: J. Verbeek and J. Mairal, Advanced Learning Models, 27h eqTD, M2, UGA, Grenoble.

10.2.2. Supervision

PhD: N. Chesneau, Learning to Recognize Actions with Weak Supervision, Univ. Grenoble Alpes, Feb 2018, Karteek Alahari and Cordelia Schmid.

PhD: A. Mensch, Apprentissage de représentations en imagerie fonctionnelle, Univ. Paris-Saclay, Sep 2018, Gael Varoquaux, Bertrand Thirion and Julien Mairal.

PhD: P. Tokmakov, Apprentissage à partir du mouvement, Univ. Grenoble Alpes, Jun 2018, Karteek Alahari and Cordelia Schmid.

10.2.3. Juries

K. Alahari: Mostafa S. Ibrahim, November 2018, external examiner, Simon Fraser University.

K. Alahari: Ignacio Rocco Spremolla, 2018, member of “comité de suivi de thèse”, ENS Paris.

J. Mairal: Magda Gregorova, 2018, rapporteur, University of Geneva.

J. Mairal: Saeed Varasteh, December 2018, examinateur, Univ. Grenoble Alpes.

J. Mairal: Olga Permiakova, November 2018, member of “comité de suivi de thèse”, Univ. Grenoble Alpes.

J. Mairal: Vincent Prost, November 2018, member of “comité de suivi de thèse”, Univ. Paris Saclay.

G. Rogez: Francisco Castro, December 2018, jury member, Univ. Malaga.

C. Schmid: Jean-Baptise Alayrac, September 2018, examinateur, ENS Paris.

C. Schmid: Stéphane Lathuilière, May 2018, presidente, UGA.

C. Schmid: Gunnar Atli Sigurdsson, May 2018, thesis proposal, CMU.

C. Schmid: Christoph Lassner, April 2018, rapporteur, Universitaet Tuebingen.

C. Schmid: Maxime Oquab, January 2018, examinatuer, ENS Paris.

J. Verbeek: Riccardo Del Chiaro, 2018-2020, member of supervisory committee, Univ. Florence.

J. Verbeek: Fabien Baradel, 2017-2019, member of supervisory committee, INSA Lyon.

J. Verbeek: Mélanie Ducoffe, 2018, rapporteur, Univ. Côte d’Azur, Sophia-Antipolis.

10.3. Popularization

- P. Luc: Interviews with “Computer Vision News” and “Les Echos Start”.
- J. Mairal: intervention grand public sur l’IA lors d’un événement organisé par PWN (Professional Woman Network) à Lyon en octobre 2018.
- J. Mairal: conférence grand public sur l’IA avec Cédric Villani en juin 2018, mairie de Lyon.
- J. Mairal: participation à un débat public sur l’IA organisé par Mme la député Cendra Motin en avril 2018, Montalieu-Vercieu.

10.3.1. Internal or external Inria responsibilities

- C. Schmid: Member, “Comité scientifique", Inria Grenoble, since 2015.
- J. Verbeek: Scientific correspondent, national project calls, Inria Grenoble, since 2017.
- J. Verbeek: Member, Inria Grenoble working group on HPC - Big Data - Machine learning, since 2018.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] N. CHESNEAU.*Learning to Recognize Actions with Weak Supervision*, Université Grenoble Alpes, February 2018, https://tel.archives-ouvertes.fr/tel-01893147

[2] G. CHÉRON.*Structured modeling and recognition of human actions in video*, Ecole normale supérieure - ENS PARIS, December 2018, https://hal.inria.fr/tel-01975247

[3] P. TOKMAKOV.*Learning from motion*, Université Grenoble Alpes, June 2018, https://tel.archives-ouvertes.fr/tel-01908817

Articles in International Peer-Reviewed Journal

[4] N. CHESNEAU, K. ALAHARI, C. SCHMID.*Learning from Web Videos for Event Classification*, in "IEEE Transactions on Circuits and Systems for Video Technology", 2018 [*DOI :* 10.1109/TCSVT.2017.2764624], https://hal.inria.fr/hal-01618400

[5] T. DIAS-ALVES, J. MAIRAL, M. BLUM.*Loter: A software package to infer local ancestry for a wide range of species*, in "Molecular Biology and Evolution", June 2018, vol. 35, n^{o} 9, p. 2318 - 2326 [*DOI :* 10.1093/MOLBEV/MSY126], https://hal.inria.fr/hal-01630228

[6] G. DURIF, L. MODOLO, J. MICHAELSSON, J. E. MOLD, S. LAMBERT-LACROIX, F. PICARD.*High Dimensional Classification with combined Adaptive Sparse PLS and Logistic Regression*, in "Bioinformatics", February 2018, vol. 34, n^{o} 3, p. 485-493, https://arxiv.org/abs/1502.05933 [*DOI :* 10.1093/BIOINFORMATICS/BTX571], https://hal.archives-ouvertes.fr/hal-01587360

[7] B. HAM, M. CHO, C. SCHMID, J. PONCE.*Proposal Flow: Semantic Correspondences from Object Proposals*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2018, vol. 40, n^{o} 7, p. 1711-1725 [*DOI :* 10.1109/TPAMI.2017.2724510], https://hal.inria.fr/hal-01644132

[8] G. HU, X. PENG, Y. YANG, T. HOSPEDALES, J. VERBEEK.*Frankenstein: Learning Deep Face Representations using Small Data*, in "IEEE Transactions on Image Processing", January 2018, vol. 27, n^{o} 1, p. 293-303, https://arxiv.org/abs/1603.06470 [*DOI :* 10.1109/TIP.2017.2756450], https://hal.inria.fr/hal-01306168

[9] H. LIN, J. MAIRAL, Z. HARCHAOUI.*Catalyst Acceleration for First-order Convex Optimization: from Theory to Practice*, in "Journal of Machine Learning Research", April 2018, vol. 18, n^{o} 212, p. 1-54, http://jmlr.org/papers/volume18/17-748/17-748.pdf, https://hal.inria.fr/hal-01664934

[10] A. MENSCH, J. MAIRAL, B. THIRION, G. VAROQUAUX.*Stochastic Subsampling for Factorizing Huge Matrices*, in "IEEE Transactions on Signal Processing", January 2018, vol. 66, n^{o} 1, p. 113-128, https://arxiv.org/abs/1701.05363 [*DOI :* 10.1109/TSP.2017.2752697], https://hal.archives-ouvertes.fr/hal-01431618

[11] G. ROGEZ, C. SCHMID.*Image-based Synthesis for Deep 3D Human Pose Estimation*, in "International Journal of Computer Vision", September 2018, vol. 126, n^{o} 9, p. 993–1008, https://arxiv.org/abs/1802.04216 [*DOI :* 10.1007/S11263-018-1071-9], https://hal.inria.fr/hal-01717188

[12] G. ROGEZ, P. WEINZAEPFEL, C. SCHMID.*LCR-Net++: Multi-person 2D and 3D Pose Detection in Natural Images*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2019, https://hal.archives-ouvertes.fr/hal-01961189

[13] J. S. SUPANCIC, G. ROGEZ, Y. YANG, J. SHOTTON, D. RAMANAN.*Depth-based hand pose estimation: methods, data, and challenges*, in "International Journal of Computer Vision", November 2018, vol. 126, n^{o} 11, p. 1180–1198 [*DOI :* 10.1007/S11263-018-1081-7], https://hal.inria.fr/hal-01759416

[14] P. TOKMAKOV, C. SCHMID, K. ALAHARI.*Learning to Segment Moving Objects*, in "International Journal of Computer Vision", 2018, https://arxiv.org/abs/1712.01127 , https://hal.archives-ouvertes.fr/hal-01653720

[15] G. VAROL, I. LAPTEV, C. SCHMID.*Long-term Temporal Convolutions for Action Recognition*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", June 2018, vol. 40, n^{o} 6, p. 1510-1517, https://arxiv.org/abs/1604.04494 [*DOI :* 10.1109/TPAMI.2017.2712608], https://hal.inria.fr/hal-01241518

[16] V. ZADRIJA, J. KRAPAC, S. ŠEGVIĆ, J. VERBEEK.*Sparse weakly supervised models for object localization in road environment*, in "Computer Vision and Image Understanding", 2018, p. 1-13 [*DOI :* 10.1016/J.CVIU.2018.10.004], https://hal.inria.fr/hal-01900418

International Conferences with Proceedings

[17] F. M. CASTRO, M. J. MARÍN-JIMÉNEZ, N. GUIL, C. SCHMID, K. ALAHARI.*End-to-End Incremental Learning*, in "ECCV 2018 - European Conference on Computer Vision", Munich, Germany, V. FERRARI, M. HEBERT, C. SMINCHISESCU, Y. WEISS (editors), Lecture Notes in Computer Science, September 2018, https://arxiv.org/abs/1807.09536 , https://hal.inria.fr/hal-01849366

[18] V. CHOUTAS, P. WEINZAEPFEL, J. REVAUD, C. SCHMID.*PoTion: Pose MoTion Representation for Action Recognition*, in "CVPR 2018 - IEEE Conference on Computer Vision and Pattern Recognition", Salt Lake City, United States, IEEE, June 2018, p. 1-10, https://hal.inria.fr/hal-01764222

[19] G. CHÉRON, J.-B. ALAYRAC, I. LAPTEV, C. SCHMID.*A flexible model for training action localization with varying levels of supervision*, in "NIPS 2018 - 32nd Conference on Neural Information Processing Systems", Montréal, Canada, December 2018, p. 1-17, https://arxiv.org/abs/1806.11328 , https://hal.inria.fr/hal-01937002

[20] C. COUPRIE, P. LUC, J. VERBEEK.*Joint Future Semantic and Instance Segmentation Prediction*, in "ECCV Workshop on Anticipating Human Behavior", Munich, Germany, 2018, https://hal.inria.fr/hal-01867746

[21] M. DOUZE, A. SABLAYROLLES, H. JÉGOU.*Link and code: Fast indexing with graphs and compact regression codes*, in "CVPR 2018 - IEEE Conference on Computer Vision & Pattern Recognition", Salt Lake City, United States, IEEE, June 2018, p. 1-9, https://hal.inria.fr/hal-01955971

[22] G. DURIF, L. MODOLO, J. E. MOLD, S. LAMBERT-LACROIX, F. PICARD.*Probabilistic Count Matrix Factorization for Single Cell Expression Data Analysis*, in "RECOMB 2018 - 22nd Annual International Conference on Research in Computational Molecular Biology", Paris, France, B. J. RAPHAEL (editor), Lecture Notes in Bioinformatics, Springer, April 2018, vol. 10812, p. 254-255, https://hal.archives-ouvertes.fr/hal-01962030

[23] N. DVORNIK, J. MAIRAL, C. SCHMID.*Modeling Visual Context is Key to Augmenting Object Detection Datasets*, in "ECCV 2018 - European Conference on Computer Vision", Munich, Germany, LNCS, Springer, September 2018, vol. 11216, p. 375-391, https://arxiv.org/abs/1807.07428 [*DOI :* 10.1007/978-3-030-01258-8_23], https://hal.archives-ouvertes.fr/hal-01844474

[24] M. ELBAYAD, L. BESACIER, J. VERBEEK.*Pervasive Attention: 2D Convolutional Neural Networks for Sequence-to-Sequence Prediction*, in "CoNLL 2018 - Conference on Computational Natural Language Learning", Brussels, Belgium, October 2018, p. 1-11, https://hal.inria.fr/hal-01851612

[25] M. ELBAYAD, L. BESACIER, J. VERBEEK.*Token-level and sequence-level loss smoothing for RNN language models*, in "ACL - 56th Annual Meeting of the Association for Computational Linguistics", Melbourne, Australia, July 2018, https://hal.inria.fr/hal-01790879

[26] C. Gu, C. Sun, D. Ross, C. Vondrick, C. Pantofaru, Y. Li, S. Vijayanarasimhan, G. Toderici, S. Ricco, R. R. Sukthankar, C. Schmid, J. Malik.*AVA: A Video Dataset of Spatio-temporally Localized Atomic Visual Actions*, in "CVPR 2018 - Computer Vision and Pattern Recognition", Salt Lake City, United States, IEEE, June 2018, p. 1-10, https://hal.inria.fr/hal-01764300

[27] X. Li, J. Ylioinas, J. Verbeek, J. Kannala.*Scene Coordinate Regression with Angle-Based Reprojection Loss for Camera Relocalization*, in "ECCV 2018 - Workshop Geometry Meets Deep Learning", Munich, Germany, September 2018, p. 1-16, https://hal.inria.fr/hal-01867143

[28] P. Luc, C. Couprie, Y. Lecun, J. Verbeek.*Predicting Future Instance Segmentation by Forecasting Convolutional Features*, in "ECCV 2018 - European Conference on Computer Vision", Munich, Germany, September 2018, p. 1-21, https://hal.inria.fr/hal-01757669

[29] T. Lucas, C. Tallec, J. Verbeek, Y. Ollivier.*Mixed batches and symmetric discriminators for GAN training*, in "ICML - 35th International Conference on Machine Learning", Stockholm, Sweden, July 2018, https://hal.inria.fr/hal-01791126

[30] T. Lucas, J. Verbeek.*Auxiliary Guided Autoregressive Variational Autoencoders*, in "ECML-PKDD 2018 - European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases", Dublin, Ireland, September 2018, p. 1-16, https://hal.inria.fr/hal-01652881

[31] C. Paquette, H. Lin, D. Drusvyatskiy, J. Mairal, Z. Harchaoui.*Catalyst for Gradient-based Nonconvex Optimization*, in "AISTATS 2018 - 21st International Conference on Artificial Intelligence and Statistics", Lanzarote, Spain, April 2018, p. 1-10, https://hal.inria.fr/hal-01773296

[32] K. Shmelkov, C. Schmid, K. Alahari.*How good is my GAN?*, in "ECCV 2018 - European Conference on Computer Vision", Munich, Germany, V. Ferrari, M. Hebert, C. Sminchisescu, Y. Weiss (editors), Lecture Notes in Computer Science, September 2018, p. 1-20, https://arxiv.org/abs/1807.09499 , https://hal.inria.fr/hal-01850447

[33] G. A. Sigurdsson, A. Gupta, C. Schmid, A. Farhadi, K. Alahari.*Actor and Observer: Joint Modeling of First and Third-Person Videos*, in "CVPR 2018 - IEEE Conference on Computer Vision & Pattern Recognition", Salt Lake City, Utah, United States, IEEE, June 2018, p. 1-6, https://arxiv.org/abs/1804.09627 , https://hal.inria.fr/hal-01755547

[34] G. Varol, D. Ceylan, B. Russell, J. Yang, E. Yumer, I. Laptev, C. Schmid.*BodyNet: Volumetric Inference of 3D Human Body Shapes*, in "ECCV 2018 - 15th European Conference on Computer Vision", Munich, Germany, September 2018, p. 1-27, https://arxiv.org/abs/1804.04875 , https://hal.inria.fr/hal-01852169

[35] N. Verma, E. Boyer, J. Verbeek.*FeaStNet: Feature-Steered Graph Convolutions for 3D Shape Analysis*, in "CVPR - IEEE Conference on Computer Vision & Pattern Recognition", Salt Lake City, United States, June 2018, p. 1-9, https://arxiv.org/abs/1706.05206 , https://hal.inria.fr/hal-01540389

[36] D. Wynen, C. Schmid, J. Mairal.*Unsupervised Learning of Artistic Styles with Archetypal Style Analysis*, in "NeurIPS 2018", Montréal, Canada, December 2018, Accepted at NIPS 2018, Montréal, Canada, https://hal.inria.fr/hal-01802131

Other Publications

[37] A. Bietti, A. Agarwal, J. Langford.*A Contextual Bandit Bake-off*, December 2018, https://arxiv.org/abs/1802.04064 - working paper or preprint, https://hal.inria.fr/hal-01708310

[38] A. Bietti, J. Mairal.*Group Invariance, Stability to Deformations, and Complexity of Deep Convolutional Representations*, October 2018, https://arxiv.org/abs/1706.03078 - working paper or preprint, https://hal.inria.fr/hal-01536004

[39] A. Bietti, G. Mialon, J. Mairal.*On Regularization and Robustness of Deep Neural Networks*, November 2018, https://arxiv.org/abs/1810.00363 - working paper or preprint, https://hal.inria.fr/hal-01884632

[40] D. Chen, L. Jacob, J. Mairal.*Biological Sequence Modeling with Convolutional Kernel Networks*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01632912

[41] G. Chéron, A. Osokin, I. Laptev, C. Schmid.*Modeling Spatio-Temporal Human Track Structure for Action Localization*, January 2019, https://arxiv.org/abs/1806.11008 - working paper or preprint, https://hal.inria.fr/hal-01979583

[42] N. Dvornik, J. Mairal, C. Schmid.*On the Importance of Visual Context for Data Augmentation in Scene Understanding*, September 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01869784

[43] H. Lin, J. Mairal, Z. Harchaoui.*An Inexact Variable Metric Proximal Point Algorithm for Generic Quasi-Newton Acceleration*, July 2018, https://arxiv.org/abs/1610.00960 - working paper or preprint, https://hal.inria.fr/hal-01376079

[44] A. Mensch, J. Mairal, B. Thirion, G. Varoquaux.*Extracting Universal Representations of Cognition across Brain-Imaging Studies*, October 2018, https://arxiv.org/abs/1809.06035 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01874713

[45] A. Pashevich, D. Hafner, J. Davidson, R. R. Sukthankar, C. Schmid.*Modulated Policy Hierarchies*, December 2018, Deep RL workshop at NIPS 2018, https://hal.archives-ouvertes.fr/hal-01963580

[46] J. Peyre, I. Laptev, C. Schmid, J. Sivic.*Detecting rare visual relations using analogies*, January 2019, https://arxiv.org/abs/1812.05736 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01975760

[47] A. Ruiz, O. Martinez, X. Binefa, J. Verbeek.*Learning Disentangled Representations with Reference-Based Variational Autoencoders*, October 2018, working paper or preprint, https://hal.inria.fr/hal-01896007

[48] K. Shmelkov, T. Lucas, K. Alahari, C. Schmid, J. Verbeek.*Coverage and quality driven training of generative image models*, October 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01886285

Team TRIPOP

Modélisation, simulation et commande des systèmes dynamiques non lisses

Inria teams are typically groups of researchers working on the definition of a common project, and objectives, with the goal to arrive at the creation of a project-team. Such project-teams may include other partners (universities or research institutions).

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Optimization and control of dynamic systems

Table of contents

Team TRIPOP

Creation of the Team: 2018 January 01

Keywords:

Computer Science and Digital Science:

A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.4. - Multiscale modeling
A6.4.1. - Deterministic control
A6.4.3. - Observability and Controlability
A6.4.4. - Stability and Stabilization
A6.4.5. - Control of distributed parameter systems
A6.4.6. - Optimal control
A6.5.1. - Solid mechanics
A6.5.4. - Waves

Other Research Topics and Application Domains:

B3.3.1. - Earth and subsoil
B5.2.1. - Road vehicles
B5.2.3. - Aviation
B5.2.4. - Aerospace
B5.4. - Microelectronics
B5.6. - Robotic systems
B7.1.2. - Road traffic
B9.5.2. - Mathematics
B9.5.5. - Mechanics
B9.11.1. - Environmental risks

1. Team, Visitors, External Collaborators

Research Scientists

Vincent Acary [Inria, Team leader, Senior Researcher, HDR]
Bernard Brogliato [Inria, Senior Researcher, HDR]
Franck Bourrier [IRSTEA, Researcher, HDR]
Arnaud Tonnelier [Inria, Researcher]

Faculty Member

Guillaume James [Institut polytechnique de Grenoble, Professor, HDR]

Technical Staff

Franck Pérignon [CNRS]
Maurice Brémond [Inria]

PhD Students

Charlelie Bertrand [ENTPE, from Nov 2018]
Rami Sayoud [Schneider Electric, from Feb 2018]
Alexandre Vieira [Institut polytechnique de Grenoble, until Sep 2018]

Post-Doctoral Fellows

Saeed Ahmed [Inria, from Dec 2018]
Achref Elmansour [Inria, from Apr 2018]
Alexandre Rocca [Inria, from Nov 2018]
Kirill Vorotnikov [Inria, until Jun 2018]

Visiting Scientists

Jozsef Kovecses [McGill University, from Mar 2018 until Jun 2018]
Mathias Legrand [McGill University, from Sep 2018 until Nov 2018]
Christophe Prieur [CNRS, Gipsa-lab]

Administrative Assistant

Diane Courtiol [Inria]

2. Overall Objectives

2.1. Introduction

The joint research team, TRIPOP, between Inria Grenoble Rhône–Alpes, Grenoble INP and CNRS, part of the Laboratoire Jean Kuntzmann (LJK UMR 5224) is a follow up of the BIPOP team (2003–2017). The team is mainly concerned by the modeling, the mathematical analysis, the simulation and the control of nonsmooth dynamical systems. Nonsmooth dynamics concerns the study of the time evolution of systems that are not smooth in the mathematical sense, i.e., systems that are characterized by a lack of differentiability, either of the mappings in their formulations, or of their solutions with respect to time. In mechanics, the main instances of nonsmooth dynamical systems are multibody systems with Signorini unilateral contact, set-valued (Coulomb-like) friction and impacts. In Electronics, examples are found in switched electrical circuits with ideal components (diodes, switches, transistors). In Control, nonsmooth systems arise in the sliding mode control theory and in optimal control. A lot of examples can also be found in cyber-physical systems (hybrid systems), in transportation sciences, in mathematical biology or in finance.

