Activity Report Grenoble - Rhône-Alpes 2019

Edition: 2020-03-23

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, until Mar 2019, HDR]

Faculty Members
Hervé Rivano [Team leader, INSA Lyon, Professor, HDR]
Walid Bechkit [INSA Lyon, Associate Professor]
Nicolas Chiabaut [Ecole de l'aménagement durable des territoires, Professor, from Apr 2019 until Aug 2019, HDR]
Oana Iova [INSA Lyon, Associate Professor]
Razvan Stanica [INSA Lyon, Associate Professor, HDR]
Fabrice Valois [INSA Lyon, Professor, HDR]

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

Technical Staff
Manoel Dahan [INSA Lyon, Engineer]

PhD Students
Abderrahman Ben Khalifa [INSA Lyon, PhD Student]
Rodrigue Domga Komguem [University of Yaoundé 1, PhD Student]
Mohamed Anis Fekih [Univ de Lyon, PhD Student]
Kawtar Lasri [INPT Rabat, Maroc, PhD Student, from Apr 2019]
Ichrak Mokhtari [INSA Lyon, PhD Student, from Nov 2019]
Romain Pujol [INSA Lyon, PhD Student]
Solohaja Rabenjamina [Inria, PhD Student]

Post-Doctoral Fellows

Panagiota Katsikouli [Inria, Post-Doctoral Fellow, until Feb 2019]
Abdoul Aziz Mbacke [INSA Lyon, Post-Doctoral Fellow, until Nov 2019]
Jad Oueis [INSA Lyon, Post-Doctoral Fellow, until Apr 2019]

Visiting Scientists

Sami Abdellatif [Biskra University, Algeria, from Nov 2019]
Abdelmalik Bachir [Biskra University, Algeria, from Nov 2019]
Abdessamed Sassi [Oum El Bouaghi University, Algeria, until Oct 2019]
Panagiota Katsikouli [Technical University of Denmark, Denmark, from Mar 2019]

Administrative Assistants

Laetitia Gauthe [Inria, Administrative Assistant]
Giulietta Randisi-Lacoste [INSA Lyon, until Sep 2019]
Linda Soumari [INSA Lyon, from Oct 2019]

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 that 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 network 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 breakthroughs in micro and nano technologies are indeed enabling dense deployments of low-cost sensing devices that produce reliable enough measurements of physical phenomena 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 are 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 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 considers 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 an 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 Atmo Auvergne-Rhône-Alpes.

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 works 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 arises: how to cope with a dense set of M2M low bit rate traffics from energy and computing power constrained devices while classic cellular infrastructures 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 itself through device heterogeneity in terms of available communication technologies (not all devices have a long-range connection, or their 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 led by compressive sensing solutions. Our objective is to get rid of the assumption of knowing 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 dynamics behind several key applications. The goal is to provide a taxonomy of the applications according to the data properties in terms of stationarity, dynamics, 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. Consequently, 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 holds the PEDR (2018-2022).
- Hervé Rivano holds the PEDR (2017-2021).
- Razvan Stanica holds the PEDR (2016-2020).
- Razvan Stanica obtained his HDR from the University Lyon 1 / INSA Lyon, in November 2019.

5.1.1. Awards

- Ahmed Boubrima is runner-up (*accessit*) for the Gilles Kahn thesis prize 2019.

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 Pi, 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 plateform 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 lead by Khaled Boussetta. The mutualization of the software tools will allow us to conduct multi sites experiments, at Lyon and at Paris. Since 2019, this platform is used in the European Project Interreg Med ESMARTCITY and for the PHC Ulysses (joint collaboration with Nimbus Center, Ireland).

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 has 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.

6.5.3. 3M'air sensor platform

We developed the 3M'Air sensor platform to be used in participatory sensing of temperature and air quality. We have built our own nodes equipped with multiple sensors measuring Nitrogen-Dioxide (NO2), Particulate Matter (PM1, PM2.5, PM10), temperature and humidity. They are battery-powered and equipped with a GPS module to have the position of the measurements. Data are stored on a micro SD card and at the same time sent over LoRa to a server we have developed that is responsible to store these data for future analyses. A web platform has also been developed to display the collected concentration measurements in real time. This developed solution is used in several participatory planned measurement campaigns in Lyon city.

7. New Results

7.1. Wireless network deployment

Participants: Walid Bechkit, Ahmed Boubrima, Oana Iova, Rodrigue D. Komguem, Abdoul-Aziz Mbacke, Jad Oueis, Hervé Rivano, Razvan Stanica, Fabrice Valois

7.1.1. Deployment of wireless sensor networks for air quality mapping

Wireless sensor networks (WSN) are widely used in environmental applications where the aim is to sense a physical phenomenon such as temperature, air pollution, etc. A careful deployment of sensors is necessary in order to get a better knowledge of these physical phenomena while ensuring the minimum deployment cost [18]. In this work, we focus on using WSN for air pollution mapping and tackle the optimization problem of sensor deployment [3]. Unlike most of the existing deployment approaches, which are either generic or assume that sensors have a given detection range, we define an appropriate coverage formulation based on an interpolation formula that is adapted to the characteristics of air pollution sensing. We derive from this formulation two deployment models for air pollution mapping using integer linear programming while ensuring the connectivity of the network and taking into account the sensing error of nodes. We analyze the theoretical complexity of our models and propose heuristic algorithms based on linear programming relaxation and binary search. We perform extensive simulations on a dataset of the Lyon city, France in order to assess the computational complexity of our proposal and evaluate the impact of the deployment requirements on the obtained results.

7.1.2. Characterization of radio links in case of a ground deployment

In this work, we are interested in characterizing the link properties of a wireless sensor network with nodes deployed at ground level [5]. Such a deployment is fairly common in practice, e.g., when monitoring the vehicular traffic on a road segment or the status of infrastructures such as bridges, tunnels or dams. However, the behavior of off-the-shelf wireless sensor nodes in these settings is not yet completely understood. Through a thorough experimentation campaign, we evaluated not only the impact of the ground proximity on the wireless links, but also the impact of some parameters such as the packet payload, the communication channel frequency and the topography of the deployment area. Our results show that a ground-level deployment has a significant negative impact on the link quality, while parameters such as the packet size produce unexpected consequences. This allows us to parameterize classical theoretical models in order to fit a ground-level deployment scenario. Finally, based on the lessons learned in our field tests, we discuss some considerations that must be taken into account during the design of communication protocols and before the sensor deployment in order to improve network performance.

7.1.3. Sensor deployment in linear wireless sensor networks using the concept of virtual node

In a multi-hop wireless sensor network with a convergecast communication model, there is a high traffic accumulation in the neighborhood of the sink. This area constitutes the bottleneck of the network since the sensors deployed within it rapidly exhaust their batteries. In this work, we consider the problem of sensors deployment for lifetime maximization in a linear wireless sensor network [6]. Existing approaches express the deployment recommendations in terms of distance between consecutive sensors. Solutions imposing such constraints on the deployment may be costly and difficult to manage. We propose a new approach where the network is formed of virtual nodes, each associated to a certain geographical area. An analytical model of the network traffic per virtual node is proposed and a greedy algorithm to calculate the number of sensors that should form each virtual node is presented. Performance evaluation shows that the greedy deployment can improve the network lifetime by up to 40%, when compared to the uniform deployment. Moreover, the proposed approach outperforms the related work when complemented by a scheduling algorithm which reduces the messages overhearing. It is also shown that the lifetime of the network can be significantly improved if the battery capacity of each sensor is dimensioned taking into account the traffic it generates or relays.

7.1.4. Core network function placement in self-deployable mobile networks

Emerging mobile network architectures (e.g., aerial networks, disaster relief networks) are disrupting the classical careful planning and deployment of mobile networks by requiring specific 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 this work, 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 [9], [15] [15]. We propose a novel centrality metric, the flow centrality, which measures a node capacity of receiving the total amount of flows in the network. We show that in order to maximize the amount of exchanged traffic 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 flow centrality. We first compare our proposed metric to other state of the art centralities. Then, we highlight the significant traffic loss occurring when the core network entity is not placed on the node with the maximum flow centrality, which could reach 55% in some cases.

7.1.5. Cyber physical systems and Internet of things: emerging paradigms on smart cities

A city is smart when investment in traditional and modern infrastructure, human and social capital, fuel well being, high quality of life, and sustainable economic development. The Smart City paradigm is driven by technological evolution in the field of Information and Communication Technologies, and more specifically the paradigms of Internet of Things, Industrial Internet of Things and their confluence with Cyber Physical Systems [12]. Smart Cities present a number of application domains that are related to their critical infrastructures, including energy and transport. These domains present needs similar to the industrial manufacturing environment utilizing smart devices and employing control automation for their applications. They could thus be labeled as *industrial domains* in the wider sense. This work presents three application domains associated with Smart Cities, namely Smart Lighting, Smart Buildings / Energy, and Smart Urban Mobility, identifies their requirements and challenges and reviews existing solutions.