2.2. General scope and motivations

Nonsmooth dynamics concerns the study of the time evolution of systems that are not smooth in the mathematical sense, *i.e.*, systems that are characterized by a lack of differentiability, either of the mappings in their formulations, or of their solutions with respect to time. The class of nonsmooth dynamical systems recovers a large variety of dynamical systems that arise in many applications. The term “nonsmooth”, as the term “nonlinear”, does not precisely define the scope of the systems we are interested in but, and most importantly, they are characterized by the mathematical and numerical properties that they share. To give more insight of what are nonsmooth dynamical systems, we give in the sequel a very brief introduction of their salient features. For more details, we refer to [1], [2] [61], [78], [94], [63], [39].

2.2.1. *A flavor of nonsmooth dynamical systems*

As a *first* illustration, let us consider a linear finite-dimensional system described by its state $x(t) \in \mathbb{R}^n$ over a time-interval $t \in [0, T]$:

$$\dot{x}(t) = Ax(t) + a, \quad A \in \mathbb{R}^{n\times n},\, a \in \mathbb{R}^n, \tag{5}$$

subjected to a set of m inequality (unilateral) constraints:

$$y(t) = Cx(t) + c \geq 0, \qquad C \in \mathbb{R}^{m\times n},\, c \in \mathbb{R}^m. \tag{6}$$

If the constraints are physical constraints, a standard modeling approach is to augment the dynamics in (1) by an input vector $\lambda(t) \in \mathbb{R}^m$ that plays the role of a Lagrange multiplier vector. The multiplier restricts the trajectory of the system in order to respect the constraints. Furthermore, as in the continuous optimization theory, the multiplier must be signed and must vanish if the constraint is not active. This is usually formulated as a complementarity condition:

$$0 \leq y(t) \perp \lambda(t) \geq 0, \tag{7}$$

which models the one-sided effect of the inequality constraints. The notation $y \geq 0$ holds component–wise and $y \perp \lambda$ means $y^T\lambda = 0$. All together we end up with a Linear Complementarity System (LCS) of the form,

$$\begin{cases} \dot{x}(t) = Ax(t) + a + B\lambda(t) \\ y(t) = Cx(t) + c \\ 0 \leq y(t) \perp \lambda(t) \geq 0 \end{cases} \tag{8}$$

where $B \in \mathbb{R}^{n\times m}$ is the matrix that models the input generated by the constraints. In a more general way, the constraints may also involve the Lagrange multiplier,

$$y(t) = Cx(t) + c + D\lambda(t) \geq 0, \qquad D \in \mathbb{R}^{m\times m}, \tag{9}$$

leading to a general definition of LCS as

$$\begin{cases} \dot{x}(t) = A\,x(t) + a + B\,\lambda(t) \\ y(t) = C\,x(t) + c + D\,\lambda(t) \\ 0 \leq y(t) \perp \lambda(t) \geq 0. \end{cases} \tag{10}$$

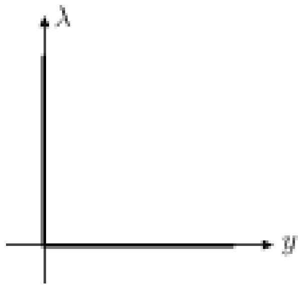


Figure 1. Complementarity condition $0 \leq y \perp \lambda \geq 0$.

The complementarity condition, illustrated in Figure 1 is the archetype of a nonsmooth graph that we extensively use in nonsmooth dynamics. The mapping $y \mapsto \lambda$ is a multi-valued (set-valued) mapping, that is nonsmooth at the origin. It has a lot of interesting mathematical properties and reformulations that come mainly from convex analysis and variational inequality theory. Let us introduce the indicator function of $\mathbb{R}_+$ as

$$\Psi_{\mathbb{R}_+}(x) = \begin{cases} 0 \text{ if } x \geq 0, \\ +\infty \text{ if } x < 0. \end{cases} \tag{11}$$

This function is convex, proper and can be sub-differentiated [67]. The definition of the subdifferential of a convex function $f : I\!R^m \to I\!R$ is defined as:

$$\partial f(x) = \{x^{\star} \in I\!R^m \mid f(z) \geq f(x) + (z-x)^\top x^{\star}, \forall z\}. \tag{12}$$

A basic result of convex analysis reads as

$$0 \leq y \perp \lambda \geq 0 \longleftrightarrow -\lambda \in \partial \Psi_{\boldsymbol{R}_+}(y) \tag{13}$$

that gives a first functional meaning to the set-valued mapping $y \mapsto \lambda$. Another interpretation of $\partial\Psi_{\boldsymbol{R}_+}$ *is based on* the normal cone to a closed and nonempty convex set C:

$$N_C(x) = \{v \in I\!R^m | v^\top (z-x) \leq 0 \text{ for all } z \in K\}. \tag{14}$$

It is easy to check that $\partial\Psi_{\boldsymbol{R}_+} = N_{\boldsymbol{R}_+}(x)$ and it follows that

$$0 \leq y \perp \lambda \geq 0 \longleftrightarrow -\lambda \in N_{\boldsymbol{R}_+}(y). \tag{15}$$

Finally, the definition of the normal cone yields a variational inequality:

$$0 \leq y \perp \lambda \geq 0 \longleftrightarrow \lambda^\top (y-z) \leq 0, \forall z \geq 0. \tag{16}$$

The relations (11) and (12) allow one to formulate the complementarity system with $D = 0$ as a differential inclusion based on a normal cone (see (15)) or as a differential variational inequality. By extending the definition to other types of convex functions, possibly nonsmooth, and using more general variational inequalities, the same framework applies to the nonsmooth laws depicted in Figure 2 that includes the case of piecewise smooth systems.

The mathematical concept of solutions depends strongly on the nature of the matrix quadruplet (A, B, C, D) in (6). If D is a positive definite matrix (or a P-matrix), the Linear Complementarity problem

$$0 \leq Cx + c + D\lambda \perp \lambda \geq 0, \tag{17}$$

admits a unique solution $\lambda(x)$ which is a Lipschitz continuous mapping. It follows that the Ordinary Differential Equation (ODE)

$$\dot{x}(t) = Ax(t) + a + B\lambda(x(t)), \tag{18}$$

is a standard ODE with a Lipschitz right-hand side with a C^1 solution for the initial value problem. If $D = 0$, the system can be written as a differential inclusion in a normal cone as

$$-\dot{x}(t) + Ax(t) + a \in BN_{\boldsymbol{R}_+}(Cx(t)), \tag{19}$$

that admits a solution that is absolutely continuous if CB is a definite positive matrix and the initial condition satisfies the constraints. The time derivative $\dot{x}(t)$ and the multiplier $\lambda(t)$ may have jumps and are generally considered as functions of bounded variations. If $CB = 0$, the order of nonsmoothness increases and the Lagrange multiplier may contain Dirac atoms and must be considered as a measure. Higher–order index, or higher relative degree systems yield solutions in terms of distributions and derivatives of distributions [33].

A lot of variants can be derived from the basic form of linear complementarity systems, by changing the form of the dynamics including nonlinear terms or by changing the complementarity relation by other multivalued maps. In particular the nonnegative orthant may be replaced by any convex closed cone $K \subset \mathbb{R}^m$ leading to complementarity over cones

$$K^{\star} \ni y \perp \lambda \in K, \tag{20}$$

where $K^{\star}$ its dual cone given by

$$K^{\star} = \{x \in \mathbb{R}^m \mid x^\top y \geq 0 \text{ for all } y \in K\}. \tag{21}$$

In Figure 2, we illustrate some other basic maps that can used for defining the relation between λ and y. The saturation map, depicted in Figure 2(a) is a single valued continuous function which is an archetype of piecewise smooth map. In Figure 2(b), the relay multi-function is illustrated. If the upper and the lower limits of λ are respectively equal to 1 and -1, we obtain the multivalued sign function defined as

$$\mathrm{Sgn}\,(y) = \begin{cases} 1, & y > 0 \\ [-1,1], & y = 0 \\ -1, & y < 0. \end{cases} \tag{22}$$

Using again convex analysis, the multivalued sign function may be formulated as an inclusion into a normal cone as

$$\lambda \in \mathrm{Sgn}\,(y) \longleftrightarrow y \in N_{[-1,1]}(\lambda). \tag{23}$$

More generally, any system of the type,

$$\begin{cases} \dot{x}(t) = Ax(t) + a + B\lambda(t) \\ y(t) = Cx(t) + a \\ -\lambda(t) \in \mathrm{Sgn}(y(t)), \end{cases} \tag{24}$$

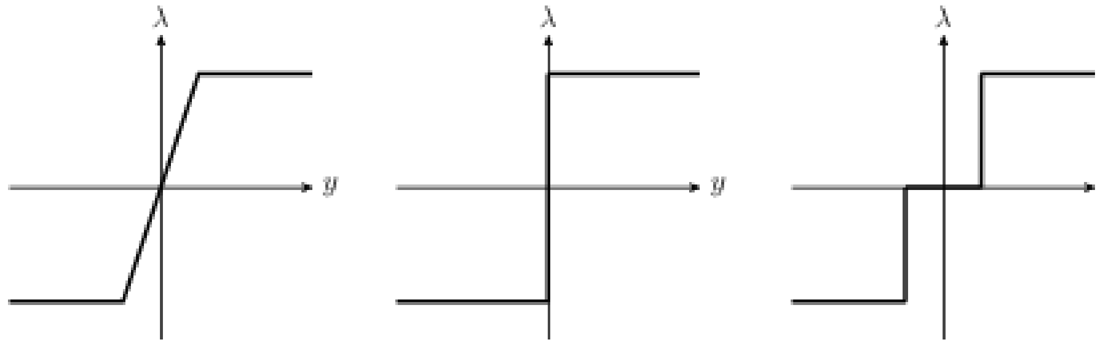
can reformulated in terms of the following set-valued system

$$\begin{cases} \dot{x}(t) = Ax(t) + a + B\lambda(t) \\ y(t) = Cx(t) + c \\ -y(t) \in N_{[-1,1]^m}(\lambda(t)). \end{cases} \tag{25}$$

The system (21) appears in a lot of applications; among them, we can cite the sliding mode control, electrical circuits with relay and Zener diodes [29], or mechanical systems with friction [31].

Though this class of systems seems to be rather specific, it includes as well more general dynamical systems such as piecewise smooth systems and discontinuous ordinary differential equations. Indeed, the system (20) for scalars y and λ can be viewed as a discontinuous differential equation:

$$\dot{x}(t) = \begin{cases} Ax + a + B & \text{if} \quad Cx + c > 0 \\ Ax + a - B & \text{if} \quad Cx + c < 0. \end{cases} \tag{26}$$



(a) Saturation (b) Relay (c) Relay with dead zone

Figure 2. Examples of multivalued piecewise linear models

One of the most well-known mathematical framework to deal with such systems is the Filippov theory [61] that embed the discontinuous differential equations into a differential inclusion. In the case of a single discontinuity surface given in our example by $S = \{x \mid Cx + c = 0\}$, the Filippov differential inclusion based on the convex hull of the vector fields in the neighborhood of S is equivalent to the use of the multivalued sign function in (20). Conversely, as it has been shown in [37], a piecewise smooth system can be formulated as a nonsmooth system *based on products of multivalued sign functions*.

2.2.2. Nonsmooth Dynamical systems in the large

Generally, the nonsmooth dynamical systems we propose to study mainly concern systems that possess the following features:

1. A nonsmooth formulation of the constitutive/behavioral laws that define the system. Examples of nonsmooth formulations are piecewise smooth functions, multi–valued functions, inequality constraints, yielding various definitions of dynamical systems such as piecewise smooth systems, discontinuous ordinary differential equations, complementarity systems, projected dynamical systems, evolution or differential variational inequalities and differential inclusions (into normal cones). Fundamental mathematical tools come from convex analysis [87], [68], [67], complementarity theory [57], and variational inequalities theory [60].
2. A concept of solutions that does not require continuously differentiable functions of time. For instance, absolutely continuous, Lipschitz continuous functions or functions of local bounded variation are the basis for solution concepts. Measures or distributions are also solutions of interest for differential inclusions or evolution variational inequalities.

2.2.3. Nonsmooth systems versus hybrid systems

The nonsmooth dynamical systems we are dealing with, have a nonempty intersection with hybrid systems and cyber-physical systems, as it is briefly discussed in Sect. 3.2.4. Like in hybrid systems, nonsmooth dynamical systems define continuous–time dynamics that can be identified to modes separated by guards, defined by the constraints. However, the strong mathematical structure of nonsmooth dynamical systems allows us to state results on the following points:

1. *Mathematical concept of solutions: well-posedness (existence, and possibly, uniqueness properties, (dis)continuous dependence on initial conditions).*
2. *Dynamical systems theoretic properties: existence of invariants (equilibria, limit cycles, periodic solutions,...) and their stability, existence of oscillations, periodic and quasi-periodic solutions and propagation of waves.*
3. *Control theoretic properties: passivity, controllability, observability, stabilization, robustness.*

These latter properties, that are common for smooth nonlinear dynamical systems, distinguish the nonsmooth dynamical systems from the very general definition of hybrid or cyber-physical systems [42], [66]. Indeed, it is difficult to give a precise mathematical concept of solutions for hybrid systems since the general definition of hybrid automata is usually too loose.

2.2.4. Numerical methods for nonsmooth dynamical systems

To conclude this brief exposition of nonsmooth dynamical systems, let us recall an important fact related to numerical methods. Beyond their intrinsic mathematical interest, and the fact that they model real physical systems, using nonsmooth dynamical systems as a model is interesting, because it exists a large set of robust and efficient numerical techniques to simulate them. Without entering into deeper details, let us give two examples of these techniques:

- *Numerical time integration methods: convergence, efficiency (order of consistency, stability, symplectic properties).* For the nonsmooth dynamical systems described above, there exist event–capturing time–stepping schemes with strong mathematical results. These schemes have the ability to numerically integrate the initial value problem without performing an event location, but by capturing the event within a time step. We call an event, or a transition, every change into the index set of the active constraints in the complementarity formulation or in the normal cone inclusion. Hence these schemes are able to simulate systems with a huge number of transitions or even worth finite accumulation of events (Zeno behavior). Furthermore, the schemes are not suffering from the weaknesses of the standard schemes based on a regularization (smoothing) of the multi-valued mapping resulting in stiff ordinary differential equations. *For the time–integration of the initial value problem (IVP), or Cauchy problem, a lot of improvements of the standard time–stepping schemes for nonsmooth dynamics (Moreau–Jean time-stepping scheme) have been proposed in the last decade, in terms of accuracy and dissipation properties [26], [27], [88], [89], [28], [56], [52], [90], [54]. An important part of these schemes has been developed by members of the BIPOP team and has been implemented in the Siconos software (see Sect. 5.1).*
- *Numerical solution procedure for the time–discretized problem, mainly through well-identified problems studied in the optimization and mathematical programming community.* Another very interesting feature is the fact that the discretized problem that we have to solve at each time–step is generally a well-known problem in optimization. For instance, for LCSs, we have to solve a linear complementarity problem [57] for which there exist efficient solvers in the literature. Comparing to the brute force algorithm with exponential complexity that consists in enumerating all the possible modes, the algorithms for linear complementarity problem have polynomial complexity when the problem is monotone.

In the Axis 2 of the research program (see Sect. 3.3), we propose to perform new research on the geometric time-integration schemes of nonsmooth dynamical systems, to develop new integration schemes for Boundary Value Problem (BVP), and to work on specific methods for two time-discretized problems: the Mathematical Program with Equilibrium Constraints (MPEC) for optimal control and Second Order Cone Complementarity Problems (SOCCP) for discrete frictional contact systems.

3. Research Program

3.1. Introduction

In this section, we develop our scientific program. In the framework of nonsmooth dynamical systems, the activities of the project–team will be on focused on the following research axes:

- *Axis 1: Modeling and analysis (detailed in Sect. 3.2).*
- *Axis 2: Numerical methods and simulation (detailed in Sect. 3.3).*
- *Axis 3: Automatic Control (detailed in Sect. 3.4)*

These research axes will be developed with a strong emphasis on the software development and the industrial transfer that are detailed respectively in Sect. 5.1 and Sect. 7.1.

3.2. Axis 1: Modeling and analysis

This axis is dedicated to the modeling and the mathematical analysis of nonsmooth dynamical systems. It consists of four main directions. Two directions are in the continuation of BIPOP activities: 1) multibody vibro-impact systems (Sect. 3.2.1) and 2) excitable systems (Sect. 3.2.2). Two directions are completely new with respect to BIPOP: 3) Nonsmooth geomechanics and natural hazards assessment (Sect. 3.2.3) and 4) Cyber-physical systems (hybrid systems) (Sect. 3.2.4).

3.2.1. Multibody vibro-impact systems

Participants: B. Brogliato, F. Bourrier, G. James, V. Acary

- *Multiple impacts with or without friction :* there are many different approaches to model collisions, especially simultaneous impacts (so-called multiple impacts) [84]. One of our objectives is on one hand to determine the range of application of the models (for instance, when can one use "simplified" rigid contact models relying on kinematic, kinetic or energetic coefficients of restitution?) on typical benchmark examples (chains of aligned beads, rocking block systems). On the other hand, try to take advantage of the new results on nonlinear waves phenomena, to better understand multiple impacts in 2D and 3D granular systems. The study of multiple impacts with (unilateral) nonlinear visco-elastic models (Simon-Hunt-Crossley, Kuwabara-Kono), or visco-elasto-plastic models (assemblies of springs, dashpots and dry friction elements), is also a topic of interest, since these models are widely used.
- *Artificial or manufactured or ordered granular crystals, meta-materials :* Granular metamaterials (or more general nonlinear mechanical metamaterials) offer many perspectives for the passive control of waves originating from impacts or vibrations. The analysis of waves in such systems is delicate due to spatial discreteness, nonlinearity and non-smoothness of contact laws [86], [72], [73], [79]. We will use a variety of approaches, both theoretical (e.g. bifurcation theory, modulation equations) and numerical, in order to describe nonlinear waves in such systems, with special emphasis on energy localization phenomena (excitation of solitary waves, fronts, breathers).
- *Systems with clearances, modeling of friction :* joint clearances in kinematic chains deserve specific analysis, especially concerning friction modeling [40]. Indeed contacts in joints are often conformal, which involve large contact surfaces between bodies. Lubrication models should also be investigated.
- *Painlevé paradoxes :* the goal is to extend the results in [65], which deal with single-contact systems, to multi-contact systems. One central difficulty here is the understanding and the analysis of singularities that may occur in sliding regimes of motion.

As a continuation of the work in the BIPOP team, our software code, Siconos (see Sect. 5.1) will be our favorite software platform for the integration of these new modeling results.

3.2.2. Excitable systems

Participants: A. Tonnelier, G. James

An excitable system elicits a strong response when the applied perturbation is greater than a threshold [81], [82], [45], [91]. This property has been clearly identified in numerous natural and physical systems. In mechanical systems, non-monotonic friction law (of spinodal-type) leads to excitability. Similar behavior may be found in electrical systems such as active compounds of neuristor type. Models of excitable systems incorporate strong non-linearities that can be captured by non-smooth dynamical systems. Two properties are deeply associated with excitable systems: oscillations and propagation of nonlinear waves (autowaves in coupled excitable systems). We aim at understanding these two dynamical states in excitable systems through theoretical analysis and numerical simulations. Specifically we plan to study:

- Threshold-like models in biology: spiking neurons, gene networks.
- Frictional contact oscillators (slider block, Burridge-Knopoff model).
- Dynamics of active electrical devices : memristors, neuristors.

3.2.3. Nonsmooth geomechanics and natural hazards assessment

Participants: F. Bourrier, B. Brogliato, G. James, V. Acary

- *Rockfall impact modeling :* Trajectory analysis of falling rocks during rockfall events is limited by a rough modeling of the impact phase [47], [46], [77]. The goal of this work is to better understand the link between local impact laws at contact with refined geometries and the efficient impact laws written for a point mass with a full reset map. A continuum of models in terms of accuracy and complexity will be also developed for the trajectory studies. In particular, nonsmooth models of rolling friction, or rolling resistance will be developed and formulated using optimization problems.
- *Experimental validation :* The participation of IRSTEA with F. Bourrier makes possible the experimental validation of models and simulations through comparisons with real data. IRSTEA has a large experience of lab and in-situ experiments for rockfall trajectories modeling [47], [46]. It is a unique opportunity to vstrengthen our model and to prove that nonsmooth modeling of impacts is reliable for such experiments and forecast of natural hazards.
- *Rock fracturing :* When a rock falls from a steep cliff, it stores a large amount of kinetic energy that is partly dissipated though the impact with the ground. If the ground is composed of rocks and the kinetic energy is sufficiently high, the probability of the fracture of the rock is high and yields an extra amount of dissipated energy but also an increase of the number of blocks that fall. In this item, we want to use the capability of the nonsmooth dynamical framework for modeling cohesion and fracture [74], [38] to propose new impact models.
- *Rock/forest interaction :* To prevent damages and incidents to infrastructures, a smart use of the forest is one of the ways to control trajectories (decrease of the run-out distance, jump heights and the energy) of the rocks that fall under gravity [58], [59]. From the modeling point of view and to be able to improve the protective function of the forest, an accurate modeling of impacts between rocks and trees is required. Due to the aspect ratio of the trees, they must be considered as flexible bodies that may be damaged by the impact. This new aspect offers interesting modeling research perspectives.

More generally, our collaboration with IRSTEA opens new long term perspectives on granular flows applications such as debris and mud flows, granular avalanches and the design of structural protections. *The numerical methods that go with these new modeling approaches will be implemented in our software code, Siconos (see Sect. 5.1)*

3.2.4. Cyber-physical systems (hybrid systems)

Participants: V. Acary, B. Brogliato, C. Prieur, A. Tonnelier

Nonsmooth systems have a non-empty intersection with hybrid systems and cyber–physical systems. However, nonsmooth systems enjoy strong mathematical properties (concept of solutions, existence and uniqueness) and efficient numerical tools. This is often the result of the fact that nonsmooth dynamical systems are models of physical systems, and then, take advantage of their intrinsic property (conservation or dissipation of energy, passivity, stability). A standard example is a circuit with n ideal diodes. From the hybrid point of view, this circuit is a piecewise smooth dynamical system with 2^n modes, that can be quite cumbersome to enumerate in order to determinate the current mode. As a nonsmooth system, this circuit can be formulated as a complementarity system for which there exist efficient time–stepping schemes and polynomial time algorithms for the computation of the current mode. The key idea of this research action is to take benefit of this observation to improve the hybrid system modeling tools.

Research actions: There are two main actions in this research direction that will be implemented in the framework of the Inria Project Lab (IPL " Modeliscale", see https://team.inria.fr/modeliscale/ for partners and details of the research program):

•*Structural analysis of multimode DAE :* When a hybrid system is described by a Differential Algebraic Equation (DAE) with different differential indices in each continuous mode, the structural analysis has to be completely rethought. In particular, the re-initialization rule, when a switching occurs from a mode to another one, has to be consistently designed. We propose in this action to use our knowledge in complementarity and (distribution) differential inclusions [33] to design consistent re-initialization rule for systems with nonuniform relative degree vector $(r_1, r_2, ..., r_m)$ and $r_i \neq r_j, i \neq j$.

•*Cyber–physical in hybrid systems modeling languages :* Nowadays, some hybrid modeling languages and tools are widely used to describe and to simulate hybrid systems (MODELICA, SIMULINK, and see [55] for references therein). Nevertheless, the compilers and the simulation engines behind these languages and tools suffer from several serious weaknesses (failure, weird output or huge sensitivity to simulation parameters), especially when some components, that are standard in nonsmooth dynamics, are introduced (piecewise smooth characteristic, unilateral constraints and complementarity condition, relay characteristic, saturation, dead zone, ...). One of the main reasons is the fact that most of the compilers reduce the hybrid system to a set of smooth modes modeled by differential algebraic equations and some guards and reinitialization rules between these modes. Sliding mode and Zeno–behaviour are really harsh for hybrid systems and relatively simple for nonsmooth systems. With B. Caillaud (Inria HYCOMES) and M. Pouzet (Inria PARKAS), we propose to improve this situation by implementing a module able to identify/describe nonsmooth elements and to efficiently handle them with SICONOS as the simulation engine. They have already carried out a first implementation [53] in Zelus, a synchronous language for hybrid systems http://zelus.di.ens.fr. Removing the weaknesses related to the nonsmoothness of solutions should improve hybrid systems towards robustness and certification.

•*A general solver for piecewise smooth systems* This direction is the continuation of the promising result on modeling and the simulation of piecewise smooth systems [37]. As for general hybrid automata, the notion or concept of solutions is not rigorously defined from the mathematical point of view. For piecewise smooth systems, multiplicity of solutions can happen and sliding solutions are common. The objective is to recast general piecewise smooth systems in the framework of differential inclusions with Aizerman–Pyatnitskii extension [37], [61]. This operation provides a precise meaning to the concept of solutions. Starting from this point, the goal is to design and study an efficient numerical solver (time–integration scheme and optimization solver) based on an equivalent formulation as mixed complementarity systems of differential variational inequalities. We are currently discussing the issues in the mathematical analysis. The goal is to prove the convergence of the time–stepping scheme to get an existence theorem. With this work, we should also be able to discuss the general Lyapunov stability of stationary points of piecewise smooth systems.

3.3. Axis 2: Numerical methods and simulation

This axis is dedicated to the numerical methods and simulation for nonsmooth dynamical systems. As we mentioned in the introduction, the standard numerical methods have been largely improved in terms of accuracy and dissipation properties in the last decade. Nevertheless, the question of the geometric time–integration techniques remains largely open. It constitutes the objective of the first research direction in Sect. 3.3.1. Beside the standard IVP, the question of normal mode analysis for nonsmooth systems is also a research topic that emerged in the recent years. More generally, the goal of the second research direction (Sect. 3.3.2) is to develop numerical methods to solve boundary value problems in the nonsmooth framework. This will serve as a basis for the computation of the stability and numerical continuation of invariants. Finally, once the time-integration method is chosen, it remains to solve the one-step nonsmooth problem, which is, most of time, a numerical optimization problem. In Sect. 3.3.3, we propose to study two specific problems with a lot of applications: the Mathematical Program with Equilibrium Constraints (MPEC) for optimal control, and Second Order Cone Complementarity Problems (SOCCP) for discrete frictional contact systems. After some possible prototypes in scripting languages (Python and Matlab), we will be attentive that all these developments of numerical methods will be integrated in Siconos.

3.3.1. Geometric time–integration schemes for nonsmooth Initial Value Problem (IVP)

Participants: V. Acary, B. Brogliato, G. James, F. Pérignon

The objective of this research item is to continue to improve classical time–stepping schemes for nonsmooth systems to ensure some qualitative properties in discrete-time. In particular, the following points will be developed

- Conservative and dissipative systems. The question of the energy conservation and the preservation of dissipativity properties in the Willems sense [64] will be pursued and extended to new kinds of systems (nonlinear mechanical systems with nonlinear potential energy, systems with limited differentiability (rigid impacts vs. compliant models)).
- Lie–group integration schemes for finite rotations for the multi-body systems extending recent progresses in that directions for smooth systems [43].
- Conservation and preservation of the dispersion properties of the (non)-dispersive system.

3.3.2. Stability and numerical continuation of invariants

Participants: G. James, V. Acary, A. Tonnelier, F. Pérignon,

By invariants, we mean equilibria, periodic solutions, limit cycles or waves. Our preliminary work on this subject raised the following research perspectives:

- Computation of periodic solutions of discrete mechanical systems . The modal analysis, *i.e.*, a spectral decomposition of the problem into linear normal modes is one of the basic tools for mechanical engineers to study dynamic response and resonance phenomena of an elastic structure. Since several years, the concept of nonlinear normal modes [75], that is closely related to the computation of quasi-periodic solutions that live in a nonlinear manifold, has emerged as the nonlinear extension of the modal analysis. One of the fundamental question is: what remains valid if we add unilateral contact conditions ? The computation of nonsmooth modes amounts to computing periodic solutions, performing the parametric continuation of solution branches and studying the stability of these branches. This calls for time integration schemes for IVP an BVP that satisfy some geometric criteria: conservation of energy, reduced numerical dispersion, symplecticity as we described before. Though the question of conservation of energy for unilateral contact has been discussed in [28], the other questions remain open. For the shooting technique and the study of stability, we need to compute the Jacobian matrix of the flow with respect to initial conditions, the so-called saltation matrix [76], [85] for nonsmooth flows. The eigenvalues of this matrix are the Floquet multipliers that give some information on the stability of the periodic solutions. The question of an efficient computation of this matrix is also an open question. For the continuation, the question is also largely open since the continuity of the solutions with respect to the parameters is not ensured.
- Extension to elastic continuum media . This is a difficult task. First of all, the question of the mathematical model for the dynamic continuum problem with unilateral contact raises some problems of well–posedness. For instance, the need for an impact law is not clear in some cases. If we perform a semi–discretization in space with classical techniques (Finite Element Methods, Finite Difference Schemes), we obtain a discrete system for which the impact law is needed. Besides all the difficulties that we enumerate for discrete systems in the previous paragraph, the space discretization also induces numerical dispersion that may destroy the periodic solutions or renders their computation difficult. The main targeted applications for this research are cable–systems, string musical instruments, and seismic response of electrical circuit breakers with Schneider Electric.
- Computation of solutions of nonsmooth time Boundary Value Problems (BVP) (collocation, shooting) . The technique developed in the two previous items can serve as a basis for the development of more general solvers for nonsmooth BVP that can be for instance found when we solve optimal control problems by direct or indirect methods, or the computation of nonlinear waves. Two directions can be envisaged:
  - Shooting and multiple shooting techniques. In such methods, we reformulate the BVP into a sequence of IVPs that are iterated through a Newton based technique. This implies the computation of Jacobians for nonsmooth flows, the question of the continuity w.r.t to initial condition and the use of semi-smooth Newton methods.

  - Finite differences and collocations techniques. In such methods, the discretization will result into a large sparse optimization problems to solve. The open questions are as follows: a) the study of convergence, b) how to locally improve the order if the solution is locally smooth, and c) how to take benefit of spectral methods.

- Continuation techniques of solutions with respect to a parameter. Standard continuation technique requires smoothness. What types of methods can be extended in the nonsmooth case (arc-length technique, nonsmooth (semi-smooth) Newton, Asymptotical Numerical Methods (ANM))

3.3.3. Numerical optimization for discrete nonsmooth problems

Participants: V. Acary, M. Brémond, F. Pérignon, B. Brogliato, C. Prieur

- Mathematical Program with Equilibrium Constraints (MPEC) for optimal control . The discrete problem that arises in nonsmooth optimal control is generally a MPEC [92]. This problem is intrinsically nonconvex and potentially nonsmooth. Its study from a theoretical point of view has started 10 years ago but there is no consensus for its numerical solving. The goal is to work with world experts of this problem (in particular M. Ferris from Wisconsin University) to develop dedicated algorithms for solving MPEC, and provide to the optimization community challenging problems.
- Second Order Cone Complementarity Problems (SOCCP) for discrete frictional systems : After some extensive comparisons of existing solvers on a large collection of examples [36], [30], the numerical treatment of constraints redundancy by the proximal point technique and the augmented Lagrangian formulation seems to be a promising path for designing new methods. From the comparison results, it appears that the redundancy of constraints prevents the use of second order methods such as semi–smooth Newton methods or interior point methods. With P. Armand (XLIM, U. de Limoges), we propose to adapt recent advances for regularizing constraints for the quadratic problem [62] for the second-order cone complementarity problem. The other question is the improvement of the efficiency of the algorithms by using accelerated schemes for the proximal gradient method that come from large–scale machine learning and image processing problems. Learning from the experience in large–scale machine learning and image processing problems, the accelerated version of the classical gradient algorithm [83] and the proximal point algorithm [44], and many of their further extensions, could be of interest for solving discrete frictional contact problems. Following the visit of Y. Kanno (University of Tokyo) and his preliminary experience on frictionless problems, we will extend its use to frictional contact problem. When we face large-scale problems, the main available solvers is based on a Gauss–Seidel strategy that is intrinsically sequential. Accelerated first-order methods could be a good alternative to take benefit of the distributed scientific computing architectures.

3.4. Axis 3: Automatic Control

Participants: B. Brogliato, C. Prieur, V. Acary

This last axis is dedicated to the automatic control of nonsmooth dynamical systems, or the nonsmooth control of smooth systems. The first item concerns the discrete-time sliding mode control for which significant results on the implicit implementation have been obtained in the BIPOP team. The idea is to pursue this research towards state observers and differentiators (Sect 3.4.1). The second direction concerns the optimal control which brings of nonsmoothness in their solution and their formulation. After the preliminary work in BIPOP on the quadratic optimal control of Linear Complementarity systems(LCS), we propose to go further to the minimal time problem, to impacting systems and optimal control with state constraints (Sect. 3.4.2). In Sect 3.4.3, the objective is to study the control of nonsmooth systems that contain unilateral constraint, impact and friction. The targeted systems are cable–driven systems, multi-body systems with clearances and granular materials. In Sect 3.4.4, we will continue our work on the higher order Moreau sweeping process. Up to now, the work of BIPOP was restricted to finite-dimensional systems. In Sect 3.4.5, we propose to extend our approach to the control of elastic structures subjected to contact unilateral constraints.

It is noteworthy that most of the problems listed below, will make strong use of the numerical tools analyzed in Axis 2, and of the Modeling analysis of Axis 1. For instance all optimal control problems yield BVPs. Control of granular materials will undoubtedly use models and numerical simulation developed in Axis 1 and 2. And so on. It has to be stressed that the type of nonsmooth models we are working with, deserve specific numerical algorithms which cannot be found in commercial software packages. One of the goals is to continue to extend our software package Siconos, and in particular the siconos/control toolbox with these developments.

3.4.1. Discrete-time Sliding-Mode Control (SMC) and State Observers (SMSO)

- *SMSO, exact differentiators:* we have introduced and obtained significant results on the implicit discretization of various classes of sliding-mode controllers [32], [34], [69], [80], [49], with successful experimental validations [70], [69], [71], [93]. Our objective is to prove that the implicit discretization can also bring advantages for sliding-mode state observers and Levant's exact differentiators, compared with the usual explicit digital implementation that generates chattering. In particular the implicit discretization guarantees Lyapunov stability and finite-time convergence properties which are absent in explicit methods.
- *High-Order SMC (HOSMC):* this family of controllers has become quite popular in the sliding-mode scientific community since its introduction by Levant in the nineties. We want here to continue the study of implicit discretization of HOSMC (twisting, super-twisting algorithms) and especially we would like to investigate the comparisons between classical (first order) SMC and HOSMC, when both are implicitly discretized, in terms of performance, accuracy, chattering suppression. Another topic of interest is stabilization in finite-time of systems with impacts and unilateral constraints, in a discrete-time setting.

3.4.2. Optimal Control

- *Linear Complementarity Systems (LCS) :* With the PhD thesis of A. Vieira, we have started to study the quadratic optimal control of LCS. Our objective is to go further with minimum-time problems. Applications of LCS are mainly in electrical circuits with set-valued components such as ideal diodes, transistors, *etc*. Such problems naturally yield MPEC when numerical solvers are sought. It is therefore intimately linked with Axis 2 objectives.
- *Impacting systems :* the optimal control of mechanical systems with unilateral constraints and impacts, largely remains an open issue. The problem can be tackled from various approaches: vibro-impact systems (no persistent contact modes) that may be transformed into discrete-time mappings *via* the impact Poincaré map; or the classical integral action minimization (Bolza problem) subjected to the complementarity Lagrangian dynamics including impacts.
- *State constraints, generalized control :* this problem differs from the previous two, since it yields Pontryagin's first order necessary conditions that take the form of an LCS with higher relative degree between the complementarity variables. This is related to the numerical techniques for the higher order sweeping process [33].

3.4.3. Control of nonsmooth discrete Lagrangian systems

- *Cable–driven systems:* these systems are typically different from the cable-car systems, and are closer in their mechanical structure to so-called tensegrity structures. The objective is to actuate a system *via* cables supposed in a first instance to be flexible (slack mode) but non-extensible in their longitudinal direction. This gives rise to complementarity conditions, one big difference with usual complementarity Lagrangian systems being that the control actions operate directly in one of the complementary variables (and not in the smooth dynamics as in cable-car systems). Therefore both the cable models and the control properties are expected to differ a lot from what we may use for cableway systems (for which guaranteeing a positive cable tension is usually not an issue, hence avoiding slack modes, but the deformation of the cables due to the nacelles and cables weights, is an important factor). Tethered systems are a close topic.

- *Multi-body systems with clearances:* our approach is to use models of clearances with dynamical impact effects, *i.e.* within Lagrangian complementarity systems. Such systems are strongly underactuated due to mechanical play at the joints. However their structure, as underactuated systems, is quite different from what has been usually considered in the Robotics and Control literature. In the recent past we have proposed a thorough numerical robustness analysis of various feedback collocated and non-collocated controllers (PD, linearization, passivity-based). We propose here to investigate specific control strategies tailored to such underactuated systems [48].
- *Granular systems:* the context is the feedback control of granular materials. To fix the ideas, one may think of a "juggling" system whose "object" (uncontrolled) part consists of a chain of aligned beads. Once the modeling step has been fixed (choice of a suitable multiple impact law), one has to determine the output to be controlled: all the beads, some of the beads, the chain's center of mass (position, velocity, vibrational magnitude and frequency), *etc*. Then we aim at investigating which type of controller may be used (output or state feedback, "classical" or sinusoidal input with feedback through the magnitude and frequency) and especially which variables may be measured/observed (positions and/or velocities of all or some of the beads, position and/or velocity of the chain's center of gravity). This topic follows previous results we obtained on the control of juggling systems [50], with increasing complexity of the "object"'s dynamics. The next step would be to extend to 2D and then 3D granular materials. Applications concern vibrators, screening, transport in mining and manufacturing processes.
- *Stability of structures:* our objective here is to study the stability of stacked blocks in 2D or 3D, and the influence on the observed behavior (numerically and/or analytically) of the contact/impact model.

3.4.4. Switching LCS and DAEs, higher-order sweeping process (HOSwP)

- We have gained a strong experience in the field of complementarity systems and distribution differential inclusions [33], [51], that may be seen as some kind of switching DAEs. We plan to go further with non-autonomous HOSwP with switching feedback inputs and non-uniform vector relative degrees. Switching linear complementarity systems can also be studied, though the exact relationships between both point of views remain unclear at the present time. This axis of research is closely related to cyber-physical systems in section 3.2.