7.2. Wireless data collection

Participants: Oana Iova, Abderrahman Ben Khalifa, Razvan Stanica

7.2.1. Reliable and efficient support for downward traffic in RPL

Modern protocols for wireless sensor networks efficiently support multi-hop upward traffic from many sensors to a collection point, a key functionality enabling monitoring applications. However, the ever-evolving scenarios involving low-power wireless devices increasingly require support also for downward traffic, e.g., enabling a controller to issue actuation commands based on the monitored data. The IETF Routing Protocol for Low-power and Lossy Networks (RPL) is among the few tackling both traffic patterns. Unfortunately, its support for downward traffic is significantly unreliable and inefficient compared to its upward counterpart. We tackle this problem by extending RPL with mechanisms inspired by opposed, yet complementary, principles [7]. At one extreme, we retain the route-based operation of RPL and devise techniques allowed by the standard but commonly neglected by popular implementations. At the other extreme, we rely on flooding as the main networking primitive. Inspired by these principles, we define three base mechanisms, integrate them in a popular RPL implementation, analyze their individual and combined performance, and elicit the resulting tradeoffs in scalability, reliability, and energy consumption. The evaluation relies on simulation, using both real-world topologies from a smart city scenario and synthetic grid ones, as well as on testbed experiments validating our findings from simulation. Results show that the combination of all three mechanisms into a novel protocol, T-RPL *i)* yields high reliability, close to the one of flooding, *ii)* with a low energy consumption, similar to route-based approaches, and *iii)* improves remarkably the scalability of RPL w.r.t. downward traffic.

7.2.2. Performance evaluation of LED-to-camera communications

The use of LED-to-camera communication opens the door to a wide range of use cases and applications, with diverse requirements in terms of quality of service. However, while analytical models and simulation tools exist for all the major radio communication technologies, the only way of currently evaluating the performance of a network mechanism over LED-to-camera is to implement and test it. Our work aims to fill this gap by proposing a Markov-modulated Bernoulli process to model the wireless channel in LED-to-camera communications, which is shown to closely match experimental results [11]. Based on this model, we develop and validate *CamComSim*, the first network simulator for LED-to-camera communications.

7.2.3. Performance evaluation of channel access methods for dedicated IoT networks

Networking technologies dedicated for the Internet of Things are different from the classical mobile networks in terms of architecture and applications. This new type of network is facing several challenges to satisfy specific user requirements. Sharing the communication medium between (hundreds of) thousands of connected nodes and one base station is one of these main requirements, hence the necessity to imagine new solutions, or to adapt existing ones, for medium access control. In this work, we start by comparing two classical medium access control protocols, CSMA/CA and Aloha, in the context of Internet of Things dedicated networks [13]. We continue by evaluating a specific adaptation of Aloha, already used in low-power wide area networks, where no acknowledgement messages are transmitted in the network. Finally, we apply the same concept to

CSMA/CA, showing that this can bring a number of benefits. The results we obtain after a thorough simulation study show that the choice of the best protocol depends on many parameters (number of connected objects, traffic arrival rate, allowed retransmission number), as well as on the metric of interest (e.g. packet reception probability or energy consumption).

7.2.4. *On the use of wide channels in WiFi networks*

An increased density of access points is common today in WiFi deployments, and more and more parameters need to be configured in such networks. In this work, we question current industrial guidelines for both residential and enterprise scenarios [14]. More precisely, we investigate the joint channel, power, and carrier sense threshold allocation problem in IEEE 802.11ac networks, showing that the current practice, which is to use narrower channels at maximum power when the deployment is dense, yields much worse performance than a solution using the widest possible channel with a much lower power.

7.3. Network data exploitation

Participants: Florent Delaine, Panagiota Katsikouli, Hervé Rivano, Razvan Stanica

7.3.1. *Calibration algorithms for environmental sensor networks*

The recent developments in both nanotechnologies and wireless technologies have enabled the rise of small, low cost and energy efficient environmental sensing devices. Many projects involving dense sensor networks deployments have followed, in particular within the Smart City trend. If such deployments are now within economical and technical reach, their maintenance and reliability remain however a challenge. In particular, reaching, then maintaining, the targeted quality of measurement throughout deployment duration is an important issue. Indeed, factory calibration is too expensive for systematic application to low-cost sensors and as these sensors are usually prone to drifting because of premature aging. In addition, there are concerns about the applicability of factory calibration to field conditions [4]. These challenges have fostered many researches on in situ calibration. In situ means that the sensors are calibrated without removing them from their deployment location, preferably without physical intervention, often leveraging their communication capabilities. It is a critical challenge for the economical sustainability of networks with large scale deployments. In this work, we focus on in situ calibration methods for environmental sensor networks. We propose a taxonomy of the methodologies in the literature. Our classification relies on both the architecture of the network of sensors and the algorithmic principles of the calibration methods. This review allows us to identify and discuss two main challenges: how to improve the performance evaluation of such methods and how to enable a quantified comparison of these strategies?

7.3.2. *Characterizing and Removing Oscillations in Mobile Phone Location Data*

Human mobility analysis is a multidisciplinary research subject that has attracted a growing interest over the last decade. A substantial amount of such recent studies is driven by the availability of original sources of real-world information about individual movement patterns. An important task in the analysis of mobility data is reliably distinguishing between the stop locations and movement phases that compose the trajectories of the monitored subjects. The problem is especially challenging when mobility is inferred from mobile phone location data: here, oscillations in the association of mobile devices to base stations lead to apparent user mobility even in absence of actual movement [10]. In this work, we leverage a unique dataset of spatiotemporal individual trajectories that allows capturing both the user and network operator perspectives in mobile phone location data, and investigate the oscillation phenomenon. We present probabilistic and machine learning approaches for detecting oscillations in mobile phone location data, and a filtering technique for removing those. Our analyses and comparison with state-of-the-art approaches demonstrate the superiority of our solution, both in terms of removed oscillations and of error with respect to ground-truth trajectories.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- We have contracted a first bilateral contract with Total (2018-2021) 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 and evaluated by LQA.
- 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.
- 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, 2019
  Participants: Razvan Stanica
  The partners of this project, supported by the *Fédération d'Informatique de Lyon*, are: CITI, LIP.
  WLANs (Wireless Local Area Networks) are typically based on IEEE 802.11 (known as WiFi). However, WLANs are prone to performance issues such as unfairness and inefficiencies. 802.11 includes a Rate Adaptation (RA) mechanism that allows user devices to change their transmission rate with regard to the current quality of the radio channel. The RA mechanism is based on preset values that may lead to suboptimal WLAN performance. Our goal is to address this issue by making fine adjustments to the parameters related to the RA mechanism. The search for an adequate setting is made complex due to the vast number of parameters to be considered that precludes the finding of general closed-form expressions. We propose to explore a data-driven approach based on techniques from Machine Learning to design an adaptive and distributed solution.
- 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 explores 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, 2016-2019.
  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, 2017-2019
  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 information 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
  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
  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)* a complete 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 2018 - 2020.
  Participants: Solohaja Rabenjamina, Razvan Stanica.
  The ANR CoWorkWorlds (Sustainability and spatiality in co-workers' mobility practices) project is led by ENTPE. Its focus is on the study of co-working environments, and more precisely on the mobility behavior 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. GDR CNRS RSD - Pôle ResCom

- Ongoing participation (since 2006)
  Communication networks, working groups of GDR ASR/RSD, CNRS (https://gdr-rsd.cnrs.fr/pole_rescom). Hervé Rivano is member of the scientific committee of ResCom.

9.2.3. EquipEx

- SenseCity
  We have coordinated the participation of several Inria teams to the SenseCity EquipEx. Within the SenseCity project, several small reproductions 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 test 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 built, the information system is operational since April 2018.

9.3. European Initiatives

9.3.1. Collaborations in European Programs, Except FP7 & H2020

- Herve Rivano is member of European COST action CA18204 - Dynamics of placemaking and digitization in Europe´s cities on behalf of Ecole Urbaine de Lyon and Labex IMU.

  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.
- **Digital Catapul, London, UK**. 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.

- **Rice University**. Collaboration around network deployment and data assimilation for air quality monitoring with the group of Prof. Edward W. Knightly.
- **University of Edinburgh, UK**. Joint publications and visits to/from the group of Dr. Paul Patras.
- **Biskra University, Algeria**. Joint publications and visits from Prof. Abdelmalik Bachir.