3.4.5. Control of Elastic (Visco-plastic) systems with contact, impact and friction

- *Stabilization, trajectory tracking:* until now we have focused on the stability and the feedback control of systems of rigid bodies. The proposal here is to study the stabilization of flexible systems (for instance, a "simple" beam) subjected to unilateral contacts with or without set-valued friction (contacts with obstacles, or impacts with external objects line particle/beam impacts). This gives rise to varying (in time and space) boundary conditions. The best choice of a good contact law is a hard topic discussed in the literature.
- *Cableway systems (STRMTG, POMA):* cable-car systems present challenging control problems because they usually are underactuated systems, with large flexibilities and deformations. Simplified models of cables should be used (Ritz-Galerkin approach), and two main classes of systems may be considered: those with moving cable and only actuator at the station, and those with fixed cable but actuated nacelles. It is expected that they possess quite different control properties and thus deserve separate studies. The nonsmoothness arises mainly from the passage of the nacelles on the pylons, which induces frictional effects and impacts. It may certainly be considered as a nonsmooth set-valued disturbance within the overall control problem.

4. Application Domains

4.1. Applications Domains

Nonsmooth dynamical systems arise in a lot of application fields. *We briefly expose here some applications that have been treated in the BIPOP team and that we will continue in the TRIPOP team, as a validation for the research axes and also in terms of transfer (Sect. 7.1).* In mechanics, the main instances of nonsmooth dynamical systems are multibody systems with Signorini's unilateral contact, set-valued (Coulomb-like) friction and impacts, or in continuum mechanics, ideal plasticity, fracture or damage. Some illustrations are given in Figure 4(a-f). Other instances of nonsmooth dynamical systems can also be found in electrical circuits with ideal components (see Figure 4(g)) and in control theory, mainly with sliding mode control and variable structure systems (see Figure 4(h)). More generally, every time a piecewise, possibly set–valued, model of systems is invoked, we end up with a nonsmooth system. This is the case, for instance, for hybrid systems in nonlinear control or for piecewise linear modeling of gene regulatory networks in mathematical biology (see Figure 4(i)). Another common example of nonsmooth dynamics is also found when the vector field of a dynamical system is defined as a solution of an optimization problem under constraints, or a variational inequality. Examples of this kind are found in the optimal control theory, in dynamic Nash equilibrium or in the theory of dynamic flows over networks.

5. New Software and Platforms

5.1. Platforms: SICONOS

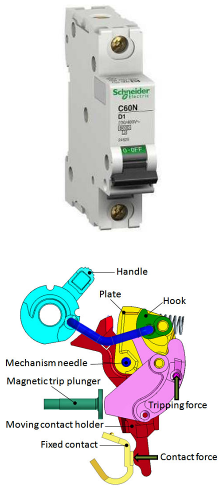
5.1.1. Platform A : SICONOS

Participants: Vincent Acary, Maurice Brémond, Olivier Huber, Franck Pérignon.

In the framework of the FP5 European project Siconos (2002-2006), Bipop was the leader of the Work Package 2 (WP2), dedicated to the numerical methods and the software design for nonsmooth dynamical systems. This has given rise to the platform SICONOS which is the main software development task in the team. The aim of this work is to provide a common platform for the simulation, modeling, analysis and control of abstract nonsmooth dynamical systems. Besides usual quality attributes for scientific computing software, we want to provide a common framework for various scientific fields, to be able to rely on the existing developments (numerical algorithms, description and modeling software), to support exchanges and comparisons of methods, to disseminate the know-how to other fields of research and industry, and to take into account the diversity of users (end-users, algorithm developers, framework builders) in building expert interfaces in Python and end-user front-end through Scilab.

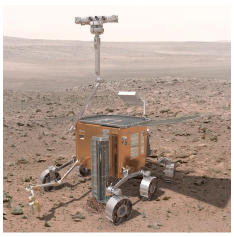
After the requirement elicitation phase, the Siconos Software project has been divided into 5 work packages which are identified to software products:

1. SICONOS/NUMERICS This library contains a set of numerical algorithms, already well identified, to solve non smooth dynamical systems. This library is written in low-level languages (C,F77) in order to ensure numerical efficiency and the use of standard libraries (Blas, Lapack, ...)
2. SICONOS/KERNEL This module is an object-oriented structure (C++) for the modeling and the simulation of abstract dynamical systems. It provides the users with a set of classes to describe their nonsmooth dynamical system (dynamical systems, intercations, nonsmooth laws, ...) and to perform a numerical time integration and solving.
3. SICONOS/FRONT-END. This module is mainly an auto-generated wrapper in Python which provides a user-friendly interface to the Siconos libraries. A scilab interface is also provided in the Front-End module.
4. SICONOS/CONTROL This part is devoted to the implementation of control strategies of non smooth dynamical systems.
5. SICONOS/MECHANICS. This part is dedicated to the modeling and the simulation of multi-body systems with 3D contacts, impacts and Coulomb's friction. It uses the Siconos/Kernel as simulation engine but relies on a industrial CAD library (OpenCascade and pythonOCC) to deal with complex body geometries and to compute the contact locations and distances between B-Rep description and on Bullet for contact detection between meshes.

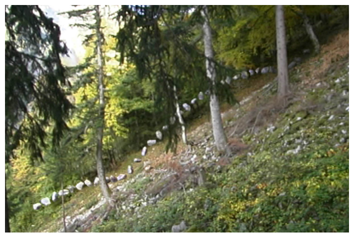


(a) Circuit breakers mechanisms [40]

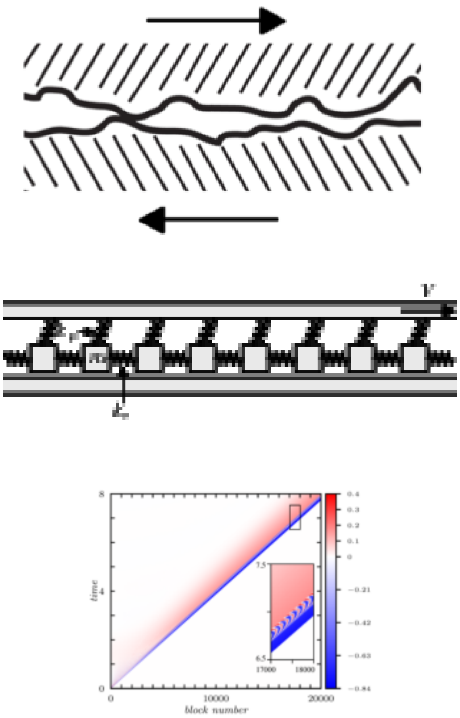
(b) Granular flows



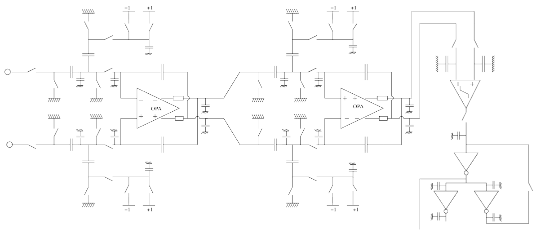
(c) Robots (ESA ExoMars Rover [35])



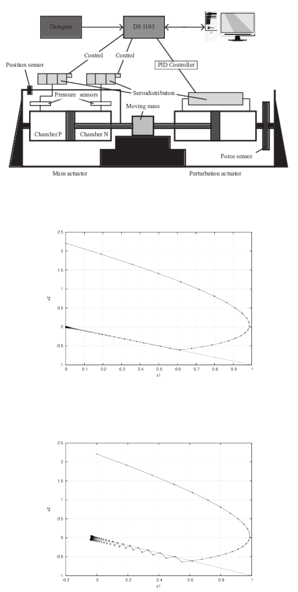
(d) Rockfall [47], [46], [59]



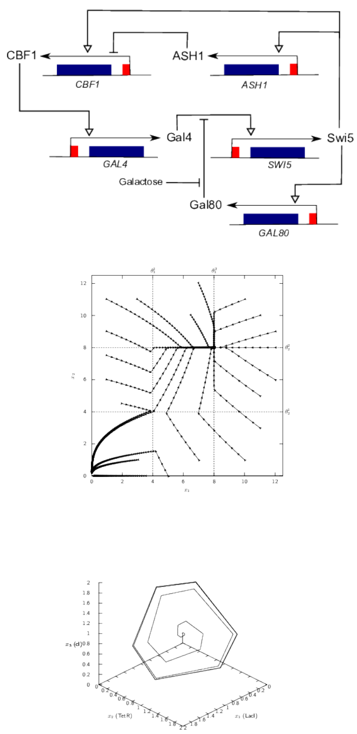
(e) Frictional interface and solitary waves in the Burridge-Knopoff model [81]



(a) Switched electrical circuits (delta-sigma converter) [29]



(b) Sliding mode control [32], [34], [70], [71], [80]



Further informations may be found at http://siconos.gforge.inria.fr/

6. New Results

6.1. Nonlinear waves in granular chains

Participants: Guillaume James, Bernard Brogliato, Kirill Vorotnikov.

Granular chains made of aligned beads interacting by contact (e.g. Newton's cradle) are widely studied in the context of impact dynamics and acoustic metamaterials. In order to describe the response of such systems to impacts or vibrations, it is important to analyze different wave effects such as the propagation of localized compression pulses (solitary waves) or oscillations (traveling breathers), or the scattering of vibrations through the chain. Such phenomena are strongly influenced by contact nonlinearities (Hertz force), spatial inhomogeneities and dissipation.

In the work [22], we analyze the Kuwabara-Kono (KK) model for contact damping, and we develop new approximations of this model which are efficient for the simulation of multiple impacts. The KK model is a simplified viscoelastic contact model derived from continuum mechanics, which allows for simpler calibration (using material parameters instead of phenomenological ones), but its numerical simulation requires a careful treatment due to its non-Lipschitzian character. Using different dissipative time-discretizations of the conservative Hertz model, we show that numerical dissipation can be tuned properly in order to reproduce the physical dissipation of the KK model and associated wave effects. This result is obtained analytically in the limit of small time steps (using methods from backward analysis) and is numerically validated for larger time steps. The resulting schemes turn out to provide good approximations of impact propagation even for relatively large time steps.

In reference [8], we analyze the discrete p-Schrödinger equation, an envelope equation that describes small oscillations in a Newton's cradle. In the limit when the exponent of the contact force lies slightly above unity, we derive three different continuum limits of the model which allow us to approximate the profiles of traveling breather solutions. One model consists of a logarithmic nonlinear Schrödinger equation which leads to a Gaussian approximation, and the two other are fully nonlinear degenerate Schrödinger equations which provide compacton approximations. These approximations are numerically validated by Newton-type computations. In the opposite (vibroimpact) limit when the exponent of the contact force is large, we obtain an analytical approximation of solitary waves in the form of a compacton.

6.2. Periodic motions of coupled impact oscillators

Participants: Guillaume James, Vincent Acary, Franck Pérignon.

In the work [17], we study the existence and stability of time-periodic oscillations in an infinite chain of linearly coupled impact oscillators, for rigid impacts without energy dissipation. We reformulate the search of periodic solutions as a boundary value problem incorporating unilateral constraints. This formulation, together with an appropriate notion of nondegenerate modes, allows us to construct nonsmooth modes of oscillations (spatially localized or extended) when the oscillators are weakly coupled (this approach is an adaptation of the idea of 'anticontinuum' limit to the nonsmooth setting). In this framework, we show the existence of exact solutions (in particular, we check the condition of non-penetration of the obstacle) for an arbitrary number of impacting particles. Different solution branches corresponding to stable or unstable breathers, multibreathers and nonsmooth normal modes are found. We provide a formula for the monodromy matrix that determines spectral stability of nonsmooth modes in the presence of simple impacts. These results are completed by a numerical computation of the time-periodic solutions at larger coupling, and the Siconos software is used to simulate the system and explore dynamical instabilities. The above approach is much more effective than numerical continuation of periodic solutions based on stiff compliant models, which leads to stiff ODEs and costly numerical continuation.

6.3. Solitary waves in the excitable Burridge-Knopoff model

Participants: Guillaume James, Jose Eduardo Morales Morales, Arnaud Tonnelier.

The Burridge-Knopoff model is a lattice differential equation describing a chain of blocks connected by springs and pulled over a surface. This model was originally introduced to investigate nonlinear effects arising in the dynamics of earthquake faults. One of the main ingredients of the model is a nonlinear velocity-dependent friction force between the blocks and the fixed surface. We introduce a simplified piecewise linear friction law (reminiscent of the McKean nonlinearity for excitable cells) which allows us to obtain analytical expression of solitary waves and study some of their qualitative properties, such as wavespeed and propagation failure. These results have been published in [11].
We have obtained an existence theorem for solitary waves in the Burridge-Knopoff model. Our approach uses a piecewise-linear friction force combined with a weak coupling strength. Using asymptotic arguments, we show that trial solutions, obtained semi-analytically, satisfy, for some parameter set, the inequality constraints associated with the threshold conditions. An approximation of the wave profile is obtained and a minimal wave speed is derived.

6.4. Signal propagation along excitable chains

Participant: Arnaud Tonnelier.

Nonlinear self-sustained waves, or *autowaves*, have been identified in a large class of discrete excitable media. We have proposed a simple continuous-time threshold model for wave propagation in excitable media. The ability of the resulting transmission line to convey a one-bit signal is investigated. Existence and multistability of signals where two successive units share the same waveform is established. We show that, depending on the connectivity of the transmission line, an arbitrary number of distinct signals can be transmitted. More precisely, we prove that, for a one-dimensional information channel with $n\mathrm{th}$-neighbor interactions, a n-fold degeneracy of the speed curve induces the coexistence of $2n$ propagating signals, n of which are stable and allow n distinct symbols transmission. The influence of model parameters (time constants, coupling strength and connectivity) on the traveling signal properties is analyzed. This work is almost finished and is going to be submitted.

6.5. Numerical analysis of multibody mechanical systems with constraints

This scientific theme concerns the numerical analysis of mechanical systems with bilateral and unilateral constraints, with or without friction [1]. They form a particular class of dynamical systems whose simulation requires the development of specific methods for analysis and dedicated simulators [6].

6.5.1. Multibody systems with clearances (dynamic backlash)

Participants: Vincent Acary, Bernard Brogliato.

The PhD thesis of N. Akadkhar under contract with Schneider Electric has concerned the numerical simulation of mechanical systems with unilateral constraints and friction, where the presence of clearances in imperfect joints plays a crucial role. A first work deals with four-bar planar mechanisms with clearances at the joints, which induce unilateral constraints and impacts, rendering the dynamics nonsmooth. The objective is to determine sets of parameters (clearance value, restitution coefficients, friction coefficients) such that the system's trajectories stay in a neighborhood of the ideal mechanism (*i.e.* without clearance) trajectories. The analysis is based on numerical simulations obtained with the projected Moreau-Jean time-stepping scheme. Circuits breakers with 3D joint clearances have been studied in [3] [41] where it is demonstrated that the nonsmooth dynamics approach as coded in our software SICONOS, allows a very good prediction of the system's dynamics, with experimental validation. An overview of various approaches for the feedback control of multibody systems with joint clearances is proposed in [4].

6.5.2. Generalized-α scheme for nonsmooth multibody systems.

Participant: Vincent Acary.

This work [16] concerns a formalism for the transient simulation of nonsmooth dynamic mechanical systems composed of rigid and flexible bodies, kinematic joints and frictionless contact conditions. The proposed algorithm guarantees the exact satisfaction of the bilateral and unilateral constraints both at position and velocity levels. Thus, it significantly differs from penalty techniques since no penetration is allowed. The numerical scheme is obtained in two main steps. Firstly, a splitting method is used to isolate the contributions of impacts, which shall be integrated with only first-order accuracy, from smooth contributions which can be integrated using a higher order scheme. Secondly, following the idea of Gear, Gupta and Leimkuhler, the equation of motion is reformulated so that the bilateral and unilateral constraints appear both at position and velocity levels. After time discretization, the equation of motion involves two complementarity conditions and it can be solved at each time step using a monolithic semi-smooth Newton method. The numerical behaviour of the proposed method is studied and compared to other approaches for a number of numerical examples. Itis shown that the formulation offers a unified and valid approach for the description of contactconditions between rigid bodies as well as between flexible bodies.

6.5.3. Mechanics of musical instruments with contact and impacts.

Participants: Vincent Acary, Franck Pérignon.

Collisions in musical string instruments play a fundamental role in explaining the sound production in various instruments such as sitars, tanpuras and electric basses. Contacts occuring during the vibration provide a nonlinear effect which shapes a specific tone due to energy transfers and enriches the hearing experience. As such, they must be carefully simulated for the purpose of physically-based sound synthesis. Most of the numerical methods presented in the literature rely on a compliant modeling of the contact force between the string and the obstacle. In this contribution, numerical methods from nonsmooth contact dynamics are used to integrate the problem in time. A Moreau-Jean time-stepping scheme is combined with an exact scheme for phases with no contact, thus controlling the numerical dispersion. Results for a two-point bridge mimicking a tanpura and an electric bass are presented, showing the ability of the method to deal efficiently with such problems while invoking, as compared to a compliant approach, less modelling parameters and a reduced computational burden [7].

6.5.4. Numerical solvers for frictional contact problems.

Participants: Vincent Acary, Maurice Brémond.

In [15] report, we review several formulations of the discrete frictional contact problem that arises in space and time discretized mechanical systems with unilateral contact and three-dimensional Coulomb's friction. Most of these formulations are well–known concepts in the optimization community, or more generally, in the mathematical programming community. To cite a few, the discrete frictional contact problem can be formulated as variational inequalities, generalized or semi–smooth equations, second–order cone complementarity problems, or as optimization problems such as quadratic programming problems over second-order cones. Thanks to these multiple formulations, various numerical methods emerge naturally for solving the problem. We review the main numerical techniques that are well-known in the literature and we also propose new applications of methods such as the fixed point and extra-gradient methods with self-adaptive step rules for variational inequalities or the proximal point algorithm for generalized equations. All these numerical techniques are compared over a large set of test examples using performance profiles. One of the main conclusion is that there is no universal solver. Nevertheless, we are able to give some hints to choose a solver with respect to the main characteristics of the set of tests

6.5.4.1. Impact laws in chains of aligned balls

In [18] several "classical" multiple-impact laws are compared on chains of aligned balls: Moreau's law, the binary collision law, and the LZB approach [2]. Short analyses of these laws are made, and thorough comparisons are led numerically. It is concluded that both Moreau and the binary collision laws, furnish good results (in terms of predictability) only in very particular cases of elasticity coefficien, contact stiffenesses ratios, and mass ratios.

6.6. Analysis and Control of Set-Valued Systems

Participants: Bernard Brogliato, Christophe Prieur, Alexandre Vieira.

6.6.1. Higher-order sweeping process

This work [5] continues our previous results in [33], to the case when exogeneous terms are present in both the unilateral constraint, and in the dynamics. A suitable change of state variables allows one to recast the dynamics in a format that is close to the autonomous case, so that the well-posedness issues (existence and uniqueness of solutions) is shown (see the preprint [20] for a complete analysis, which in fact differs only slightly from the original one in [33]). The link with switching DAEs is made.

6.6.2. Robust sliding-mode control: continuous and discrete-time

This work [10] concerns the robust control of linear time-invariant systems, subjected to nonlinear varying state dependent disturbances as well as parameter uncertainties. A specific set-valued class of sliding-mode controllers is designed, and its discretization (with the implicit method introduced in [32], [34]) is analysed. One difficulty is that the parameter uncertainties, as well as the discretization, create unmatched disturbances. Stability and convergence results are proved. Let us mention also [9] that corrects a slight mistake in [80]. In the same way it is worth citing [14], [14] which continues the analysis of the implicit discretization of set-valued systems, this time oriented towards the consistency of time-discretizations for homogeneous systems, with one discontinuity at zero (sometimes called quasi-continuous, strangely enough).

6.6.3. Evolution variational inequalities

In [13] we continue our previous works on well-posedness and stabilization/control of a class of set-valued systems, that take the form of evolution variational inequalities. Dissipativity is then a key property. Regulation with state and output feedback, viability issues, are solved, with absolutely continuous and bounded variations solutions. Applications are in power converters.

6.6.4. Optimal control of LCS

The quadratic and minimum time optimal control of LCS as in (6) is tackled in [24], [25]. This work relies on the seminal results by Guo and ye (SIAM 2016), and aims at particularizing their results for LCS, so that they become numerically tractable and one can compute optimal controllers and optimal trajectories. The basic idea is to take advantage of the complementarity, to construct linear complementarity problems in the Pontryagin's necessary conditions which can then be integrated numerically, without having to guess a priori the switching instants (the optimal controller can eb discontinuous and the optimal trajectories can visit several modes of the complementarity conditions).

7. Bilateral Contracts and Grants with Industry

7.1. Schneider Electric

This action started in 2001 with my post-doc co–supported by Schneider Electric and CNRS. With some brief interruptions, this action is still active and should further continue. It concerns mainly the simulation and modeling of multi–body systems with contact, friction and impacts with the application for the virtual prototyping of electrical circuit breakers. During these years, various forms of collaborations have been held. Two PhD thesis have been granted by Schneider Electric (D.E. Taha and N. Akhakdar) accompanied with research contracts between Inria and Schneider Electric. Schneider Electric participated also the ANR project Saladyn as a main partner. Without going into deep details of the various actions over the years, the major success of this collaboration is the statistical tolerance analysis of the functional requirements of the circuit breakers with respect to clearance in joints and geometrical tolerances on the parts. Starting from the geometrical descriptions (CAD files) of a mechanism with prescribed tolerances on the manufacturing process, we perform worst-case analysis and Monte–Carlo simulations of the circuit breaker with Siconos and we record the variations in the functional requirements. The difficulty in such simulations are the modeling of contact with friction that models the joints with clearances. The results of these analysis enable Schneider Electric to define the manufacturing precision that has a huge impact of the production cost (Schneider Electric produces several millions of C60-type circuit breaker per year). Note that it is not possible to perform such simulations with the existing software codes of the market. At the beginning, our interlocutor at Schneider Electric was the innovation (R&D) department. Now, we are working and discussing with the business unit, Division Power and Dinnov (M. Abadie, E. Boumediene, X. Herreros) in charge of designing and producing the circuit–breakers. The targeted users are the R&D engineers of Schneider Electric that use simulation tools for designing new models or improving existing circuit breakers. This collaboration continues with new modeling and simulation challenges (flexible parts, multiple impact laws) with the CIFRE PhD of Rami Sayoud.

7.2. STRMTG

We have recently started with STRMTG a research contract about modelling, simulation and control of cable-transport systems. In such systems, the question of the coupling between the nonlinear dynamics of cables and their supports with unilateral contact and friction appears now to be determinant in order to increase the performances of the cableway systems, especially for urban transportation systems.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

B. Brogliato coordinates the ANR project Digitslid (PRC, ANR-18-CE40-0008-01), Differentiateurs et commandes homogenes par modes glissants en temps discret: l'approche implicite. Partners: LS2N (Ecole Centrale de Nantes), Inria Lille Nord Europe (team Non-A-Post), and Tripop. October 2018-September 2021. 12 participants overall (3 post-doc students recruited by the project, 3 Ph.D. students supported by other means). Total financial support by the ANR: 338 362 euros (100 762 for Tripop, 18 months of post-doc to be recruited in 2019).

8.1.2. FUI Modeliscale.

https://team.inria.fr/modeliscale/

The ModeliScale FUI focuses on the modeling, simulation and analysis of large cyber-physical systems. It federates the research activities of several teams, covering a broad spectrum of topics, namely hybrid systems modeling & verification, numerical analysis, programming language design and automatic control. Our research agenda includes the following tracks:

- New compilation techniques for Modelica modelers: structural analysis of multimode DAE (Differential Algebraic Equations) systems, modular compilation, combining state-machines and non-smooth dynamical systems (complementarity dynamical systems and Filippov differential inclusions), contract-based specification of cyber-physical systems requirements, requirements capture using under-/over-determined DAE systems.
- Simulation of large cyber-physical systems: distributed simulation, discretization methods for non-smooth dynamical systems, space-/time-adaptive discretization methods for multimode DAE systems, quantized state solvers (QSS).
- Guaranteed numerics: guaranteed simulation of non-smooth and hybrid dynamical systems, numerical methods preserving invariant properties of hybrid systems, contract-based reasoning methods.

Table 1. Member of IPL Modeliscale

Name	Team	Inria Center or Laboratory
Vincent Acary	Bipop	Inria Grenoble Rhône Alpes
Bernard Brogliato		
Albert Benveniste	Hycomes Inria Rennes	
Benoît Caillaud		Bretagne Atlantique
Khalil Ghorbal		
Marc Pouzet	Parkas	ENS
Tim Bourke		Inria Paris
Goran Frehse	Tempo	Verimag-univ. Grenoble Alpes
Antoine Girard		L2S-CNRS, Saclay
Eric Goubault	Cosynus	LIX, École Polytechnique,
Sylvie Putot		Saclay

8.1.3. Inria Project Lab (IPL): ModeliScale, Languages and Compilation for Cyber-Physical System Design

https://team.inria.fr/modeliscale/

The project gathers researchers from three Inria teams, and from three other research labs in Grenoble and Paris area.

The main objective of ModeliScale is to advance modeling technologies (languages, compile-time analyses, simulation techniques) for CPS combining physical interactions, communication layers and software components. We believe that mastering CPS comprising thousands to millions of components requires radical changes of paradigms. For instance, modeling techniques must be revised, especially when physics is involved. Modeling languages must be enhanced to cope with larger models. This can only be done by combining new compilation techniques (to master the structural complexity of models) with new mathematical tools (new numerical methods, in particular).

ModeliScale gathers a broad scope of experts in programming language design and compilation (reactive synchronous programming), numerical solvers (nonsmooth dynamical systems) and hybrid systems modeling and analysis (guaranteed simulation, verification). The research program is carried out in close cooperation with the Modelica community as well as industrial partners, namely, Dassault Systèmes as a Modelica/FMI tool vendor, and EDF and Engie as end users.

8.2. International Research Visitors

8.2.1. Visits of International Scientists

- Mathias Legrand (McGill University, Mechanical Engineering).

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Journal

9.1.1.1. Member of the Editorial Boards

- B. Brogliato is Associate Editor at Nonlinear Analysis: Hybrid Systems.
- B. Brogliato is Associate Editor at ASME Journal of Computational and Nonlinear Dynamics.

9.1.1.2. Reviewer - Reviewing Activities

- A. Tonnelier has been reviewer for PRE (Physical Review E) and DCDS (Discrete and Continuous Dynamical Systems).
- G. James has been reviewer for IMA J. Appl. Math.
- V. Acary has been for Optimization Methods and Software, Nonlinear Dynamics, Applied Mechanics Reviews, IEEE Transactions on Robotics, Multibody Systems Dynamics, International Journal for Numerical Methods in Engineering, IFAC Conference on Modelling, Identification and Control of Nonlinear Systems, CDC 2018, SIMPAR 2018.
- B. Brogliato has been reviewer for IEEE Transactions on Automatic Control, IEEE Transactions on Robotics, IEEE Control Systems Letters, Automatica, SIAM Journal on Control and Optimization, Multibody System Dynamics.

9.1.2. Invited Talks

- G. James : invited talk at the International Symposium on Intrinsic Localized Modes (Kyoto, 01/18), speaker at the Conference on Advances in Nonsmooth Mechanics (Bristol, 06/18), lecture at the thematic school Méthodes de Dynamique Non-Linéaire pour l'Ingénierie des Structures, GDR Dynolin (Fréjus, 05/18).
- V. Acary : lecture at the thematic school Méthodes de Dynamique Non-Linéaire pour l'Ingénierie des Structures, GDR Dynolin (Fréjus, 05/18).

9.1.3. Leadership within the Scientific Community

- V. Acary is coordinator with R. Leine of the Europe Network for Nonsmooth Dynamics http://ennsd.gforge.inria.fr/.

9.1.4. Research Administration

- A. Tonnelier is member of the CED and the CLHSCT.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Licence : G. James, Introduction to Dynamical Systems in UE MAP201, 18 hETD, L1, Université Grenoble Alpes.
- Licence : G. James, Normed Vector Spaces, 25 hETD, L2, Prépa INP, Grenoble.
- Master : G. James, Numerical Methods, 94 hETD, M1, Grenoble INP - Ensimag (1st year).
- Master : G. James, Dynamical Systems, 45 hETD, M1, Grenoble INP - Ensimag (2nd year).
- Master : Vincent Acary, 17H éq TD Systèmes dynamiques, ENSIMAG 2A.
- Master : Vincent Acary, 12H éq TD Systèmes dynamiques, Master ACSYON. Université de Limoges.

9.2.2. Supervision

Rami Sayoub, Influence of vibrations on multibody systems, September 2018, V. Acary and B. Brogliato.

Charlélie Bertrand, Mechanical model fro cable vibrations, October 2018, V. Acary and C.H. Lamarque.

Christelle Kozaily. Structural analysis for multi-mode DAE systems, Octobre 2018, V. Acary and B. Caillaud.

Alexandre Vieira. Commande optimale de systèmes linéaires de complémentarité, université Grenoble Alpes, 24 septembre 2018, B. Brogliato and C. Prieur.

9.2.3. Juries

- Vincent Acary, president of Ph.D. Thesis committee of Thomas Catterou (22 October 2018), Universite Aix–Marseille.
- Bernard Brogliato, examinateur of Ph.D. thesis of C. Beneux (12 July 2018), Université Lorraine.

10. Bibliography

Major publications by the team in recent years

[1] V. ACARY, B. BROGLIATO.*Numerical methods for nonsmooth dynamical systems. Applications in mechanics and electronics*, Lecture Notes in Applied and Computational Mechanics 35. Berlin: Springer. xxi, 525 p. , 2008

[2] B. BROGLIATO.*Nonsmooth mechanics*, Communications and Control Engineering Series, Third, Springer, [Cham], 2016, xxii+629, Models, dynamics and control, http://dx.doi.org/10.1007/978-3-319-28664-8

Publications of the year

Articles in International Peer-Reviewed Journal

[3] N. AKHADKAR, V. ACARY, B. BROGLIATO.*Multibody systems with 3D revolute joints with clearances: an industrial case study with an experimental validation*, in "Multibody System Dynamics", March 2018, vol. 42, n° 3, p. 249–282 [*DOI :* 10.1007/s11044-017-9584-5], https://hal.inria.fr/hal-01562703

[4] B. BROGLIATO.*Feedback control of multibody systems with joint clearance and dynamic backlash: a tutorial*, in "Multibody System Dynamics", March 2018, vol. 42, n° 3, p. 283–315 [*DOI :* 10.1007/s11044-017-9585-4], https://hal.inria.fr/hal-01499581

[5] B. BROGLIATO.*Non-autonomous higher-order Moreau's sweeping process: Well-posedness, stability and Zeno trajectories*, in "European Journal of Applied Mathematics", October 2018, vol. 29, n° 5, p. 941-968, Special issue Theory and applications of nonsmooth dynamical systems [*DOI :* 10.1017/S0956792518000256], https://hal.inria.fr/hal-01893796

[6] F. DUBOIS, V. ACARY, M. JEAN.*The Contact Dynamics method: A nonsmooth story* , in "Comptes Rendus Mécanique", March 2018, vol. 346, n° 3, p. 247-262 [*DOI :* 10.1016/j.crme.2017.12.009], https://hal.archives-ouvertes.fr/hal-01676287

[7] C. ISSANCHOU, V. ACARY, F. PÉRIGNON, C. TOUZÉ, J.-L. LE CARROU.*Nonsmooth contact dynamics for the numerical simulation of collisions in musical string instruments*, in "Journal of the Acoustical Society of America", May 2018, vol. 143, n° 5, p. 1-13 [*DOI :* 10.1121/1.5039740], https://hal-ensta.archives-ouvertes.fr/hal-01806133

[8] G. JAMES.*Travelling breathers and solitary waves in strongly nonlinear lattices*, in "Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences", July 2018, vol. 376, n° 2127, p. 20170138:1-25 [*DOI :* 10.1098/rsta.2017.0138], https://hal.inria.fr/hal-01951239

[9] F. A. MIRANDA-VILLATORO, B. BROGLIATO, F. CASTAÑOS.*Errata to "Multivalued Robust Tracking Control of Lagrange Systems: Continuous and Discrete-Time Algorithms" [Sep 17 4436-4450]*, in "IEEE Transactions on Automatic Control", August 2018, vol. 63, n° 8, 2750 [*DOI :* 10.1109/TAC.2018.2827992], https://hal.inria.fr/hal-01927000

[10] F. MIRANDA-VILLATORO, B. BROGLIATO, F. CASTAÑOS.*Set-valued sliding-mode control of uncertain linear systems: continuous and discrete-time analysis*, in "SIAM Journal on Control and Optimization", 2018, vol. 56, n$^{\text{o}}$ 3, p. 1756-1793 [*DOI :* 10.1137/16M1077362], https://hal.inria.fr/hal-01317948

[11] J. E. MORALES, G. JAMES, A. TONNELIER.*Solitary waves in the excitable Burridge-Knopoff model*, in "Wave Motion", January 2018, vol. 76, p. 103-121 [*DOI :* 10.1016/J.WAVEMOTI.2017.10.001], https://hal.archives-ouvertes.fr/hal-01634470

[12] A. POLYAKOV, D. EFIMOV, B. BROGLIATO.*Consistent Discretization of Finite-time and Fixed-time Stable Systems*, in "SIAM Journals", 2018, p. 1-28, https://hal.inria.fr/hal-01838712

[13] A. TANWANI, B. BROGLIATO, C. PRIEUR.*Well-Posedness and Output Regulation for Implicit Time-Varying Evolution Variational Inequalities*, in "SIAM Journal on Control and Optimization", March 2018, vol. 56, n$^{\text{o}}$ 2, p. 751 - 781 [*DOI :* 10.1137/16M1083657], https://hal.archives-ouvertes.fr/hal-01360325

Invited Conferences

[14] A. POLYAKOV, D. EFIMOV, B. BROGLIATO.*Consistent Discretization of Finite-time Stable Homogeneous Systems*, in "VSS 2018 - 15th International Workshop on Variable Structure Systems and Sliding Mode Control", Graz, Austria, IEEE, July 2018, p. 360-365 [*DOI :* 10.1109/VSS.2018.8460306], https://hal.inria.fr/hal-01794386

Scientific Books (or Scientific Book chapters)

[15] V. ACARY, M. BRÉMOND, O. HUBER.*On solving contact problems with Coulomb friction: formulations and numerical comparisons*, in "Advanced Topics in Nonsmooth Dynamics - Transactions of the European Network for Nonsmooth Dynamics", Springer International Publishing, June 2018, p. 375-457 [*DOI :* 10.1007/978-3-319-75972-2_10], https://hal.inria.fr/hal-01878539

[16] O. BRÜLS, V. ACARY, A. CARDONA.*On the Constraints Formulation in the Nonsmooth Generalized-α Method*, in "Advanced Topics in Nonsmooth Dynamics. Transactions of the European Network for Nonsmooth Dynamics", Springer International Publishing, June 2018, p. 335-374 [*DOI :* 10.1007/978-3-319-75972-2_9], https://hal.inria.fr/hal-01878550

[17] G. JAMES, V. ACARY, F. PÉRIGNON.*Periodic motions of coupled impact oscillators*, in "Advanced Topics in Nonsmooth Dynamics, Transactions of the European Network for Nonsmooth Dynamics", Springer International Publishing, 2018, 457, https://hal.archives-ouvertes.fr/hal-01660891

[18] N. S. NGUYEN, B. BROGLIATO.*Comparisons of multiple-impact laws for multibody systems: Moreau's law, binary impacts, and the LZB approach*, in "Advanced Topics in Nonsmooth Dynamics. Transactions of the European Network for Nonsmooth Dynamics", Advanced Topics in Nonsmooth Mechanics. Transactions of the European Network for Nonsmooth Dynamics, Springer International Publishing AG, June 2018, p. 1-45 [*DOI :* 10.1007/978-3-319-75972-2_1], https://hal.inria.fr/hal-01932551

Research Reports

[19] B. BROGLIATO.*Nonsmooth Mechanics. Models, Dynamics and Control : Erratum/Addendum*, Inria Grenoble - Rhone-Alpes, June 2018, https://hal.inria.fr/hal-01331565

Other Publications

[20] B. BROGLIATO.*Non-autonomous higher-order Moreau's sweeping process: Well-posedness, stability and Zeno trajectories*, March 2018, working paper or preprint, https://hal.inria.fr/hal-01722772

[21] A. FERRANTE, A. LANZON, B. BROGLIATO.*A direct proof of the equivalence of side conditions for strictly positive real matrix transfer functions*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01947938

[22] G. JAMES, K. VOROTNIKOV, B. BROGLIATO.*Kuwabara-Kono numerical dissipation: a new method to simulate granular matter*, September 2018, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01878973

[23] F. MIRANDA-VILLATORO, F. CASTAÑOS, B. BROGLIATO.*Continuous and discrete-time stability of a robust set-valued nested controller*, December 2018, working paper or preprint, https://hal.inria.fr/hal-01944920

[24] A. VIEIRA, B. BROGLIATO, C. PRIEUR.*Optimality conditions for the minimal time problem for Complementarity Systems*, August 2018, working paper or preprint, https://hal.inria.fr/hal-01856054

[25] A. VIEIRA, B. BROGLIATO, C. PRIEUR.*Quadratic Optimal Control of Linear Complementarity Systems: First order necessary conditions and numerical analysis*, April 2018, working paper or preprint, https://hal.inria.fr/hal-01690400

References in notes

[26] V. ACARY.*Higher order event capturing time–stepping schemes for nonsmooth multibody systems with unilateral constraints and impacts.*, in "Applied Numerical Mathematics", 2012, vol. 62, p. 1259-1275 [*DOI :* 10.1016/J.APNUM.2012.06.026], http://hal.inria.fr/inria-00476398

[27] V. ACARY.*Projected event-capturing time-stepping schemes for nonsmooth mechanical systems with unilateral contact and Coulomb's friction*, in "Computer Methods in Applied Mechanics and Engineering", 2013, vol. 256, p. 224 - 250 [*DOI :* 10.1016/J.CMA.2012.12.012], http://www.sciencedirect.com/science/article/pii/S0045782512003829

[28] V. ACARY.*Energy conservation and dissipation properties of time-integration methods for the nonsmooth elastodynamics with contact*, in "Zeitschrift für Angewandte Mathematik und Mechanik", 2016, vol. 95, n^{o} 5, p. 585–603

[29] V. ACARY, O. BONNEFON, B. BROGLIATO.*Nonsmooth modeling and simulation for switched circuits.*, Lecture Notes in Electrical Engineering 69. Dordrecht: Springer. xxiii, 284 p., 2011, http://dx.doi.org/10.1007/978-90-481-9681-4

[30] V. ACARY, M. BRÉMOND, O. HUBER.*On solving frictional contact problems: formulations and comparisons of numerical methods.*, in "Advanced Topics in Nonsmooth Dynamics", Acary, V. and Brüls. O. and Leine, R. (eds). Springer Verlag, 2018, To appear

[31] V. ACARY, B. BROGLIATO.*Numerical methods for nonsmooth dynamical systems. Applications in mechanics and electronics.*, Lecture Notes in Applied and Computational Mechanics 35. Berlin: Springer. xxi, 525 p. , 2008

[32] V. ACARY, B. BROGLIATO.*Implicit Euler numerical scheme and chattering-free implementation of sliding mode systems*, in "Systems and Control Letters.", 2010, vol. 59, n^{o} 5, p. 284-295, doi:10.1016/j.sysconle.2010.03.002

[33] V. ACARY, B. BROGLIATO, D. GOELEVEN.*Higher order Moreau's sweeping process: mathematical formulation and numerical simulation*, in "Mathematical Programming Ser. A", 2008, vol. 113, p. 133-217