9.4.2. Participation in Other International Programs

9.4.2.1. PHC Campus France

- **University College Cork, Ireland**. PHC Ulysses (2019-2021) on real-world characterisation of long range wireless newtorks, a collaboration with Khaled Abdelfadeel.
- **INPT Rabat, Morocco**. PHC Toubkal (2019-2021) on efficient data collection for smart building and smart city applications, a collaboration with the group of Prof. Loubna Echabbi.
- **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: visiting professor at INSA Lyon (november, 2019).
- Ravi Mazumdar, Waterloo University, Canada, visiting scientist at INSA Lyon (february, 2019).
- Priscilla Solis, Professor, Brasilia University, Brazil, visiting the Agora team to prepare a sabbatical.

9.5.1.1. Internships

- Sami Abdelatif, PhD student, Biskra University, Algeria: visiting professor at INSA Lyon (november, 2019).

9.5.2. Visits to International Teams

9.5.2.1. Research Stays Abroad

- Mihai Popescu visited the group of Prof. Gabriela Czibula, at University of Cluj-Napoca, Romania (2 periods of 1 month duration: April and July 2019).
- Fabrice Valois visited Prof. Catherine Rosenberg, University of Waterloo, Canada (6 weeks between January and March 2019).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- Oana Iova was the scientific chair of the first Low Power and Wide Area Networks Days (GDR RSD), 11-12 July 2019, Lyon.
- Oana Iova was co-chair of NewNets 2019 - The 1st Workshop on Emerging Technologies and Trends in Engineering Low-Power Networks (collocated with IEEE WF-IoT), April 15, 2019, Limerick, Ireland.
- Razvan Stanica co-organized a workshop on New Architectures and Services for Autonomous and Flexible Cellular Networks, within the Jacques Cartier Meetings organised in November in Montreal (Canada).

10.1.1.2. Member of the Organizing Committees

- Oana Iova was on the Organizing Committee of the first Low Power and Wide Area Networks Days (GDR RSD), 11-12 July 2019, Lyon.
- Hervé Rivano was on the Organizing Committee of the first Low Power and Wide Area Networks Days (GDR RSD), 11-12 July 2019, Lyon.
- Fabrice Valois was on the Organizing Committee of the first Low Power and Wide Area Networks Days (GDR RSD), 11-12 July 2019, Lyon.

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- Oana Iova was program co-chair of CoRes 2019: 4ème Rencontre Francophone sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication, June 3-4, 2019, Toulouse (France).

10.1.2.2. Member of the Conference Program Committees

- Oana Iova was in the TPC of the following conferences: IAUV, UrbCom, IoT-HPC, ACM SIGCOMM AINTEC, AdHoc-Now, EWSN, IEEE DIPI.
- Razvan Stanica was in the TPC of the following conferences: IEEE ICC, IEEE GlobeCom, IEEE CCNC, IEEE WCNC, IEEE PIMRC, IFIP WD, ICIN, ISNCC, GIoT.
- Fabrice Valois was in the TPC of the following conferences: IEEE Globecom, IEEE ICC, IEEE ICT, IEEE WCNC, WiSARN.

10.1.2.3. Reviewer

- Oana Iova was a reviewer for the following conference: IEEE INFOCOM 2019 (demo and posters).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Fabrice Valois is associated editor for Annals of Telecommunications (IF: 1.412).

10.1.3.2. Reviewer - Reviewing Activities

- Walid Bechkit was reviewer for the following journals: Elsevier Ad hoc Networks.
- Oana Iova was reviewer for the following journals: IEEE Transactions on Mobile Computing, Elsevier Sustainable Computing, Informatics and Systems.
- Razvan Stanica was a reviewer for the following journals: IEEE Transactions on Mobile Computing, IEEE Communication Letters, Computer Networks, Ad-Hoc Networks, Wireless Networks, IEEE Transactions on Network and Service Management, IEEE Access, Mobile Information Systems.

10.1.4. Invited Talks

- Hervé Rivano was invited speaker IMERIR, Perpignan, France.
- Hervé Rivano was invited speaker Forum Numerica - Learning Center, Campus SophiaTech, France.
- Hervé Rivano was invited speaker Institut d'Automne en Intelligence Artificielle, GDR IA, Lyon, France.
- Hervé Rivano was invited speaker Les Entretiens Jacques Cartier, Montréal, Canada.
- Hervé Rivano was invited speaker IEEE Societal Automation, Krakow, Poland.
- Hervé Rivano was chair of the panelist A l'école de l'anthropocène, Lyon, France.
- Fabrice Valois was invited speaker IEEE Canada Distinguished Lecture, University of Waterloo, Canada.
- Fabrice Valois was invited speaker Les Entretiens Jacques Cartier, Montréal, Canada.
- Fabrice Valois was invited speaker Mois des Cultures Numériques, Institut Français, Sfax, Tunisie.

10.1.5. Leadership within the Scientific Community

- Walid Bechkit is a nominated member in the scientific committee of the Fédération d'Informatique de Lyon (FR 2000 CNRS).
- Oana Iova is the scientific leader of the steering committee of the LPWAN Days (GDR RSD).
- Hervé Rivano is member of the steering committee of the ResCom axis of the RSD CNRS GdR.
- Hervé Rivano is member ot the steering committee of Ecole Urbaine de Lyon (PIA Institut Convergence), in charge of Smart and Learning Cities.
- Hervé Rivano was co-representative of the Labex IMU at the Smart City Expo World Congres (Barcelona).
- 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).

10.1.6. Scientific Expertise

- Hervé Rivano is member of the Scientific Committee of the Digital League Regional Cluster.
- Fabrice Valois was a member in the recruitment committee of an Associate Professor in Computer Science at Université de Clermont Auvergne.
- Fabrice Valois was a member in two recruitment committees of Associate Professors in Computer Science at Université de Cergy-Pontoise.

10.1.7. 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 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: Fabrice Valois, IP Networks, 36h, L3, Telecom. Dpt. INSA Lyon.

Licence: Fabrice Valois, Medium Access Control, 54h, L3, Telecom. Dpt. INSA Lyon.

Licence: Fabrice Valois, Basic networking, 20h, L3, IST / Telecom. Dpt. INSA Lyon (lectures given in english).

Licence: Oana Iova, Network and System Programming, 108h, L3, Telecom. Dpt. INSA Lyon.

Licence: Oana Iova, Introduction to research, 20h, L3, Telecom. Dpt. INSA Lyon.

Licence: Oana Iova, Smart city: networking challenges, 8h, L3, Smart Program, INSA Lyon (lectures given in english).

Licence: Oana Iova, Internet of connected things, 3h, L3, Innov Program, INSA Lyon (lectures given in english).

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

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

Licence: Hervé Rivano, Programming robot control, 20h, L2, INSA Lyon.

Licence: 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, 66h, M1, Telecom. Dpt. INSA Lyon.

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

Master: Oana Iova, IoT technical project, 8h, M2, Telecom. Dpt. INSA Lyon.

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

Master: Hervé Rivano, Smart Cities, Master Cities, Environment and Urbanism, University of Lyon.

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

Master: Razvan Stanica, Network Science, 10h, M2, 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.

Hervé Rivano is responsible of the Smart program (international teaching program with Tohoku University and Tokyo University) about Smart Cities.

Hervé Rivano is responsible of the IoT specialization of the Innov program (INSA Lyon and US students).

Hervé 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 : 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: Kawtar Lasri, Data collection and distributed spatial coordination in LPWAN networks, since 01/2019. Advisors: Oana Iova, Fabrice Valois.

PhD in progress: Ichrak Mokhtari, Spatio-temporal analysis of pollution data from low cost sensors, since 11/2019. 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 committees:
  - G. Bu, Fiabilité, sécurité et vie privée dans les WBAN, LIP6, Sorbonne Université
  - B. Grozev, Conférences vidéos WebRTC à routage optimisé, Icube, Université de Strasbourg
- Hervé Rivano was a member in the following PhD defense committees:
  - M. Smache, La sécurité des réseaux déterministe de l'Internet industriel des objets (IIoT), CEA Leti, Mines Saint Etienne
  - H. Mazouzi, Algorithmes pour le déchargement de tâches sur serveurs de périphérie mobile, L2TI, Université de Paris Nord
- Razvan Stanica was a member in the following PhD defense committee:
  - M. M. Merah, Conception et réalisation d'un lien Light-Fidelity multi-utilisateur en intérieur, LISV, Université de Versailles Saint Quentin en Yvelines
- Fabrice Valois was a reviewer in the following PhD defense committee:
  - H. Mroue, Développement de réseaux radio d'objets connectés pour les Villes Intelligentes : améioration de la Qualité de Service du réseau LoRa, IETR, Université de Nantes
- Fabrice Valois was a member in the following PhD defense committees:
  - R. T. Hermeto, Standard Improvements and Predictable Performance for Industrial Internet of Things in Indoor Deployments, ICube, Université de Strasbourg
  - H. Chour, Full-Duplex Device-to-Device Communication for 5G Network, IETR, Centrale Supélec

10.3. Popularization

- Hervé Rivano, Atelier de recherche participative sur la mesure de la qualité de l'air, Ecole Urbaine de Lyon, France.
- Hervé Rivano, Développement et expérimentation du kit pédagogique *Ca va chauffer*, Ecole Jules Guesde, Villeurbanne, France.
- Hervé Rivano, Aterlier autour du kit pédagogique *Ca va chauffer*, Super Demain, Métropole de Lyon.

10.3.1. Articles and contents

- Hervé Rivano is co-author of the blog post *Représenter l'intangible : les défis de la visualisation des données numériques environnementales. Montrer la pollution de l'air, le cas d'Air To Go*https://medium.com/anthropocene2050/
- Hervé Rivano is organizer and moderator of the podcast *Quelle ville intelligente est possible ?*, https://www.sondekla.com/user/event/10011

- Hervé Rivano is moderator of the podcast *Anthropocène et outils numériques*, https://www.sondekla.com/user/event/9805

10.3.2. Interventions

- Walid Bechkit was invited speaker about *3M'Air Mesures citoyennes Mobiles et Modélisation: qualité de l'air et îlots de chaleur à Lyon*, Colloque National Capteurs et Sciences Participatives, Paris, 2019.
- Oana Iova, Hervé Rivano and Fabrice Valois was involved in *Les rêveries lumineuses de Léonard*, https://popsciences.universite-lyon.fr/agenda/les-reveries-lumineuses-de-leonard-fete-des-lumieres-2019/, contribution de l'Ecole Urbaine de Lyon pour l'Université de Lyon, Fête des Lumières, 2019.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. BOUBRIMA.*Deployment and Scheduling of Wireless Sensor Networks for Air Pollution Monitoring*, Insa Lyon, March 2019, https://hal.inria.fr/tel-02446568

[2] R. STANICA.*From Networks to Data and Back Again : A Story of Wireless Networks in the 21st Century*, Institut National des Sciences Appliquées de Lyon, November 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02446174

Articles in International Peer-Reviewed Journal

[3] A. BOUBRIMA, W. BECHKIT, H. RIVANO.*On the Deployment of Wireless Sensor Networks for Air Quality Mapping: Optimization Models and Algorithms*, in "IEEE/ACM Transactions on Networking", August 2019, vol. 27, n^{o} 4, p. 1629-1642 [*DOI :* 10.1109/TNET.2019.2923737], https://hal.inria.fr/hal-02157476

[4] F. DELAINE, B. LEBENTAL, H. RIVANO.*In Situ Calibration Algorithms for Environmental Sensor Networks: a Review*, in "IEEE Sensors Journal", January 2019, vol. 19, n^{o} 15, p. pp 5968 - 5978 [*DOI :* 10.1109/JSEN.2019.2910317], https://hal.archives-ouvertes.fr/hal-02174938

[5] R. DOMGA KOMGUEM, R. STANICA, M. TCHUENTÉ, F. VALOIS.*Ground Level Deployment of Wireless Sensor Networks: Experiments, Evaluation and Engineering Insight*, in "Sensors", July 2019, vol. 19, n^{o} 15, 3358 [*DOI :* 10.3390/S19153358], https://hal.sorbonne-universite.fr/hal-02290185

[6] R. DOMGA KOMGUEM, R. STANICA, M. TCHUENTÉ, F. VALOIS.*Sensor Deployment in Wireless Sensor Networks with Linear Topology using Virtual Node Concept*, in "Wireless Networks", November 2019, vol. 25, n^{o} 8, p. 4947-4962 [*DOI :* 10.1007/S11276-019-02071-X], https://hal.inria.fr/hal-02354755

[7] T. ISTOMIN, O. IOVA, G. P. PICCO, C. KIRALY.*Route or Flood? Reliable and Efficient Support for Downward Traffic in RPL*, in "ACM Transactions on Sensor Networks", 2019, p. 1-36, forthcoming, https://hal.archives-ouvertes.fr/hal-02301882

[8] P. LEJOUX, A. FLIPO, N. ORTAR, N. OVTRACHT, S. SOUCHE-LECORVEC, R. STANICA.*Coworking, a Way to Achieve Sustainable Mobility? Designing an Interdisciplinary Research Project*, in "Sustainability", December 2019, vol. 11, n^{o} 24, 7161 [*DOI :* 10.3390/SU11247161], https://hal.archives-ouvertes.fr/hal-02419717

[9] 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

Invited Conferences

[10] P. KATSIKOULI, M. FIORE, A. FURNO, R. STANICA.*Characterizing and Removing Oscillations in Mobile Phone Location Data*, in "WoWMoM 2019 - 20th IEEE International symposium on a World of Wireless, Mobile and Multimedia Networks", Washington DC, United States, IEEE, June 2019, p. 1-9 [*DOI :* 10.1109/WOWMOM.2019.8793034], https://hal.inria.fr/hal-02110719

International Conferences with Proceedings

[11] A. DUQUE, A. DESPORTES, R. STANICA, H. RIVANO.*Performance Evaluation of LED-to-Camera Communications*, in "MSWiM 2019 - 22nd ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems", Miami Beach, United States, ACM, November 2019, p. 135-142 [*DOI :* 10.1145/3345768.3355922], https://hal.inria.fr/hal-02270454

[12] A. P. KALOGERAS, H. RIVANO, L. FERRARINI, C. ALEXAKOS, O. IOVA, S. RASTEGARPOUR, A. A. MBACKÉ.*Cyber Physical Systems and Internet of Things: Emerging Paradigms on Smart Cities*, in "SA 2019 - First International Conference on Societal Automation", Krakow, Poland, IEEE, September 2019, p. 1-13 [*DOI :* 10.1109/SA47457.2019.8938083], https://hal.inria.fr/hal-02427190

[13] A. B. KHALIFA, R. STANICA.*Performance Evaluation of Channel Access Methods for Dedicated IoT Networks*, in "WD 2019 - 11th IFIP Annual Wireless Days Conference", Manchester, United Kingdom, April 2019, p. 1-6 [*DOI :* 10.1109/WD.2019.8734186], https://hal.inria.fr/hal-02042688

[14] S. MALEKMOHAMMADI, C. ROSENBERG, R. STANICA.*On the Use of Wide Channels in WiFi Networks*, in "LCN 2019 - 44th IEEE Conference on Local Computer Networks", Osnabruck, Germany, IEEE, October 2019, p. 1-4, https://hal.inria.fr/hal-02336041

[15] 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

[16] R. PUJOL, R. STANICA, F. VALOIS.*LTE User Association for Self-Deployable Networks in Disaster Management Scenarios*, in "ICT-DM 2019 - 6th International Conference on Information and Communication Technologies for Disaster Management", Paris, France, December 2019, https://hal.archives-ouvertes.fr/hal-02327011

[17] A. SASSI, M. BRAHIMI, W. BECHKIT, A. BACHIR.*Location Embedding and Deep Convolutional Neural Networks for Next Location Prediction*, in "LCN 2019 - 44th Annual IEEE Conference on Local Computer Networks", Osnabrück, Germany, IEEE, October 2019, p. 1-9, https://hal.archives-ouvertes.fr/hal-02357778

Scientific Books (or Scientific Book chapters)

[18] A. BOUBRIMA, W. BECHKIT, H. RIVANO.*On the Optimization of WSN Deployment for Sensing Physical Phenomena: Applications to Urban Air Pollution Monitoring*, in "Mission-Oriented Sensor Networks and Systems: Art and Science", Studies in Systems, Decision and Control, Springer, 2019, vol. 163, p. 99-145 [*DOI :* 10.1007/978-3-319-91146-5_4], https://hal.inria.fr/hal-02159770

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)
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.5. - Bayesian methods
A3.4.7. - Kernel methods
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
A6.3.5. - Uncertainty Quantification
A6.5.2. - Fluid mechanics
A6.5.4. - Waves

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, until Sep 2019]
Eugene Kazantsev [Inria, Researcher]
Florian Lemarié [Inria, Researcher]
Gurvan Madec [CNRS, Researcher, HDR]
Farshid Nazari [Inria, Starting Research Position, until Oct 2019]
Arthur Vidard [Inria, Researcher, HDR]

Olivier Zahm [Inria, Researcher]

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]
Clémentine Prieur [Univ Grenoble Alpes, Professor, HDR]

External Collaborator

Maëlle Nodet [Univ Grenoble Alpes, until Aug 2019]