[34] V. ACARY, B. BROGLIATO, Y. ORLOV.*Chattering-Free Digital Sliding-Mode Control With State Observer and Disturbance Rejection*, in "IEEE Transactions on Automatic Control", may 2012, vol. 57, n^{o} 5, p. 1087 -1101, http://dx.doi.org/10.1109/TAC.2011.2174676

[35] V. ACARY, M. BRÉMOND, K. KAPELLOS, J. MICHALCZYK, R. PISSARD-GIBOLLET.*Mechanical simulation of the Exomars rover using Siconos in 3DROV*, in "ASTRA 2013 - 12th Symposium on Advanced Space Technologies in Robotics and Automation", Noordwijk, Netherlands, ESA/ESTEC, May 2013, http://hal.inria.fr/hal-00821221

[36] V. ACARY, M. BRÉMOND, T. KOZIARA, F. PÉRIGNON.*FCLIB: a collection of discrete 3D Frictional Contact problems*, Inria, February 2014, n^{o} RT-0444, 34, https://hal.inria.fr/hal-00945820

[37] V. ACARY, H. DE JONG, B. BROGLIATO.*Numerical Simulation of Piecewise-Linear Models of Gene Regulatory Networks Using Complementarity Systems Theory*, in "Physica D", January 2013, vol. 269, p. 103–199

[38] V. ACARY, Y. MONERIE.*Nonsmooth fracture dynamics using a cohesive zone approach*, Inria, 2006, n^{o} RR-6032, 56, http://hal.inria.fr/inria-00110560/en/

[39] S. ADLY.*A Variational Approach to Nonsmooth Dynamics. Applications in Unilateral Mechanics and Electronics*, SpringerBriefs in Mathematics, Springer Verlag, 2017

[40] N. AKHADKAR, V. ACARY, B. BROGLIATO.*Multibody systems with 3D revolute joint clearance, modelling, numerical simulation and experimental validation: an industrial case study*, in "Multibody System Dynamics", 2017

[41] N. AKHADKAR, V. ACARY, B. BROGLIATO.*3D Revolute Joint with Clearance in Multibody Systems*, in "CK 2017 - 7th International Workshop on Computational Kinematics", Poitiers, France, Springer, May 2017, p. 11-18 [*DOI :* 10.1007/978-3-319-60867-9_2], https://hal.archives-ouvertes.fr/hal-01625733

[42] R. ALUR, C. COURCOUBETIS, N. HALBWACHS, T. HENZINGER, P. HO, X. NICOLLIN, A. OLIVERO, J. SIFAKIS, S. YOVINE.*The algorithmic analysis of hybrid systems*, in "Theoretical Computer Science", 1995, vol. 138, n^{o} 1, p. 3–34

[43] M. ARNOLD, O. BRÜLS, A. CARDONA.*Error analysis of generalized-α Lie group time integration methods for constrained mechanical systems*, in "Numerische Mathematik", Jan 2015, vol. 129, n^{o} 1, p. 149–179, https://doi.org/10.1007/s00211-014-0633-1

[44] A. BECK, M. TEBOULLE.*A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems*, in "SIAM Journal on Imaging Sciences", 2009, vol. 2, n^{o} 1, p. 183-202, http://dx.doi.org/10.1137/080716542

[45] V. N. BIKTASHEV, M. A. TSYGANOV.*Solitary waves in excitable systems with cross-diffusion*, in "Proc. R. Soc. A", 2005, vol. 461, p. 3711-3730

[46] F. BOURRIER, F. BERGER, P. TARDIF, L. DORREN, O. HUNGR.*Rockfall rebound: comparison of detailed field experiments and alternative modelling approaches*, in "Earth Surface Processes and Landforms", 2012, vol. 37, n^o 6, p. 656–665, http://dx.doi.org/10.1002/esp.3202

[47] F. BOURRIER, L. DORREN, F. NICOT, F. BERGER, F. DARVE.*Toward objective rockfall trajectory simulation using a stochastic impact model*, in "Geomorphology", 2009, vol. 110, n^o 3, p. 68 - 79 [*DOI :* 10.1016/J.GEOMORPH.2009.03.017], http://www.sciencedirect.com/science/article/pii/S0169555X09001251

[48] B. BROGLIATO.*Feedback control of multibody systems with joint clearance and dynamic backlash: a tutorial*, in "Multibody System Dynamics", Aug 2017, https://doi.org/10.1007/s11044-017-9585-4

[49] B. BROGLIATO, A. POLYAKOV.*Globally stable implicit Euler time-discretization of a nonlinear single-input sliding-mode control system*, in "2015 54th IEEE Conference on Decision and Control (CDC)", Dec 2015, p. 5426-5431, http://dx.doi.org/10.1109/CDC.2015.7403069

[50] B. BROGLIATO, A. Z. RIO.*On the control of complementary-slackness juggling mechanical systems*, in "IEEE Transactions on Automatic Control", Feb 2000, vol. 45, n^o 2, p. 235-246, http://dx.doi.org/10.1109/9.839946

[51] B. BROGLIATO, L. THIBAULT.*Existence and uniqueness of solutions for non-autonomous complementarity dynamical systems*, in "J. Convex Anal.", 2010, vol. 17, n^o 3-4, p. 961–990

[52] O. BRÜLS, V. ACARY, A. CARDONA.*Simultaneous enforcement of constraints at position and velocity levels in the nonsmooth generalized-α scheme*, in "Computer Methods in Applied Mechanics and Engineering", November 2014, vol. 281, p. 131-161 [*DOI :* 10.1016/J.CMA.2014.07.025], http://hal.inria.fr/hal-01059823

[53] B. CAILLAUD.*Hybrid vs. nonsmooth dynamical systems*, 2014, synchron, http://synchron2014.inria.fr/wp-content/uploads/sites/13/2014/12/Caillaud-nsds.pdf

[54] G. CAPOBIANCO, S. R. EUGSTER.*Time finite element based Moreau-type integrators*, in "International Journal for Numerical Methods in Engineering", 2018, vol. 114, n^o 3, p. 215-231, https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.5741

[55] L. CARLONI, R. PASSERONE, A. PINTO, A. SANGIOVANNI–VICENTELLI.*Languages and tools for hybrid systems design*, in "Foundations and Trends in Electronic Design Automation", 2006, vol. 1, n^o 1/2, p. 1–193

[56] Q. Z. CHEN, V. ACARY, G. VIRLEZ, O. BRÜLS.*A nonsmooth generalized-α scheme for flexible multibody systems with unilateral constraints*, in "International Journal for Numerical Methods in Engineering", 2013, vol. 96, n^o 8, p. 487–511, http://dx.doi.org/10.1002/nme.4563

[57] R. W. COTTLE, J. PANG, R. E. STONE.*The Linear Complementarity Problem*, Academic Press, Inc., Boston, MA, 1992

[58] L. DORREN, F. BERGER, M. JONSSON, M. KRAUTBLATTER, M. MÖLK, M. STOFFEL, A. WEHRLI.*State of the art in rockfall – forest interactions*, in "Schweizerische Zeitschrift fur Forstwesen", 2007, vol. 158, n° 6, p. 128-141, https://doi.org/10.3188/szf.2007.0128

[59] S. DUPIRE, F. BOURRIER, J.-M. MONNET, S. BIGOT, L. BORGNIET, F. BERGER, T. CURT.*Novel quantitative indicators to characterize the protective effect of mountain forests against rockfall*, in "Ecological Indicators", 2016, vol. 67, p. 98 - 107 [*DOI :* 10.1016/J.ECOLIND.2016.02.023], http://www.sciencedirect.com/science/article/pii/S1470160X16300449

[60] F. FACCHINEI, J. S. PANG.*Finite-dimensional Variational Inequalities and Complementarity Problems*, Springer Series in Operations Research, Springer Verlag NY. Inc., 2003, vol. I & II

[61] A. F. FILIPPOV.*Differential Equations with Discontinuous Right Hand Sides*, Kluwer, Dordrecht, the Netherlands, 1988

[62] M. P. FRIEDLANDER, D. ORBAN.*A primal–dual regularized interior-point method for convex quadratic programs*, in "Mathematical Programming Computation", 2012, vol. 4, n° 1, p. 71–107, http://dx.doi.org/10.1007/s12532-012-0035-2

[63] D. GOELEVEN.*Complementarity and Variational Inequalities in Electronics*, Academic Press, 2017

[64] S. GREENHALGH, V. ACARY, B. BROGLIATO.*On preserving dissipativity properties of linear complementarity dynamical systems with the θ-method*, in "Numerische Mathematik", 2013, vol. 125, p. 601–637

[65] F. GÉNOT, B. BROGLIATO.*New results on Painlevé Paradoxes*, in "European Journal of Mechanics - A. solids", 1999, vol. 18, p. 653-677

[66] T. HENZINGER.*The Theory of Hybrid Automata*, in "Proceedings of the Eleventh Annual IEEE Symposium on Logic in Computer Science (LICS)", 1996, p. 278-292

[67] J. HIRIART-URRUTY, C. LEMARÉCHAL.*Fundamentals of Convex Analysis*, Springer Verlag, 2001

[68] J. HIRIART-URRUTY, C. LEMARÉCHAL.*Convex Analysis and Minimization Algorithms*, Springer Verlag, Berlin, 1993, vol. I and II

[69] O. HUBER, V. ACARY, B. BROGLIATO.*Lyapunov stability and performance analysis of the implicit discrete sliding mode control*, in "IEEE Transactions on Automatic Control", 2016, vol. 61, n° 10, p. 3016-3030, In press

[70] O. HUBER, V. ACARY, B. BROGLIATO, F. PLESTAN.*Implicit discrete-time twisting controller without numerical chattering: analysis and experimental results*, in "Control Engineering Practice", 2016, vol. 46, 129—141

[71] O. HUBER, B. BROGLIATO, V. ACARY, A. BOUBAKIR, F. PLESTAN, B. WANG.*Experimental results on implicit and explicit time-discretization of equivalent-control-based sliding-mode control*, in "Recent Trends in Sliding Mode Control", L. FRIDMAN, J. BARBOT, F. PLESTAN (editors), IET, 2016, https://hal.inria.fr/hal-01238120

[72] G. JAMES.*Periodic travelling waves and compactons in granular chains*, in "J. Nonlinear Sci.", 2012, vol. 22, p. 813-848

[73] G. JAMES, P. KEVREKIDIS, J. CUEVAS.*Breathers in oscillator chains with Hertzian interactions*, in "Physica D", 2013, vol. 251, p. 39–59

[74] M. JEAN, V. ACARY, Y. MONERIE.*Non Smooth Contact dynamics approach of cohesive materials*, in "Philisophical Transactions : Mathematical, Physical & Engineering Sciences, The Royal Society, London A", 2001, vol. A359, n^o 1789, p. 2497–2518

[75] G. KERSCHEN, M. PEETERS, J. GOLINVAL, A. VAKAKIS.*Nonlinear normal modes, Part I: A useful framework for the structural dynamicist*, in "Mechanical Systems and Signal Processing", 2009, vol. 23, n^o 1, p. 170 - 194, Special Issue: Non-linear Structural Dynamics [*DOI :* 10.1016/J.YMSSP.2008.04.002], http://www.sciencedirect.com/science/article/pii/S0888327008001015

[76] R. LEINE.*Bifurcation in Discontinuous Mechanical Systems of Filippov-Type*, Technische Universiteit Eindhoven, 2000

[77] R. I. LEINE, A. SCHWEIZER, M. CHRISTEN, J. GLOVER, P. BARTELT, W. GERBER.*Simulation of rockfall trajectories with consideration of rock shape*, in "Multibody System Dynamics", Aug 2014, vol. 32, n^o 2, p. 241–271, https://doi.org/10.1007/s11044-013-9393-4

[78] R. LEINE, N. VAN DE WOUW.*Stability and Convergence of Mechanical Systems with Unilateral Constraints*, Lecture Notes in Applied and Computational Mechanics, Springer Verlag, 2008, vol. 36

[79] L. LIU, G. JAMES, P. KEVREKIDIS, A. VAINCHTEIN.*Nonlinear waves in a strongly resonant granular chain*, in "Nonlinearity", 2016, p. 3496-3527

[80] F. A. MIRANDA-VILLATORO, B. BROGLIATO, F. CASTAÑOS.*Multivalued Robust Tracking Control of Lagrange Systems: Continuous and Discrete–Time Algorithms*, in "IEEE Transactions on Automatic Control", 2017, vol. PP, n^o 99, 1, http://dx.doi.org/10.1109/TAC.2017.2662804

[81] J. E. MORALES, G. JAMES, A. TONNELIER.*Solitary waves in the excitable Burridge-Knopoff model*, Inria Grenoble - Rhône-Alpes, December 2016, n^o RR-8996, To appear in Wave Motion, https://hal.inria.fr/hal-01411897

[82] J. MORALES, G. JAMES, A. TONNELIER.*Traveling waves in a spring-block chain sliding down a slope*, in "Phys. Rev. E 96", 2017, vol. 96, n^o 012227

[83] Y. NESTEROV.*A method of solving a convex programming problem with convergence rate $O(1/k^2)$*, in "Soviet Mathematics Doklady", 1983, vol. 27, n^o 2, p. 372–376

[84] N. NGUYEN, B. BROGLIATO.*Multiple Impacts in Dissipative Granular Chains*, Lecture Notes in Applied and Computational Mechanics, Springer Verlag, 2014, vol. 72, XXII, 234 p. 109 illus.

[85] F. Z. NQI, M. SCHATZMAN.*Computation of Lyapunov Exponents for dynamical system with impact*, in "Applied Mathematial Sciences", 2010, vol. 4, n^o 5, p. 237–252

[86] M. PORTER, P. KEVREKIDIS, C. DARAIO. *Granular crystals: Nonlinear dynamics meets materials engineering*, in "Physics Today", 2015, vol. 68, n^{o} 44

[87] R. ROCKAFELLAR. *Convex Analysis*, Princeton University Press, 1970

[88] T. SCHINDLER, V. ACARY. *Timestepping schemes for nonsmooth dynamics based on discontinuous Galerkin methods: Definition and outlook*, in "Mathematics and Computers in Simulation", 2013, vol. 95, p. 180–199 [*DOI :* 10.1016/J.MATCOM.2012.04.012], http://www.sciencedirect.com/science/article/pii/S0378475412001231

[89] T. SCHINDLER, S. REZAEI, J. KURSAWE, V. ACARY. *Half-explicit timestepping schemes on velocity level based on time-discontinuous Galerkin methods*, in "Computer methods in Applied Mechanics in Engineering", 2015, vol. 290, n^{o} 15, p. 250–276

[90] C. STUDER. *Numerics of Unilateral Contacts and Friction. – Modeling and Numerical Time Integration in Non-Smooth Dynamics*, Lecture Notes in Applied and Computational Mechanics, Springer Verlag, 2009, vol. 47

[91] A. TONNELIER. *McKean caricature of the FitzHugh-Nagumo model: traveling pulses in a discrete diffusive medium*, in "Phys. Rev. E", 2003, vol. 67, n^{o} 036105

[92] A. VIEIRA, B. BROGLIATO, C. PRIEUR. *Optimal control of linear complementarity systems*, in "IFAC World Congress on Automatic Control", Toulouse, France, 2017

[93] B. WANG, B. BROGLIATO, V. ACARY, A. BOUBAKIR, F. PLESTAN. *Experimental comparisons between implicit and explicit implementations of discrete-time sliding mode controllers: towards input and output chattering suppression*, in "IEEE Transactions on Control Systems Technology", 2015, vol. 23, n^{o} 5, p. 2071–2075

[94] M. DI BERNARDO, C. BUDD, A. CHAMPNEYS, P. KOWALCZYK. *Piecewise-smooth dynamical systems : theory and applications*, Applied mathematical sciences, Springer, London, 2008, http://opac.inria.fr/record=b1122347

Project-Team TYREX

Types and Reasoning for the Web

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

IN PARTNERSHIP WITH:
CNRS

Institut polytechnique de Grenoble

Université de Grenoble Alpes

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team TYREX

Creation of the Team: 2012 November 01, updated into Project-Team: 2014 July 01

Keywords:

Computer Science and Digital Science:

A2.1.1. - Semantics of programming languages
A2.1.4. - Functional programming
A2.1.7. - Distributed programming
A2.1.10. - Domain-specific languages
A2.2.1. - Static analysis
A2.2.4. - Parallel architectures
A2.2.8. - Code generation
A2.4. - Formal method for verification, reliability, certification
A3.1. - Data
A3.1.1. - Modeling, representation
A3.1.2. - Data management, quering and storage
A3.1.3. - Distributed data
A3.1.6. - Query optimization
A3.1.9. - Database
A3.1.10. - Heterogeneous data
A3.1.11. - Structured data
A3.2.1. - Knowledge bases
A3.2.2. - Knowledge extraction, cleaning
A3.2.6. - Linked data
A3.3.3. - Big data analysis
A3.4. - Machine learning and statistics
A3.4.1. - Supervised learning
A5.6. - Virtual reality, augmented reality
A6.3.3. - Data processing
A7. - Theory of computation
A7.1. - Algorithms
A7.2. - Logic in Computer Science
A9.1. - Knowledge
A9.2. - Machine learning
A9.7. - AI algorithmics
A9.8. - Reasoning

Other Research Topics and Application Domains:

B6.1. - Software industry
B6.3.1. - Web
B6.5. - Information systems
B8.2. - Connected city
B9.5.1. - Computer science

B9.5.6. - Data science
B9.7.2. - Open data
B9.11. - Risk management
B9.11.2. - Financial risks

1. Team, Visitors, External Collaborators

Research Scientists
Pierre Genevès [Team leader, CNRS, Researcher, HDR]
Nabil Layaïda [Inria, Senior Researcher, HDR]

Faculty Members
Angela Bonifati [Univ de Claude Bernard, Professor, from Sep 2018, HDR]
Nils Gesbert [Institut polytechnique de Grenoble, Associate Professor]
Cécile Roisin [Univ Pierre Mendès France, Professor, HDR]

Technical Staff
Thomas Calmant [Inria]
Thibaud Michel [Inria, from Jun 2018]

PhD Students
Fateh Boulmaiz [Société privée Qualifret]
Sarah Chlyah [Inria, from Mar 2018]
Amela Fejza [Univ Grenoble Alpes, from Oct 2018]
Louis Jachiet [Ecole Normale Supérieure Paris, until Aug 2018]
Muideen Lawal [Univ Grenoble Alpes]

Post-Doctoral Fellow
Thibaud Michel [Univ Grenoble Alpes, until May 2018]

Administrative Assistant
Helen Pouchot-Rouge-Blanc [Inria]

2. Overall Objectives

2.1. Objectives

We work on the foundations of the next generation of data analytics and data-centric programming systems. These systems extend ideas from programming languages, artificial intelligence, data management systems, and theory. Data-intensive applications are increasingly more demanding in sophisticated algorithms to represent, store, query, process, analyse and interpret data. We build and study data-centric programming methods and systems at the core of artificial intelligence applications. Challenges include the robust and efficient processing of large amounts of structured, heterogeneous, and distributed data.

On the data-intensive application side, our current focus is on building efficient and scalable analytics systems. Our technical contributions particularly focus on the optimization, compilation, and synthesis of information extraction and analytics code, in particular with large amounts of data.

On the theoretical side, we develop the foundations of data-centric systems and analytics engines with a particular focus on the analysis and typing of data manipulations. We focus in particular on the foundations of programming with distributed data collections. We also study the algebraic and logical foundations of query languages, for their analysis and their evaluation.

3. Research Program

3.1. Foundations for Data Manipulation Analysis: Logics and Type Systems

We develop methods for the static analysis of queries based on logical decision procedures. Static analysis can be used to optimize runtime performance by compile-time automated modification of the code. For example, queries can be substituted by more efficient — yet equivalent — variants. The query containment problem has been a central point of research for major query languages due to its vital role in query optimization. Query containment is defined as determining if the result of one query is included in the result of another one for any dataset. We explore techniques for deciding query containment for expressive languages for querying richly structured data such as knowledge graphs. One major scientific difficulty here consists in dealing with problems close to the frontier of decidability, and therefore in finding useful trade-offs between programming expressivity, complexity, succinctness, algorithmic techniques and effective implementations. We also investigate type systems and type-checking methods for the analysis of the manipulations of structured data.