Technical Staff

Nicolas Ducousso [Inria, Engineer, from Apr 2019]
Laurence Viry [Univ Grenoble Alpes, Engineer]

PhD Students

Rishabh Bhatt [Inria, from Dec 2019]
Simon Clement [Univ Grenoble Alpes, from Oct 2019]
Emilie Duval [Inria]
Mohamed Reda El Amri [IFPEN, until Apr 2019]
Maria Belen Heredia Guzman [IRSTEA]
Adrien Hirvoas [IFPEN]
Philomene Le Gall [Univ Grenoble Alpes]
Long Li [Université étrangère]
Arthur Macherey [Ecole centrale de Nantes]
Antoine Alexis Nasser [Inria, from Nov 2019]
Sophie Thery [Univ Grenoble Alpes]
Victor Trappler [Univ Grenoble Alpes]

Post-Doctoral Fellows

Matthieu Brachet [Inria]
Anass El Aouni [Inria, from Oct 2019]

Visiting Scientists

Nicholas Kevlahan [Université étrangère, until Aug 2019]
Victor Shutyaev [Université étrangère, Nov 2019]
Jose-Raphael Leon-Ramos [Université étrangère, from Dec 2019]

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 6.1.2) 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 6.4) 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 6.3 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 (6.3.1). 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 2019, a new contract has been signed with CMEMS (Copernicus Marine Environment Moniroting Service) in order to extent the multiresolution capabilities of the AGRIF and its integration into the NEMO ocean system.

- 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: Arthur 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: Arthur 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. Numerical Schemes for Ocean Modelling

Participants: Eric Blayo, Matthieu Brachet, Laurent Debreu, Emilie Duval, Christopher Eldred, Nicholas Kevlahan, Florian Lemarié, Gurvan Madec, 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). As an example, the stabilization of the coupling between barotropic (fast) and baroclinic (slow) modes in a three dimensional ocean model is a source of large numerical dissipation. This has been studied in details in [6].
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. The next COMMODORE meeting is planned for February 2020 and will take place in Hamburg. https://www.conferences.uni-hamburg.de/event/76.

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 favour the use of the non-hydrostatic compressible equations that removes the need for a 3D resolution at the price of reincluding acoustic waves [29].
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. In the context of the HEAT project, we worked on the analysis of dispersion analysis of continuous and discontinuous Galerkin methods of arbitrary degree of approximation ([31]), on compatible Galerkin schemes for shallow water model in 2D ([9]). 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 ([8]).

Accurate and stable implementation of bathymetry boundary conditions in ocean models remains a challenging problem. The dynamics of ocean flow often depend sensitively on satisfying bathymetry boundary conditions and correctly representing their complex geometry. Generalized (e.g.) terrain-following coordinates are often used in ocean models, but they require smoothing the bathymetry to reduce pressure gradient errors. Geopotential -coordinates are a common alternative that avoid pressure gradient and numerical diapycnal diffusion errors, but they generate spurious flow due to their “staircase” geometry. In [5], we introduce a new Brinkman volume penalization to approximate the no-slip boundary condition and complex geometry of bathymetry in ocean models. This approach corrects the staircase effect of -coordinates, does not introduce any new stability constraints on the geometry of the bathymetry and is easy to implement in an existing ocean model. The porosity parameter allows modelling subgrid scale details of the geometry. We illustrate the penalization and confirm its accuracy by applying it to three standard test flows: upwelling over a sloping bottom, resting state over a seamount and internal tides over highly peaked bathymetry features.

Figure (1) shows strong improvements obtained when the penalization method is used in comparison with traditional terrain following σ simulations. At 6km resolution, the penalization methods (Figure (1) d)), that takes into account details of bathymetry, allows to recover internal tide wave beams closed to the 3km simulation. (Figure (1) a)).

6.1.2. Coupling Methods for Oceanic and Atmospheric Models

Participants: Eric Blayo, Florian Lemarié, Sophie Thery, Simon Clément.

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 [49]. 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 (see [21] for an overview). Our activities can be divided into four general topics

1. *Stability and consistency analysis of existing coupling methods*: in [49] 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 in [37] 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 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 [56]. 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 [57]. More recently we have started to work specifically on the discretization methods for the parameterization of planetary boundary layers in climate models [51] 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).This work will continue in the framework of the PhD-thesis of C. Simon.

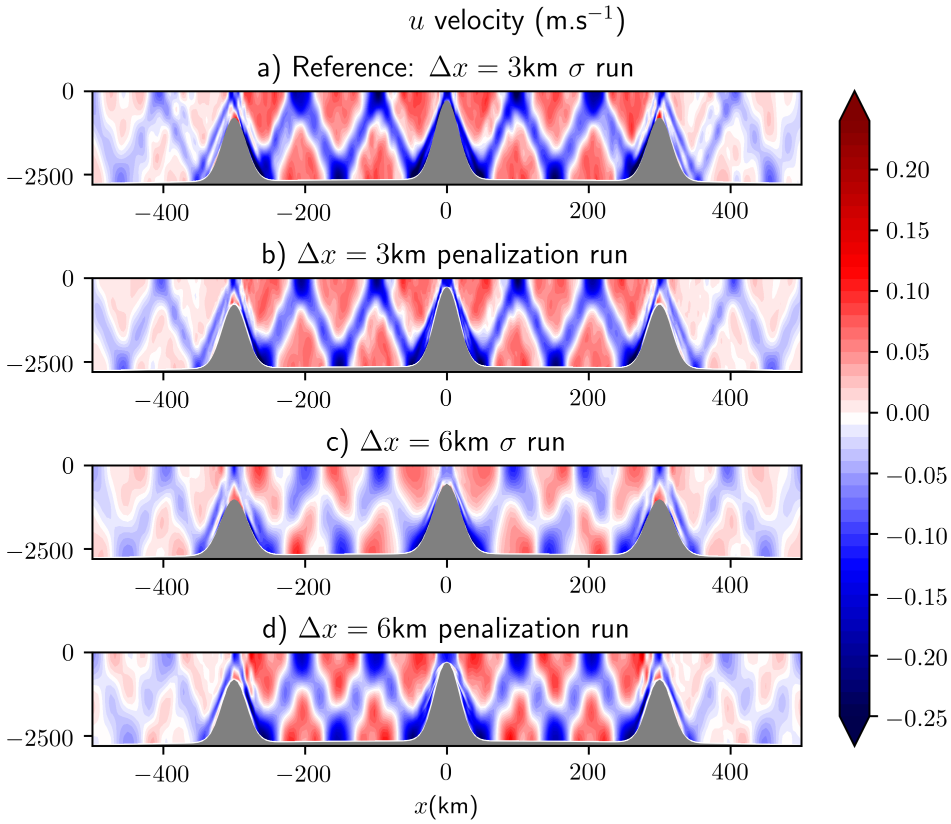


Figure 1. u velocity. Instantaneous solutions of the internal tide test case after 12 M2 tidal cycles of integration. (a) The reference σ coordinate run at 3 km resolution. (b) The penalized run at 3 km resolution. (c) The σ-coordinate run at 3 km resolution. (d) The penalized run at 6 km resolution.

3. *A simplified atmospheric boundary layer model for oceanic purposes*: Part of our activities within the IMMERSE project is dedicated to the development of a simplified model of the marine atmospheric boundary layer of intermediate complexity between a bulk parameterization and a full three-dimensional atmospheric model and to its integration to the NEMO general circulation model [24]. A constraint in the conception of such a simplified model is to allow an apt representation of the downward momentum mixing mechanism and partial re-energization of the ocean by the atmosphere while keeping the computational efficiency and flexibility inherent to ocean only modeling. A paper is in preparation and will be submitted in early 2020.
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. [50]) in large-scale realistic ocean-atmosphere coupled simulations [16], [53]. Moreover, within the ALBATROS project [23], 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 open discussion in Geoscientific Model Development journal on that topic [4].
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 [61] (paper in preparation). This work is done in the framework of S. Théry PhD (started in fall 2017). During the internship of C. Simon, efficient interface conditions have been derived at a (semi)-discrete level and can thus be systematically compared with the ones obtained from the continuous problem.

These topics are addressed through strong collaborations between the applied mathematicians and the climate community (Meteo-France, Ifremer, LMD, and LOCEAN). 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. In this context, a single-column version of the CNRM climate models has been designed and several coupling algorithms have been implemented (work done by S. Valcke, CERFACS). This model will be used to illustrate the relevance of our theoretical work in a semi-realistic context.

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 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.

Following this line of research, CASIS, a collaborative project with Mercator Océan started late 2017 until end of 2019 in order to extend these 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 [28], 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. Machine learning for parametrisation of the model dissipation.

Participants: Laurent Debreu, Eugene Kazantsev, Arthur Vidard, Olivier Zahm.

Artificial intelligence and machine learning may be considered as a potential way to address unresolved model scales and to approximate poorly known processes such as dissipation that occurs essentially at small scales. In order to understand the possibility to combine numerical model and neural network learned with the aid of external data, we develop a network generation and learning algorithm and use it to approximate nonlinear model operators. Beginning with a simple nonlinear equations like transport-diffusion and Burgers ones, we use artificially generated external data to learn the network by Adam algorithm [47]. Results show the possibility to approximate nonlinear, and even discontinuous dissipation operator with a quite good accuracy, however, several millions iterations are necessary to learn.

6.1.6. 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 is working on the design of methods to couple local nonhydrostatic models to larger scale hydrostatic ones (PhD started in Oct. 2018). 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). A prototype has been implemented, which allows for analytical solutions in simplified configurations and will make it possible to test different numerical approaches.

6.2. Assimilation of spatially dense observations

Participants: Elise Arnaud, François-Xavier Le Dimet, Arthur Vidard, Long Li, Emilie Rouzies.

6.2.1. 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 Jianwei Ma.

6.2.2. Observation error representation

Accounting for realistic observations errors is a known bottleneck in data assimilation, because dealing with error correlations is complex. Following a previous study on this subject, we propose to use multiscale modelling, more precisely wavelet transform, to address this question. In [3] we investigate the problem further by addressing two issues arising in real-life data assimilation: how to deal with partially missing data (e.g., concealed by an obstacle between the sensor and the observed system); how to solve convergence issues associated to complex observation error covariance matrices? Two adjustments relying on wavelets modelling are proposed to deal with those, and offer significant improvements. The first one consists in adjusting the variance coefficients in the frequency domain to account for masked information. The second one consists in a gradual assimilation of frequencies. Both of these fully rely on the multiscale properties associated with wavelet covariance modelling.

A collaborative project started with C. Lauvernet (IRSTEA) in order to make use of this kind of assimilation strategies on the control of pesticide transfer and it led to the co supervision of E. Rouzies PhD, starting in Dec 2019.

6.2.3. 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. This has been extended to water pollutant tracking as part of Long Li's PhD and published in [13]

6.3. Model reduction / multiscale algorithms

6.3.1. Parameter space dimension reduction and Model order reduction

Participants: Mohamed Reda El Amri, Arthur Macherey, Youssef Marzouk, Clémentine Prieur, Alessio Spantini, Ricardo Baptista, Daniele Bigoni, Olivier Zahm.

Numerical models describing the evolution of the system (ocean + atmosphere) contain a large number of parameters which are generally poorly known. The reliability of the numerical simulations strongly depends on the identification and calibration of these parameters from observed data. In this context, it seems important to understand the kinds of low-dimensional structure that may be present in geophysical models and to exploit this low-dimensional structure with appropriate algorithms. We focus in the team, on parameter space dimension reduction techniques, low-rank structures and transport maps techniques for probability measure approximation.

In [25], we proposed a framework for the greedy approximation of high-dimensional Bayesian inference problems, through the composition of multiple low-dimensional transport maps or flows. Our framework operates recursively on a sequence of “residual” distributions, given by pulling back the posterior through the previously computed transport maps. The action of each map is confined to a low-dimensional subspace that we identify by minimizing an error bound. At each step, our approach thus identifies (i) a relevant subspace of the residual distribution, and (ii) a low-dimensional transformation between a restriction of the residual onto this sub-space and a standard Gaussian. We prove weak convergence of the approach to the posterior distribution, and we demonstrate the algorithm on a range of challenging inference problems in differential equations and spatial statistics.

The paper [34] introduces a novel error estimator for the Proper Generalized Decomposition (PGD) approximation of parametrized equations. The estimator is intrinsically random: It builds on concentration inequalities of Gaussian maps and an adjoint problem with random right-hand side, which we approximate using the PGD. The effectivity of this randomized error estimator can be arbitrarily close to unity with high probability, allowing the estimation of the error with respect to any user-defined norm as well as the error in some quantity of interest. The performance of the error estimator is demonstrated and compared with some existing error estimators for the PGD for a parametrized time-harmonic elastodynamics problem and the parametrized equations of linear elasticity with a high-dimensional parameter space.

In the framework of Arthur Macherey’s PhD, we have also proposed in [26] algorithms for solving high-dimensional Partial Differential Equations (PDEs) that combine a probabilistic interpretation of PDEs, through Feynman-Kac representation, with sparse interpolation. Monte-Carlo methods and time-integration schemes are used to estimate pointwise evaluations of the solution of a PDE. We use a sequential control variates algorithm, where control variates are constructed based on successive approximations of the solution of the PDE. We are now interested in solving parametrized PDE with stochastic algorithms in the framework of potentially high dimensional parameter space. In [36], we consider 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. We are now working on the nonlinear generalization of active subspaces. Reduced models are also developed In the framework of robust inversion. In [43], we have combined a new greedy algorithm for functional quantization with a Stepwise Uncertainty Reduction strategy to solve a robust inversion problem under functional uncertainties. An ongoing work aims at further reducing the number of simulations required to solve the same robust inversion problem, based on Gaussian process meta-modeling on the joint input space of deterministic control parameters and functional uncertain variable. These results are applied to automotive depollution. This research axis is conducted in the framework of the Chair OQUAIDO.

6.4. Sensitivity analysis

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

6.4.1. 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.

6.4.2. Global sensitivity analysis

Participants: Elise Arnaud, Eric Blayo, Laurent Gilquin, Maria Belén Heredia, Adrien Hirvoas, Alexandre Janon, Henri Mermoz Kouye, Clémentine Prieur, Laurence Viry.

6.4.2.1. Global sensitivity analysis with dependent inputs

An important challenge for stochastic sensitivity analysis is to develop methodologies which work for dependent inputs. Recently, the Shapley value, from econometrics, was proposed as an alternative to quantify the importance of random input variables to a function. Owen [54] derived Shapley value importance for independent inputs and showed that it is bracketed between two different Sobol' indices. Song et al. [60] recently advocated the use of Shapley value for the case of dependent inputs. In a recent work [55], 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 [30].

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

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 [45]

In [11] 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 [15]. It has allowed to perform a global sensitivity analysis with input space dimension more than eighty, without any screening preliminary step.

6.4.2.3. Green sensitivity for multivariate and functional outputs
Participants: María Belén Heredia, Clémentine Prieur.

Another research direction for global SA algorithm starts with the report that most of the algorithms to compute sensitivity measures require special sampling schemes or additional model evaluations so that available data from previous model runs (e.g., from an uncertainty analysis based on Latin Hypercube Sampling) cannot be reused. One challenging task for estimating global sensitivity measures consists in recycling an available finite set of input/output data. Green sensitivity, by recycling, avoids wasting. These given data have been discussed, e.g., in [59], [58]. Most of the given data procedures depend on parameters (number of bins, truncation argument. . .) not easy to calibrate with a bias-variance compromise perspective. Adaptive selection of these parameters remains a challenging issue for most of these given-data algorithms. In the context of María Belén Heredia's PhD thesis, we have proposed a non-parametric given data estimator for agregated Sobol' indices, introduced in [48] and further developed in [44] for multivariate or functional outputs. This last work should be submitted soon.

6.4.2.4. Global sensitivity analysis for parametrized stochastic differential equations
Participants: Henri Mermoz Kouye, 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 (GSA) 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 [10]. The research on GSA for stochastic simulators is still ongoing, first in the context of the MATH-AmSud project FANTASTIC (Statistical inFerence and sensitivity ANalysis for models described by sTochASTIC differential equations) with Chile and Uruguay, secondly through the PhD thesis of Henri Mermoz Kouye, co-supervised by Clémentine Prieur, in collaboration with INRA Jouy.

6.5. Model calibration and statistical inference

6.5.1. Bayesian calibration

Participants: Maria Belén Heredia, Adrien Hirvoas, Clémentine Prieur.

Physically-based avalanche propagation models must still be locally calibrated to provide robust predictions, e.g. in long-term forecasting and subsequent risk assessment. Friction parameters cannot be measured directly and need to be estimated from observations. Rich and diverse data is now increasingly available from test-sites, but for measurements made along ow propagation, potential autocorrelation should be explicitly accounted for. In the context of María Belén Heredia's PhD, in collaboration with IRSTEA Grenoble, we have proposed in a comprehensive Bayesian calibration and statistical model selection framework with application to an avalanche sliding block model with the standard Voellmy friction law and high rate photogrammetric images. An avalanche released at the Lautaret test-site and a synthetic data set based on the avalanche were used to test the approach. Results have demonstrated i) the effciency of the proposed calibration scheme, and ii) that including autocorrelation in the statistical modelling definitely improves the accuracy of both parameter estimation and velocity predictions. 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. In the context of this work, it is necessary to develop algorithms of identification quantifying and ranking all the uncertainty sources. This work in done in collaboration with IFPEN.