3.2. Algebraic Foundations for Query Optimization and Code Synthesis

We consider intermediate languages based on algebraic foundations for the representation, characterization, transformations and compilation of queries. We investigate extensions of the relational algebra for optimizing expressive queries, and in particular recursive queries. We explore monads and in particular monad comprehensions and monoid calculus for the generation of efficient and scalable code on big data frameworks. When transforming and optimizing algebraic terms, we rely on cost-based searches of equivalent terms. We thus develop cost models whose purpose is to estimate the time, space and network costs of query evaluation. One difficulty is to estimate these costs in architectures where data and computations are distributed, and where the modeling of data transfers is essential.

4. Application Domains

4.1. Querying Large Graphs

Increasingly large amounts of graph-structured data become available. The methods we develop apply for the efficient evaluation of graph queries over large — and potentially distributed — graphs. In particular, we consider the SPARQL query language, which is the standard language for querying graphs structured in the Resource Description Format (RDF). We also consider other increasingly popular graph query languages such as Cypher queries for extracting information from property graphs.

We compile graph queries into lower-level distributed primitives found in big data frameworks such as Apache Spark, Flink, etc. Applications of graph querying are ubiquitous and include: large knowledge bases, social networks, road networks, trust networks and fraud detection for cryptocurrencies, publications graphs, web graphs, recommenders, etc.

4.2. Predictive Analytics for Healthcare

One major expectation of data science in healthcare is the ability to leverage on digitized health information and computer systems to better apprehend and improve care. The availability of large amounts of clinical data and in particular electronic health records opens the way to the development of quantitative models for patients that can be used to predict health status, as well as to help prevent disease and adverse effects.

In collaboration with the CHU Grenoble, we explore solutions to the problem of predicting important clinical outcomes such as patient mortality, based on clinical data. This raises many challenges including dealing with the very high number of potential predictor variables and very resource-consuming data preparation stages.

4.3. Mobile and Augmented Reality Applications

The term Augmented Environments refers collectively to ubiquitous computing, context-aware computing, and intelligent environments. The goal of our research on these environments is to introduce personal Augmented Reality (AR) devices, taking advantage of their embedded sensors. These environments offer the possibility of using ubiquitous computation, communication, and sensing to enable the presentation of context-sensitive information and services to the user. AR applications often rely on 3D content and employ specialized hardware and computer vision techniques for both tracking and scene reconstruction and exploration. Our approach tries to seek a balance between these traditional AR contexts and what has come to be known as mobile AR browsing, based for instance on attitude estimation.

5. New Software and Platforms

5.1. SPARQLGX

KEYWORDS: RDF - SPARQL - Distributed computing
SCIENTIFIC DESCRIPTION: SPARQL is the W3C standard query language for querying data expressed in RDF (Resource Description Framework). The increasing amounts of RDF data available raise a major need and research interest in building efficient and scalable distributed SPARQL query evaluators.

In this context, we propose and share SPARQLGX: our implementation of a distributed RDF datastore based on Apache Spark. SPARQLGX is designed to leverage existing Hadoop infrastructures for evaluating SPARQL queries. SPARQLGX relies on a translation of SPARQL queries into executable Spark code that adopts evaluation strategies according to (1) the storage method used and (2) statistics on data. Using a simple design, SPARQLGX already represents an interesting alternative in several scenarios.
FUNCTIONAL DESCRIPTION: This software system is an implementation of a distributed evaluator of SPARQL queries. It makes it possible to evaluate SPARQL queries on billions of triples distributed across multiple nodes in a cluster, while providing attractive performance figures.

RELEASE FUNCTIONAL DESCRIPTION: - Faster load routine which widely improves this phase perfomances by reading once the initial triple file and by partitioning data in the same time into the correct predicate files. - Improving the generated Scala-code of the translation process with mapValues. This technic allows not to break the partitioning of KeyValueRDD while applying transformations to the values instead of the traditional map that was done prior. - Merging and cleaning several scripts in bin/ such as for example sgx-eval.sh and sde-eval.sh - Improving the compilation process of compile.sh - Cleaner test scripts in tests/ - Offering the possibility of an easier deployment using Docker.

- Participants: Damien Graux, Thomas Calmant, Louis Jachiet, Nabil Layaïda and Pierre Genevès
- Contact: Pierre Genevès
- Publications: Optimizing SPARQL query evaluation with a worst-case cardinality estimation based on statistics on the data - The SPARQLGX System for Distributed Evaluation of SPARQL Queries
- URL: https://github.com/tyrex-team/sparqlgx

5.2. musparql

KEYWORDS: SPARQL - RDF - Property paths
FUNCTIONAL DESCRIPTION: reads a SPARQL request and translates it into an internal algebra. Rewrites the resulting term into many equivalent versions, then choses one of them and executes it on a graph.

- Participant: Louis Jachiet
- Contact: Nabil Layaïda
- Publication: Extending the SPARQL Algebra for the optimization of Property Paths
- URL: https://gitlab.inria.fr/tyrex/musparql

5.3. MRB

Mixed Reality Browser

KEYWORDS: Augmented reality - Geolocation - Indoor geolocalisation - Smartphone

FUNCTIONAL DESCRIPTION: MRB displays PoI (Point of Interest) content remotely through panoramics with spatialized audio, or on-site by walking to the corresponding place, it can be used for indoor-outdoor navigation, with assistive audio technology for the visually impaired. It is the only browser of geolocalized data to use XML as a native format for PoIs, panoramics, 3D audio and to rely on HTML5 both for the iconic and full information content of PoIs. Positioning in MRB is based on a PDR library, written in C++ and Java and developed by the team, which provides the user's location in real time based on the interpretation of sensors. Three main modules have been designed to build this positioning system: (i) a pedometer that estimates the distance the user has walked and his speed, (ii) a motion manager that enables data set recording and simulation but also the creation of virtual sensors or filters (e.g gyroscope drift compensation, linear acceleration, altimeter), and (iii) a map-matching algorithm that provides a new location based on a given OpenStreetMap file description and the current user's trajectory.

- Participant: Thibaud Michel
- Contact: Nabil Layaïda
- Publications: On Mobile Augmented Reality Applications based on Geolocation - Attitude Estimation for Indoor Navigation and Augmented Reality with Smartphones
- URL: http://tyrex.inria.fr/projects/mrb.html

5.4. Benchmarks Attitude Smartphones

KEYWORDS: Experimentation - Motion analysis - Sensors - Performance analysis - Smartphone

SCIENTIFIC DESCRIPTION: We investigate the precision of attitude estimation algorithms in the particular context of pedestrian navigation with commodity smartphones and their inertial/magnetic sensors. We report on an extensive comparison and experimental analysis of existing algorithms. We focus on typical motions of smartphones when carried by pedestrians. We use a precise ground truth obtained from a motion capture system. We test state-of-the-art attitude estimation techniques with several smartphones, in the presence of magnetic perturbations typically found in buildings. We discuss the obtained results, analyze advantages and limits of current technologies for attitude estimation in this context. Furthermore, we propose a new technique for limiting the impact of magnetic perturbations with any attitude estimation algorithm used in this context. We show how our technique compares and improves over previous works.

- Participants: Hassen Fourati, Nabil Layaïda, Pierre Genevès and Thibaud Michel
- Partner: GIPSA-Lab
- Contact: Pierre Genevès
- URL: http://tyrex.inria.fr/mobile/benchmarks-attitude/

5.5. MedAnalytics

KEYWORDS: Big data - Predictive analytics - Distributed systems

FUNCTIONAL DESCRIPTION: We implemented a method for the automatic detection of at-risk profiles based on a fine-grained analysis of prescription data at the time of admission. The system relies on an optimized distributed architecture adapted for processing very large volumes of medical records and clinical data. We conducted practical experiments with real data of millions of patients and hundreds of hospitals. We demonstrated how the various perspectives of big data improve the detection of at-risk patients, making it possible to construct predictive models that benefit from volume and variety. This prototype implementation is described in the 2017 preprint available at: https://hal.inria.fr/hal-01517087/document.

- Participants: Pierre Genevès and Thomas Calmant
- Partner: CHU Grenoble
- Contact: Pierre Genevès
- Publication: Scalable Machine Learning for Predicting At-Risk Profiles Upon Hospital Admission

5.6. MuIR

Mu Intermediate Representation

KEYWORDS: Optimizing compiler - Querying

FUNCTIONAL DESCRIPTION: This is a prototype of an intermediate language representation, i.e. an implementation of algebraic terms, rewrite rules, query plans, cost model, query optimizer, and query evaluators (including a distributed evaluator of algebraic terms using Apache Spark).

- Contact: Pierre Genevès

6. New Results

6.1. On the Optimization of Recursive Relational Queries

Graph databases have received a lot of attention recently as they are particularly useful in many applications such as social networks or for the semantic web. Various languages have emerged to query such graph databases. At the heart of many of those query languages, there is a construction to navigate through the graph which allows some form of recursion.The relational model has benefited from a huge body of research in the last half century and that is why many graph databases either rely on, or have adopted the techniques of, relational-based query engines. Since its introduction, the relational model has seen various attempts to extend it with recursion and it is now possible to use recursion in several SQL- or Datalog-based database systems. The optimization of recursive queries remains, however, a challenge. In this work, we introduce μ-RA, a variation of the Relational Algebra that allows for the expression of relational queries with recursion. μ-RA can express unions of conjunctive regular path queries as well as certain non-regular properties. We present its syntax, semantics and the rewriting rules we specifically devised to tackle the optimization of recursive queries. A prototype evaluator implementing these rewriting rules is shown to be more efficient than previous approaches.

These results were presented at the BDA 2018 conference [14].

6.2. A Multi-Criteria Experimental Ranking of Distributed SPARQL Evaluators

SPARQL is the standard language for querying RDF data. There exists a variety of SPARQL query evaluation systems implementing different architectures for the distribution of data and computations. Differences in architectures coupled with specific optimizations, for e.g. preprocessing and indexing, make these systems incomparable from a purely theoretical perspective. This results in many implementations solving the SPARQL query evaluation problem while exhibiting very different behaviours, not all of them being adapted to any context. We provide a new perspective on distributed SPARQL evaluators, based on multi-criteria experimental rankings. Our suggested set of 5 features (namely velocity, immediacy, dynamicity, parsimony, and resiliency) provides a more comprehensive description of the behaviours of distributed evaluators when compared to traditional runtime performance metrics. We show how these features help in more accurately evaluating to which extent a given system is appropriate for a given use case. For this purpose, we systematically benchmarked a panel of 10 state-of-the-art implementations. We ranked them using a reading grid that helps in pinpointing the advantages and limitations of current technologies for the distributed evaluation of SPARQL queries.

These results were presented at the IEEE Big Data 2018 conference [13].

6.3. SPARQL Query Containment under Schema

Query containment is defined as the problem of determining if the result of a query is included in the result of another query for any dataset. It has major applications in query optimization and knowledge base verification. The main objective of this work is to provide sound and complete procedures to determine containment of SPARQL queries under expressive description logic schema axioms. Beyond that, these procedures are experimentally evaluated. To date, testing query containment has been performed using different techniques: containment mapping, canonical databases, automata theory techniques and through a reduction to the validity problem in logic. In this work, we use the latter technique to test containment of SPARQL queries using an expressive modal logic called μ-calculus. For that purpose, we define an RDF graph encoding as a transition system which preserves its characteristics. In addition, queries and schema axioms are encoded as μ-calculus formulae. Thereby, query containment can be reduced to testing validity in the logic. We identify various fragments of SPARQL and description logic schema languages for which containment is decidable. Additionally, we provide theoretically and experimentally proven procedures to check containment of these decidable fragments. Finally, we propose a benchmark for containment solvers which is used to test and compare the current state-of-the-art containment solvers.

These results were published in the Journal on Data Semantics [4].

6.4. Selectivity Estimation for SPARQL Triple Patterns with Shape Expressions

ShEx (Shape Expressions) is a language for expressing constraints on RDF graphs. In this work we optimize the evaluation of conjunctive SPARQL queries, on RDF graphs, by taking advantage of ShEx constraints. Our optimization is based on computing and assigning ranks to query triple patterns, dictating their order of execution. We first define a set of well-formed ShEx schemas that possess interesting characteristics for SPARQL query optimization. We then define our optimization method by exploiting information extracted from a ShEx schema. We finally report on evaluation results performed showing the advantages of applying our optimization on the top of an existing state-of-the-art query evaluation system.

These results were presented at the 2018 International Conference on Web Engineering [9].

6.5. Evaluation of Query Transformations without Data

Query transformations are ubiquitous in semantic web query processing. For any situation in which transformations are not proved correct by construction, the quality of these transformations has to be evaluated. Usual evaluation measures are either overly syntactic and not very informative — the result being: correct or incorrect — or dependent from the evaluation sources. Moreover, both approaches do not necessarily yield the same result. We suggest that grounding the evaluation on query containment allows for a data-independent evaluation that is more informative than the usual syntactic evaluation. In addition, such evaluation modalities may take into account ontologies, alignments or different query languages as soon as they are relevant to query evaluation.

These results were presented at a workshop of the 2018 International Conference on World Wide Web [10].

6.6. Graph Queries: From Theory to Practice

In this work, we review various graph query language fragments that are both theoretically tractable and practically relevant. We focus on the most expressive one that retains these properties and use it as a stepping stone to examine the underpinnings of graph query evaluation along graph view maintenance. Further broadening the scope of the discussion, we then consider alternative processing techniques for graph queries, based on graph summarization and path query learning. We conclude by pinpointing the open research directions in this emerging area. These results were published in Sigmod Record Journal [3].

6.7. Query-based Linked Data Anonymization

In this work, we introduce and develop a declarative framework for privacy-preserving Linked Data publishing in which privacy and utility policies are specified as SPARQL queries. Our approach is data independent and leads to inspect only the privacy and utility policies in order to determine the sequence of anonymization operations applicable to any graph instance for satisfying the policies. We prove the soundness of our algorithms and gauge their performance through experiments.

These results were presented in the International Semantic Web Conference (ISWC 2018) [11].

6.8. Querying Graphs

Graph data modeling and querying arises in many practical application domains such as social and biological networks where the primary focus is on concepts and their relationships and the rich patterns in these complex webs of interconnectivity. In this book, we present a concise unified view on the basic challenges which arise over the complete life cycle of formulating and processing queries on graph databases. To that purpose, we present all major concepts relevant to this life cycle, formulated in terms of a common and unifying ground: the property graph data model — the predominant data model adopted by modern graph database systems.

In this book [17], we aim especially to give a coherent and in-depth perspective on current graph querying and an outlook for future developments. Our presentation is self-contained, covering the relevant topics from: graph data models, graph query languages and graph query specification, graph constraints, and graph query processing. We conclude by indicating major open research challenges towards the next generation of graph data management systems.

6.9. Backward Type Inference for XML Queries

Although XQuery is a statically typed, functional query language for XML data, some of its features such as upward and horizontal XPath axes are typed imprecisely. The main reason is that while the XQuery data model allows to navigate upwards and between siblings from a given XML node, the type model, e.g., regular tree types, can describe only the subtree structure of the given node. Recently, Giuseppe Castagna and our team independently proposed in 2015 a precise forward type inference system for XQuery using an extended type language that can describe not only a given XML node but also its context. In this work, as a complementary method to such forward type inference systems, we propose an enhanced backward type inference system for XQuery, based on an extended type language. Results include an exact type system for XPath axes and a sound type system for XQuery expressions [19].

6.10. Scalable and Interpretable Predictive Models for Electronic Health Records

Early identification of patients at risk of developing complications during their hospital stay is currently one of the most challenging issues in healthcare. Complications include hospital-acquired infections, admissions to intensive care units, and in-hospital mortality. Being able to accurately predict the patients' outcomes is a crucial prerequisite for tailoring the care that certain patients receive, if it is believed that they will do poorly without additional intervention. We consider the problem of complication risk prediction, such as patient mortality, from the electronic health records of the patients. We study the question of making predictions on the first day at the hospital, and of making updated mortality predictions day after day during the patient's stay. We develop distributed models that are scalable and interpretable. Key insights include analysing diagnoses known at admission and drugs served, which evolve during the hospital stay. We leverage a distributed architecture to learn interpretable models from training datasets of gigantic size. We test our analyses with more than one million of patients from hundreds of hospitals, and report on the lessons learned from these experiments.

These results were presented at the 2018 International Conference on Data Science and Applications [12].

6.11. Scalable Machine Learning for Predicting At-Risk Profiles Upon Hospital Admission

We show how the analysis of very large amounts of drug prescription data make it possible to detect, on the day of hospital admission, patients at risk of developing complications during their hospital stay. We explore, for the first time, to which extent volume and variety of big prescription data help in constructing predictive models for the automatic detection of at-risk profiles. Our methodology is designed to validate our claims that: (1) drug prescription data on the day of admission contain rich information about the patient's situation and perspectives of evolution, and (2) the various perspectives of big medical data (such as veracity, volume, variety) help in extracting this information.We build binary classification models to identify at-risk patient profiles. We use a distributed architecture to ensure scalability of model construction with large volumes of medical records and clinical data. We report on practical experiments with real data of millions of patients and hundreds of hospitals. We demonstrate how the fine-grained analysis of such big data can improve the detection of at-risk patients, making it possible to construct more accurate predictive models that significantly benefit from volume and variety, while satisfying important criteria to be deployed in hospitals.

These results were published in the Big Data Research journal [6].

6.12. ProvSQL: Provenance and Probability Management in PostgreSQL

This demonstration showcases ProvSQL, an open-source module for the PostgreSQL database management system that adds support for computation of provenance and probabilities of query results. A large range of provenance formalisms are supported, including all those captured by provenance semirings, provenance semirings with monus, as well as where-provenance. Probabilistic query evaluation is made possible through the use of knowledge compilation tools, in addition to standard approaches such as enumeration of possible worlds and Monte-Carlo sampling. ProvSQL supports a large subset of non-aggregate SQL queries.

These results were published in the PVLDB journal [8].

6.13. A Method to Quantitatively Evaluate Geo Augmented Reality Applications

We propose a method for quantitatively assessing the quality of Geo AR browsers. Our method aims at measuring the impact of attitude and position estimations on the rendering precision of virtual features. We report on lessons learned by applying our method on various AR use cases with real data. Our measurement technique allows shedding light on the limits of what can be achieved in Geo AR with current technologies. This also helps in identifying interesting perspectives for the further development of high-quality Geo AR applications.

These results were presented at the ISMAR 2018 conference [15].

6.14. Attitude Estimation for Indoor Navigation and Augmented Reality with Smartphones

We investigate the precision of attitude estimation algorithms in the particular context of pedestrian navigation with commodity smartphones and their inertial/magnetic sensors. We report on an extensive comparison and experimental analysis of existing algorithms. We focus on typical motions of smartphones when carried by pedestrians. We use a precise ground truth obtained from a motion capture system. We test state-of-the-art and built-in attitude estimation techniques with several smartphones, in the presence of magnetic perturbations typically found in buildings. We discuss the obtained results, analyze advantages and limits of current technologies for attitude estimation in this context. Furthermore, we propose a new technique for limiting the impact of magnetic perturbations with any attitude estimation algorithm used in this context. We show how our technique compares and improves over previous works. A particular attention was paid to the study of attitude estimation in the context of augmented reality motions when using smartphones.

These results were published in the Pervasive and Mobile Computing journal [7].