6.5.2. Non-Parametric statistical inference 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 [39] and when only partial observational data are available. A paper on the non parametric estimation of the drift has been accepted recently [40] (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 [40], in collaboration with J.R. Leon (Montevideo, Uruguay) and P. Cattiaux (Toulouse). Recursive estimators have been also proposed by the same authors in [41], 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 [42], 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 [52].

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 [38], and would like to extend that results to hypo-elliptic diffusions.

6.6. Modeling and inference for extremes

Participants: Philomène Le Gall, Clémentine Prieur, Patricia Tencaliec.

In [19], we are 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 Anne-Catherine Favre (LGGE-Lab in Grenoble) and Philippe Naveau (LSCE, Paris).

In the context of Philomène Le Gall's PhD thesis, we are applying the aforementioned modeling of extreme precipitation with the aim of regionalizing extreme precipitation.

6.7. Land Use and Transport Models Calibration

Participants: 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 [46] and Thomas Capelle defended his in April 2017 and published his latest results in [2].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

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).

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 [1] has been funded by OQUAIDO. The Chair is reconducted for one year in 2020 and then a new contract should be approved by all partners for a new 4-years period.

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 : ANR MELODY (Bridging geophysics and MachinE Learning for the modeling, simulation and reconstruction of Ocean DYnamic)

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. Inria Challenge

Sea Uncertainty Representation and Forecast (SURF),

Coord : Airsea (A. Vidard),

Partenaires Inria : Ange, Cardamom, Fluminance, Lemon, Mingus, Defi

Partenaires extérieurs: BRGM, Ifremer, SHOM

8.2.3. 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 is co-advising the PhD thesis of Henri Mermoz Kouye, ine the framework of the Inria-INRA collaboration.

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 the group (co-chair). In particular, she will co-chair next GdR annual meeting in Aussois (May 2020). 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 and https://immerse-ocean.eu/

8.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: C3S

Project acronym: ERGO

Project title: Enabling an Ensemble of Data Assimilation for the Ocean

Duration: Février 2019 - juillet 2021

Coordinator: Arthur Vidard

Other partners: Cerfacs (France), Met Office (U.K.), CMRE (int, Italie)

Abstract: The scope of this contract is to improve ocean data assimilation capabilities at ECMWF, used in both initialization of seasonal forecasts and generation of coupled Earth System reanalyses. In particular it shall focus on i) improving ensemble capabilities in NEMO and NEMOVAR and the use of their information to represent background error statistics; ii) extend NEMOVAR capabilities to allow for multiple resolution in multi-incremental 3D-Var; iii) make better use of ocean surface observations. It shall also involve performing scout experiments and providing relevant diagnostics to evaluate the benefit coming from the proposed developments.

8.3.3. 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.

Partner : SAMO board

SAMO board is in charge of the organization of the SAMO (sensitivity analysis of model outputs) conferences, every three years. It is strongly supporter by the Joint Research Center of the European Commission. In 2019, Clémentine Prieur, which is part of this board, as also co-chair of a satellite event on the future of sensitivity analysis. A position paper is under construction, as a synthesis of the discussions hold in Barcelona (autumn 2019).

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.

8.4.2. Inria International Partners

F. Lemarié and L. Debreu collaborate with Hans Burchard and Knut Klingbeil from the Leibniz-Institut für Ostseeforschung in Warnemünde (Germany) [32], [12].

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

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

F.-X. Le Dimet is a Honorary Professor of the Institut of Mechanics, Ac.Sci. Vietnam.

F.-X. Le Dimet is a Honorary Professor of the Institut of Numerical Mathematics, Russian Ac.Sci.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

Alistair Adcroft (Princeton Univ.) visited the team in Jan. 2019

Jose R. León was visiting the team during two weeks. He is working with Clémentine Prieur, in collaboration with Pierre Etoré and Adeline Samson (DATA department of LJK) on UQ for models described by SDE.

Nicholas Kevlahan, from McMaster University (Canada) was a visiting scientist of the AIRSEA team for 10 months in 2018-2019.

Victor Shutyaev, from the Institut of Numerical Mathematics (Moscow, Russian Ac.Sci.) was visiting the team during two weeks to collaborate with F.-X. Le Dimet [17].

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Journal

9.1.1.1. Member of the Editorial Boards

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

9.1.1.2. Reviewer - Reviewing Activities

E. Blayo: reviewer for Ocean Modelling, Journal of Scientific Computing

F. Lemarié reviewed papers for Q. J. Roy. Meteorol. Soc. , Ocean Model. and J. Adv. Mod. Earth Sys.

9.1.2. Invited Talks

E. Blayo: Journées Tarantola : défis en géosciences, Paris, juin 2019

E. Blayo: Workshop Modélisation océan-atmosphère, Rennes, septembre 2019

F.-X. Le Dimet: The Mathematics of Climate and the Environment, IHP Paris 2019, Nov. 12-15

F. Lemarié has given an invited talk at the Banff International Research Station during the Physics-Dynamics Coupling in Earth System Models workshop [21]

9.1.3. Leadership within the Scientific Community

L. Debreu is the chair of the CNRS-INSU research program LEFE-MANU on mathematical and numerical methods for ocean and atmosphere https://insu.cnrs.fr/fr/methodes-mathematiques-et-numeriques-manu 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

9.1.4. Scientific Expertise

E. Blayo was the head of the HCERES evaluation committee for the Maison de la Simulation (Jan. 2019)

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.

9.1.5. Research Administration

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

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

E. Arnaud is in charge of the AMAC (Algorithmes, Modeles, Analyses, Calcul) department of the Jean Kuntzmann Lab.

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

Licence: E. Blayo, Mathematical analysis, 80h, L1, University Grenoble Alpes, France

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

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

Licence: C.Kazantsev, Mathematical tools for scientists and engineers, 100h, L1, University Grenoble Alpes, France

Licence: C.Kazantsev, Mathematics for engineer, 60h, L2, University Grenoble Alpes, France

Master: E. Blayo, Partial Differential Equations, 45h, M1, University Grenoble Alpes, France

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

Master: E. Arnaud: Advising students on apprenticeship, 28h, M2, University of Grenoble, France.

Doctorat: E. Blayo, A. Vidard, Introduction to Data Assimilation, 20h,University Grenoble Alpes, France

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 : Simon Clément, Study of discretized Schwarz waveform relaxation algorithms, M2R, applied mathematics, UGA, 6 months, E. Blayo and F. Lemarié

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

PhD in progress : Rishabh Batth, Asynchronousparallel in time schemes for data assimilation, December 2019, L. Debreu and A. Vidard.

PhD in progress : Simon Clément, Numerical analysis for reconciling in space and time the air-sea exchanges and their parameterization. October 2019, E. Blayo and F. Lemarié.

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: Mohamed Reda El Amri, Analyse d'incertitudes et de robustesse pour les modèles à entrées et sorties fonctionnelles,[1], Univ.Grenoble Alpes, April, 29, 2019. Clémentine Prieur, Céline Helbert (Centrale Lyon), funded by IFPEN, in the OQUAIDO chair program.

Post Doc in progress : Anass El Aouni, Multi-resolution techniques for ocean data assimilation, October 2019, A. Vidard

9.2.3. Juries

Arthur Vidard

Feb. 2019: PhD thesis of Olivier Guillet, INPT Toulouse (reporter)

E. Blayo:

Feb. 12, 2019: PhD thesis of Benoît Pinier, University of Rennes (reporter)

Apr. 5, 2019: HDR thesis of Sophie Ricci, University of Toulouse (reporter)

Oct. 21, 2019: PhD thesis of Jai Chowdhry Beeman, University Grenoble Alpes (president)

Dec. 18, 2019: PhD thesis of Alexandre Devers, University Grenoble Alpes (president)

Arthur Vidard was part of Inria CRCN recruitment juries in Bordeaux–Sud-ouest and Lille-Nord Europe

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

9.3. Popularization

9.3.1. Internal or external Inria responsibilities

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

9.3.2. Education

Ch. Kazantsev and E. Blayo are strongly involved in the creation and dissemination of pedagogic suitcases with mathematical activities designed for primary and secondary school (used by 10,000 – 12,000 pupils in 2018-2019). This is done in collaboration with the Rectorat de Grenoble.

E. Arnaud has animation of a "laboratoire des mathématiques", Pablo Neruda School, Saint Martin d'Hères

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

C. Kazantsev is a member of an International Inter-IREM commission, which work on the multi-languages problem for children in the mathematical learning. Three meetings take place in Paris during the year, the first was on September 28.