6.15. A Hybrid Approach for Spatio-Temporal Validation of Declarative Multimedia

Declarative multimedia documents represent the description of multimedia applications in terms of media items and relationships among them. Relationships specify how media items are dynamically arranged in time and space during runtime. Although a declarative approach usually facilitates the authoring task, authors can still make mistakes due to incorrect use of language constructs or inconsistent or missing relationships in a document. In order to properly support multimedia application authoring, it is important to provide tools with validation capabilities. Document validation can indicate possible inconsistencies in a given document to an author so that it can be revised before deployment. Although very useful, multimedia validation tools are not often provided by authoring tools. This work proposes a multimedia validation approach that relies on a formal model called Simple Hyper-media Model (SHM). SHM is used for representing a document for the purpose of validation. An SHM document is validated using a hybrid approach based on two complementary techniques. The first one captures the document's spatio-temporal layout in terms of its state throughout its execution by means of a rewrite theory, and validation is performed through model checking. The second one captures the document's layout in terms of intervals and event occurrences by means of Satisfiability Modulo Theories (SMT) formulas, and validation is performed through SMT solving. Due to different characteristics of both approaches, each validation technique complements the other in terms of expressiveness of SHM and tests to be checked. We briefly present validation tools that use our approach. They were evaluated with real NCL documents and by usability tests.

These results were published in the ACM Transactions on Multimedia Computing, Communications and Applications journal [5].

7. Partnerships and Cooperations

7.1. Regional Initiatives

Data-CILE

Title: Query Compilation

Call: Appel à projet Grenoble Innovation Recherche (AGIR-Pôle)

Duration: 2016-2018

Coordinator: Nabil Layaïda

Abstract: The goal of this project is to contribute to foundational and algorithmic challenges introduced by increasingly popular data-centric paradigms for programming on distributed architectures such as spark and the massive production of big linked open data. The focus of the project is on building robust and more efficient workflows of transformations of semantic and graph web data.

BioQurate

Title: Querying and Curating Hierarchies of Biological Graphs

Funding: Fédération Informatique de Lyon (FIL)

Duration: 2018-2020

Coordinator: Angela Bonifati

Others partners: LIP/LIRIS. The project involves a bio-computing team and a database team on a common research problem

Abstract: This project aims at leveraging graph rewriting techniques of ReGraph and graph data management techniques in order to provide a persistent, robust and scalable substrate for the construction and manipulation of hierarchies of biological graphs. Moreover, we wish to investigate whether the involved graphs need further expressive graph constraints for enforcing consistency and performing data cleansing.

7.2. National Initiatives

7.2.1. ANR

CLEAR

Title: Compilation of intermediate Languages into Efficient big dAta Runtimes

Call: Appel à projets générique 2016 défi 'Société de l'information et de la communication' – JCJC

Duration: January 2017 – September 2021

Coordinator: Pierre Genevès

See also: http://tyrex.inria.fr/clear

Abstract: This project addresses one fundamental challenge of our time: the construction of effective programming models and compilation techniques for the correct and efficient exploitation of big and linked data. We study high-level specifications of pipelines of data transformations and extraction for producing valuable knowledge from rich and heterogeneous data. We investigate how to synthesize code which is correct and optimized for execution on distributed infrastructures.

DataCert

Title: Coq deep specification of security aware data integration

Call: Appel à projets Sciences et technologies pour la confiance et la sécurité numérique

Duration: January 2016 – January 2020

Participant: Angela Bonifati

Others partners: Université Paris Sud/Laboratoire de Recherche en Informatique, Université de Lille/Centre de Recherche en Informatique, Signal et Automatique de Lille, Université de Lyon/Laboratoire d'InfoRmatique en Image et Systèmes d'information.

See also: http://datacert.lri.fr/

Abstract: This project's aim is to develop a comprehensive framework handling the fundamental problems underlying security-aware data integration and sharing, resulting in a paradigm shift in the design and implementation of security-aware data integration systems. To fill the gap between both worlds, we strongly rely on deep specifications and proven-correct software, develop formal models yielding highly reliable technology while controlling the disclosure of private or confidential information.

QualiHealth

Title: Enhancing the Quality of Health Data

Call: Appel à projets Projets de Recherche Collaborative – Entreprise (PRCE)

Duration: 2018-2022

Coordinator: Angela Bonifati

Others partners: LIMOS, Université Clermont Auvergne. LIS, Université d'Aix-Marseille. HEGP, INSERM, Paris. Inst. Cochin, INSERM, Paris. Gnubila, Argonay. The University of British Columbia, Vancouver (Canada)

Abstract: This research project is geared towards a system capable of capturing and formalizing the knowledge of data quality from domain experts, enriching the available data with this knowledge and thus exploiting this knowledge in the subsequent quality-aware medical research studies. We expect a quality-certified collection of medical and biological datasets, on which quality-certified analytical queries can be formulated. We envision the conception and implementation of a quality-aware query engine with query enrichment and answering capabilities.

To reach this ambitious objectives, the following concrete scientific goals must be fulfilled : (1) An innovative research approach, that starts from concrete datasets and expert practices and knowledge to reach formal models and theoretical solutions, will be employed to elicit innovative quality dimensions and to identify, formalize, verify and finally construct quality indicators able to capture the variety and complexity of medical data; those indicators have to be composed, normalized and aggregated when queries involve data with different granularities (e.g., accuracy indications on pieces of information at the patient level have to be composed when one queries cohort) and of different quality dimensions (e.g., mixing incomplete and inaccurate data); and (2) In turn, those complex aggregated indicators have to be used to provide new quality-driven query answering, refinement, enrichment and data analytics techniques. A key novelty of this project is the handling of data which are not rectified on the original database but sanitized in a query-driven fashion: queries will be modified, rewritten and extended to integrate quality parameters in a flexible and automatic way.

7.2.2. PERSYVAL-lab LabEx

Title: Mobile Augmented Reality Applications for Smart Cities

Call: Persyval Labex (“Laboratoire d’excellence”).

Duration: 2014 – 2018

Coordinators: Pierre Genevès and Nabil Layaïda

Others partners: NeCS team at GIPSA-Lab laboratory.

Abstract: The goal of this project is to increase the relevance and reliability of augmented reality (AR) applications, through three main objectives:

1. Finding and developing appropriate representations for describing the physical world (3D maps, indoor buildings, ways...), integrated advanced media types (3D, 3D audio, precisely geo-tagged pictures with lat., long. and orientation, video...)
2. Integrating the different abstraction levels of these data streams (ranging from sensors data to high level rich content such as 3D maps) and bridging the gap with Open Linked Data (the semantic World). This includes opening the way to query the environment (filtering), and adapt AR browsers to users’ capabilities (e.g. blind people). The objective here is to provide an open and scalable platform for mobile-based AR systems (just like the web represents).
3. Increasing the reliability and accuracy of localization technologies. Robust and high-accuracy localization technologies play a key role in AR applications. Combined with geographical data, they can also be used to identify user-activity patterns, such as walking, running or being in an elevator. The interpretation of sensor values, coupled with different walking models, allows one to ensure the continuity of the localization, both indoor and outdoor. However, dead reckoning based on Inertial Navigation Systems (INS) or Step-and-Heading Systems (SHS) is subject to cumulative errors due to many factors (sensor drift (accelerometers, gyroscopes, etc.), missed steps, bad estimation of the length of each stride, etc.). One objective is to reduce such errors by merging and mixing these approaches with various external signals such as GPS and Wi-Fi or relying on the analyses of user trajectories with the help of a structured map of the environment. Some filtering methods (Kalman Filter, observer, etc.) will be useful to achieve this task.

7.3. International Research Visitors

7.3.1. Visits of International Scientists

We had short visits from Wim Martens (University of Bayreuth, Germany) and Efthymia Tsamoura (University of Oxford, UK).

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events Organisation

8.1.1.1. Member of the Organizing Committees

- P. Genevès is member of the Organizing Committee of BDA 2019.
- A. Bonifati is a permanent member of ICDT Council (The International Conference on Database Theory).

8.1.2. Scientific Events Selection

8.1.2.1. Chair of Conference Program Committees

- A. Bonifati is Co-chair of the SIGMOD 2019 Workshops.

8.1.2.2. Member of the Conference Program Committees

- P. Genevès has been program committee member for the 27th International Joint Conference on Artificial Intelligence (IJCAI'18) and for the 23rd European Conference on Artificial Intelligence (ECAI'18) and for the 18th ACM Symposium on Document Engineering (DocEng'18).
- A. Bonifati has been program committee member of VLDB 2019, AAAI 2019, ICDE 2019, EDBT 2019, SIGMOD 2019 (senior PC member), PODS 2019, DEBS 2019, ICDT 2020.

8.1.2.3. Reviewer

- P. Genevès has been reviewer for the ICALP 2018, CHI 2018, and BDA 2018 conferences.
- C. Roisin has been reviewer for the 18th ACM Symposium on Document Engineering (DocEng'18).

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

- A. Bonifati is Associate Editor of ACM Trans. on Database Systems.
- A. Bonifati is Associate Editor of the VLDB Journal.

8.1.3.2. Reviewer - Reviewing Activities

- P. Genevès has been reviewer for the Sensors journal.

8.1.4. Invited Talks

- P. Genevès gave an invited talk for the Data Science in the Alps Workshop (March 20th, 2018): "On the Prediction of At-Risk Patient Profiles with Big Prescription Data"
- P. Genevès gave an invited seminar at the University of Fribourg (March 27th, 2018): "Queries for Trees and Graphs: Static Analysis and Code Synthesis"

8.1.5. Scientific Expertise

- C. Roisin has been designed as expert for the reviewing of a research proposal of the American University of Beirut.
- P. Genevès has been a scientific expert at ANRT for the CIFRE funding process.

8.1.6. Research Administration

- P. Genevès is responsible for the Computer Science Specialty at the Doctoral School MSTII (ED 217)
- P. Genevès has been Member of a Hiring Committee at IRIF.
- C. Roisin is a member of the CNU (Conseil National des Universités).
- C. Roisin is a member of the Inria Grenoble Inria-Hub committee.

- N. Layaïda is a member of the experts pool (selection committee) of the minalogic competitive cluster.
- A. Bonifati and N. Layaïda are members of the Scientific Board of Digital League, the digital cluster of Auvergne-Rhone-Alpes.
- A. Bonifati is coordinator of the theme « Masses de Données » at Liris and at « Fédération d'Informatique de Lyon » (FIL).

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

- Licence : C. Roisin, Programmation C, 12h eq TD, L2, IUT2, Univ. Grenoble-Alpes
- Licence : C. Roisin, Architecture des réseaux, 112h eq TD, L1, IUT2, Univ. Grenoble-Alpes
- Licence : C. Roisin, Services réseaux, 22h eq TD, L2, IUT2, Univ. Grenoble-Alpes
- Licence : C. Roisin, Introduction système Linux, 21h eq TD, L1, IUT2, Univ. Grenoble-Alpes
- Licence : C. Roisin, Système et réseaux, 14h eq TD, L3, IUT2, Univ. Grenoble-Alpes
- Licence : C. Roisin, Tutorat pédagogique de 4 apprentis, 20h eq TD, L3, IUT2, Univ. Grenoble-Alpes
- Licence : C. Roisin, Suivi pédagogique de 20 étudiants (responsable de la Licence Professionnelle MI-ASSR), 13h eq TD, L1, IUT2, Univ. Grenoble-Alpes
- Licence : N. Gesbert, 'Logique pour l'informatique', 45 h eq TD, L3, Grenoble INP
- Licence : N. Gesbert, 'Bases de la programmation impérative', 30 h eq TD, L3, Grenoble INP
- Master : N. Gesbert, academic tutorship of an apprentice, 10 h eq TD, M1, Grenoble INP
- Master : N. Gesbert, 'Fondements logiques pour l'informatique', 16 h 30 eq TD, M1, Grenoble INP
- Master : N. Gesbert, 'Construction d'applications Web', 21 h eq TD, M1, Grenoble INP
- Master : N. Gesbert, 'Analyse, conception et validation de logiciels', 30 h eq TD, M1, Grenoble INP
- N. Gesbert is responsible of the L3-level course 'logique pour l'informatique' (25 apprentices) and of the M1-level course 'construction d'applications Web' (72 students).
- P. Genevès is responsible and teacher in the the M2-level course 'Semantic Web: from XML to OWL' of the MOSIG program at UGA (36h)
- P. Genevès is responsible and teacher in the the M2-level course 'Accès à l'information: du web des données au web sémantique' of the ENSIMAG ISI 3A program at Grenoble-INP (30h)

8.2.2. Supervision

- PhD: Louis Jachiet, On the foundations for the compilation of web data queries: optimization and distributed evaluation of SPARQL, University Grenoble Alpes. PhD Thesis defended on September 15th, 2018. Co-supervised by Nabil Layaïda and Pierre Genevès.
- PhD in progress: Muideen Lawal, Cost models for optimizing compilers based on mu-terms, PhD started in October 2017, supervised by Pierre Genevès.
- PhD in progress: Raouf Kerkouche, Privacy-preserving predictive analytics with big prescription data, PhD started in October 2017, co-supervised by Pierre Genevès and Claude Castelluccia.
- PhD in progress: Fateh Boulmaiz, Distributed representations of large-scale graphs, PhD started in November 2017, co-supervised by Pierre Genevès and Nabil Layaïda.
- PhD in progress: Sarah Chlyah, Algebraic foundations for the synthesis of optimized distributed code, PhD started in March 2018, supervised by Pierre Genevès.

- PhD in progress: Amela Fejza, On the extended algebraic representations for analytical workloads, PhD started in October 2018, supervised by Pierre Genevès.

8.3. Popularization

- N. Layaïda contributed to the special édition of the newspaper : Le Dauphiné libéré des enfants – C'est quoi le numérique ? – n°14 November-December 2018.
- P. Genevès presented the Tyrex team activities in 180 seconds in a dissemination event organized at Inria Montbonnot in June 2018.

8.3.1. Creation of media or tools for science outreach

- T. Michel built a mobile application for the Inria public showroom, Espace Login, called Login'AR. The application was designed to showcase the combination of indoor positioning with Augmented Reality capabilities. The application is now part of the permanent Login public exhibition. The application was developed with the help of four Grenoble INP interns (A. Convert, A. Fortecof, G. Montano and S.J. Tourangeau).

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] L. Jachiet.*On the foundations for the compilation of web data queries: optimization and distributed evaluation of SPARQL*, Communauté Université Grenoble Alpes, September 2018, https://hal.inria.fr/tel-01891444

[2] L. Jachiet.*On the foundations for the compilationof web data queries : optimization anddistributed evaluation of SPARQL*, Université Grenoble Alpes, September 2018, https://tel.archives-ouvertes.fr/tel-01960209

Articles in International Peer-Reviewed Journal

[3] A. Bonifati, S. Dumbrava.*Graph Queries: From Theory to Practice*, in "ACM SIGMOD Record", December 2018, vol. 47, n^o 4, https://hal.inria.fr/hal-01977048

[4] M. Chekol, J. Euzenat, P. Genevès, N. Layaïda.*SPARQL Query Containment under Schema*, in "Journal on Data Semantics", April 2018, vol. 7, n^o 3, p. 133-154 [*DOI :* 10.1007/s13740-018-0087-1], https://hal.inria.fr/hal-01767887

[5] J. A. F. Dos Santos, D. C. Muchaluat-Saade, C. Roisin, N. Layaïda.*A Hybrid Approach for Spatio-Temporal Validation of Declarative Multimedia Documents*, in "ACM Transactions on Multimedia Computing, Communications and Applications", November 2018, vol. 14, n^o 4, p. 1-24 [*DOI :* 10.1145/3267127], https://hal.inria.fr/hal-01946641

[6] P. Genevès, T. Calmant, N. Layaïda, M. Lepelley, S. Artemova, J.-L. Bosson.*Scalable Machine Learning for Predicting At-Risk Profiles Upon Hospital Admission*, in "Big Data Research", March 2018, vol. 12, p. 23-34 [*DOI :* 10.1016/j.bdr.2018.02.004], https://hal.inria.fr/hal-01517087

[7] T. Michel, P. Genevès, H. Fourati, N. Layaïda.*Attitude Estimation for Indoor Navigation and Augmented Reality with Smartphones*, in "Pervasive and Mobile Computing", March 2018 [*DOI :* 10.1016/j.pmcj.2018.03.004], https://hal.inria.fr/hal-01650142

[8] P. SENELLART, L. JACHIET, S. MANIU, Y. RAMUSAT.*ProvSQL: Provenance and Probability Management in PostgreSQL*, in "Proceedings of the VLDB Endowment (PVLDB)", August 2018, vol. 11, n^o 12, p. 2034-2037 [*DOI :* 10.14778/3229863.3236253], https://hal.inria.fr/hal-01851538

International Conferences with Proceedings

[9] A. ABBAS, P. GENEVÈS, C. ROISIN, N. LAYAÏDA.*Selectivity Estimation for SPARQL Triple Patterns with Shape Expressions*, in "ICWE'18 - 18th International Conference on Web Engineering", Cáceres, Spain, Springer, June 2018, p. 195-209 [*DOI :* 10.1007/978-3-319-91662-0_15], https://hal.inria.fr/hal-01673013

[10] J. DAVID, J. EUZENAT, P. GENEVÈS, N. LAYAÏDA.*Evaluation of Query Transformations without Data*, in "WWW 2018 - Companion of The Web Conference", Lyon, France, ACM Press, April 2018, p. 1599-1602 [*DOI :* 10.1145/3184558.3191617], https://hal.inria.fr/hal-01891182

[11] R. DELANAUX, A. BONIFATI, M.-C. ROUSSET, R. THION.*Query-based Linked Data Anonymization*, in "The 17th International Semantic Web Conference (ISWC 2018)", Monterey, United States, October 2018, p. 530-546, https://hal.archives-ouvertes.fr/hal-01896276

[12] A. FEJZA, P. GENEVÈS, N. LAYAÏDA, J.-L. BOSSON.*Scalable and Interpretable Predictive Models for Electronic Health Records*, in "DSAA 2018 - 5th IEEE International Conference on Data Science and Advanced Analytics", Turin, Italy, IEEE, October 2018, p. 1-10, https://hal.inria.fr/hal-01877742

[13] D. GRAUX, L. JACHIET, P. GENEVÈS, N. LAYAÏDA.*A Multi-Criteria Experimental Ranking of Distributed SPARQL Evaluators*, in "Big Data 2018 - IEEE International Conference on Big Data", Seattle, United States, IEEE, December 2018, p. 1-10, https://hal.inria.fr/hal-01381781

[14] L. JACHIET, N. GESBERT, P. GENEVÈS, N. LAYAÏDA.*On the Optimization of Recursive Relational Queries*, in "BDA 2018, 34ème Conférence sur la Gestion de Données - Principes, Technologies et Applications", Bucarest, Romania, October 2018, https://hal.inria.fr/hal-01673025

[15] T. MICHEL, P. GENEVÈS, N. LAYAÏDA.*A Method to Quantitatively Evaluate Geo Augmented Reality Applications*, in "ISMAR 2018 - International Symposium on Mixed and Augmented Reality", Munich, Germany, October 2018, p. 1-6, https://hal.inria.fr/hal-01890838

[16] D. SCHWAB, A. FEJZA, L. VIAL, Y. ROBERT.*The GazePlay Project: Open and Free Eye-trackers Games and a Community for People with Multiple Disabilities*, in "ICCHP 2018 - 16th International Conference on Computers Helping People with Special Needs", Linz, Austria, LNCS, Springer, July 2018, vol. 10896, p. 254-261 [*DOI :* 10.1007/978-3-319-94277-3_41], https://hal.archives-ouvertes.fr/hal-01804271

Scientific Books (or Scientific Book chapters)

[17] A. BONIFATI, G. FLETCHER, H. VOIGT, N. YAKOVETS.*Querying Graphs*, Synthesis Lectures on Data Management, Morgan & Claypool Publishers, October 2018, vol. 10, n^o 3, p. 1-184, https://hal.inria.fr/hal-01974379

Research Reports

[18] D. SCHWAB, A. FEJZA, L. VIAL, Y. ROBERT.*The GazePlay Project : Overview in February 2018*, LIG lab, February 2018, https://hal.archives-ouvertes.fr/hal-01981318

Other Publications

[19] H. IM, P. GENEVÈS, N. GESBERT, N. LAYAÏDA.*Backward Type Inference for XML Queries*, September 2018, working paper or preprint, https://hal.inria.fr/hal-01497857