C. Kazantsev participated to the collaboration program with the Ecole Normale Supérieure de Bamako, Mali. She spent ten days in Bamako to present the modelisation teaching and the purpose and activities of "La Grange des Maths" group. Bamako, ENSup, January, 13-25

C. Kazantsev participated to the "Colloque du cinquantenaire des IREM" with a presentation of the activities of "La Grange des Maths", Besançon, May, 9-11

9.3.3. Interventions

National events:

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 at the "Culture and mathematical games Salon", Place St Sulpice, Paris, 23-26 May.

Ch. Kazantsev and E. Blayo participated in the "Fête de la Science", October, 12.

E. Arnaud. gave a presentation "(Se) tromper avec les chiffres" at the conference "Sciences et esprit critique, interroger les certitudes", Maison pour la sciences, Acadèmie de Grenoble, 8 nov. 2018, Grenoble.

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, 10.

C.Kazantsev participated at the "Remue-méninges festival" with the presentation of the "Maths à modeler" activities. Echirolles, April, 11.

C.Kazantsev participated to the "Sou des écoles", with the presentation of the "La Grange des Maths" activities. Varces, June, 15.

C.Kazantsev participated at the "Raout de Domène" with the presentation of mathematical animations. Domene, September, 1.

In educational institutions

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

C.Kazantsev presented mathematical animations to pupils of the Jean Mermoz scholl in Poisat for about 80 children during 3 hours. Poisat, May, 29.

E. Arnaud gave a talk "(Se) tromper avec les chiffres" for pupils of secondary schools on the 8th of March 2019 and the 2nd of May 2019.

9.3.4. 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] M. R. El Amri. *Uncertainty and robustness analysis for models with functional inputs and outputs*, Université Grenoble Alpes, April 2019, https://tel.archives-ouvertes.fr/tel-02433324

Articles in International Peer-Reviewed Journal

[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", September 2019, vol. 77, p. 101146:1-13 [*DOI :* 10.1016/J.COMPENVURBSYS.2017.04.009], https://hal.inria.fr/hal-01519654

[3] V. CHABOT, M. NODET, A. VIDARD.*Multiscale Representation of Observation Error Statistics in Data Assimilation*, in "Sensors", 2019, p. 1-19, forthcoming, https://hal.inria.fr/hal-02421699

[4] X. COUVELARD, F. LEMARIÉ, G. SAMSON, J.-L. REDELSPERGER, F. ARDHUIN, R. BENSHILA, G. MADEC.*Development of a 2-way coupled ocean-wave model: assessment on a global NEMO(v3.6)-WW3(v6.02) coupled configuration*, in "Geoscientific Model Development Discussions", August 2019, p. 1-36 [*DOI :* 10.5194/GMD-2019-189], https://hal.inria.fr/hal-02267188

[5] L. DEBREU, N. K.-R. KEVLAHAN, P. MARCHESIELLO.*Brinkman volume penalization for bathymetry in three-dimensional ocean models*, in "Ocean Modelling", January 2020, vol. 145, p. 1-13 [*DOI :* 10.1016/J.OCEMOD.2019.101530], https://hal.inria.fr/hal-02416084

[6] 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*, in "Journal of Computational Physics", December 2019, vol. 398, n^{o} 108875, p. 1-26 [*DOI :* 10.1016/J.JCP.2019.108875], https://hal.inria.fr/hal-01947706

[7] M. R. EL AMRI, C. HELBERT, O. LEPREUX, M. MUNOZ ZUNIGA, C. PRIEUR, D. SINOQUET.*Data-driven stochastic inversion via functional quantization*, in "Statistics and Computing", 2019, p. 1-17, forthcoming [*DOI :* 10.1007/S11222-019-09888-8], https://hal-ifp.archives-ouvertes.fr/hal-02291766

[8] C. ELDRED, T. DUBOS, E. KRITSIKIS.*A Quasi-Hamiltonian Discretization of the Thermal Shallow Water Equations*, in "Journal of Computational Physics", February 2019, vol. 379, p. 1-31 [*DOI :* 10.1016/J.JCP.2018.10.038], https://hal.inria.fr/hal-01847698

[9] C. ELDRED, D. LE ROUX.*Dispersion analysis of compatible Galerkin schemes on quadrilaterals for shallow water models*, in "Journal of Computational Physics", June 2019, vol. 387, p. 539-568 [*DOI :* 10.1016/J.JCP.2019.02.009], https://hal.archives-ouvertes.fr/hal-01916382

[10] P. ETORÉ, C. PRIEUR, D. K. PHAM, L. LI.*Global sensitivity analysis for models described by stochastic differential equations*, in "Methodology and Computing in Applied Probability", July 2019, p. 1-29, https://arxiv.org/abs/1811.08101 [*DOI :* 10.1007/S11009-019-09732-6], https://hal.archives-ouvertes.fr/hal-01926919

[11] 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", July 2019, vol. 187, p. 28-39 [*DOI :* 10.1016/J.RESS.2018.09.010], https://hal.inria.fr/hal-01558915

[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", 2019, vol. 100, p. ES109–ES115 [*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] A. W. MOORE, M. J. MARTIN, S. AKELLA, H. G. ARANGO, M. BALMASEDA, L. BERTINO, S. CIAVATTA, B. D. CORNUELLE, J. CUMMINGS, S. FROLOV, P. LERMUSIAUX, P. ODDO, P. R. OKE, A. STORTO, A. TERUZZI, A. VIDARD, A. T. WEAVER.*Synthesis of Ocean Observations Using Data Assimilation for Operational, Real-Time and Reanalysis Systems: A More Complete Picture of the State of the Ocean*, in "Frontiers in Marine Science", March 2019, vol. 6, n^{o} 90, p. 1-7 [*DOI :* 10.3389/FMARS.2019.00090], https://hal.inria.fr/hal-02421672

[15] C. PRIEUR, L. VIRY, E. BLAYO, J.-M. BRANKART.*A global sensitivity analysis approach for marine biogeochemical modeling*, in "Ocean Modelling", July 2019, vol. 139, n^{o} 101402, p. 1-38 [*DOI :* 10.1016/J.OCEMOD.2019.101402], https://hal.inria.fr/hal-01952797

[16] L. RENAULT, F. LEMARIÉ, T. ARSOUZE.*On the implementation and consequences of the oceanic currents feedback in ocean-atmosphere coupled models*, in "Ocean Modelling", September 2019, vol. 141, 101423 [*DOI :* 10.1016/J.OCEMOD.2019.101423], https://hal.inria.fr/hal-02190847

[17] V. SHUTYAEV, F.-X. LE DIMET, E. PARMUZIN.*Sensitivity of response functions in variational data assimilation for joint parameter and initial state estimation*, in "Journal of Computational and Applied Mathematics", 2019, p. 1-14 [*DOI :* 10.1016/J.CAM.2019.112368], https://hal.inria.fr/hal-02431701

[18] K. SMETANA, O. ZAHM, A. T. PATERA.*Randomized residual-based error estimators for parametrized equations*, in "SIAM Journal on Scientific Computing", March 2019, vol. 41, n^{o} 2, p. A900-A926, https://arxiv.org/abs/1807.10489 [*DOI :* 10.1137/18M120364X], https://hal.archives-ouvertes.fr/hal-01851462

[19] P. TENCALIEC, A.-C. FAVRE, P. NAVEAU, C. PRIEUR, G. NICOLET.*Flexible semiparametric Generalized Pareto modeling of the entire range of rainfall amount*, in "Environmetrics", 2019, p. 1-28, forthcoming [*DOI :* 10.1002/ENV.2582], https://hal.inria.fr/hal-01709061

Invited Conferences

[20] E. BLAYO.*Vers une meilleure simulation du couplage océan-atmosphère*, in "2019 - Journées Tarantola : défis en géosciences", Paris, France, June 2019, https://hal.inria.fr/hal-02415133

[21] F. LEMARIÉ.*An overview of the ocean-atmosphere coupling*, in "2019 - Physics-Dynamics Coupling in Earth System Models", Banff, Canada, October 2019, https://hal.inria.fr/hal-02418164

Conferences without Proceedings

[22] E. BLAYO, F. LEMARIÉ, C. PELLETIER, S. THÉRY.*Toward an improved simulation of ocean-atmosphere interactions*, in "2019 - conférence Modélisation Océan-Atmosphère", Rennes, France, September 2019, https://hal.inria.fr/hal-02415136

[23] F. LEMARIÉ, G. SAMSON, X. COUVELARD, G. MADEC, R. BOURDALLÉ-BADIE.*Recent developments in NEMO within the Albatross project*, in "DRAKKAR 2019 - Drakkar annual workshop", Grenoble, France, January 2019, https://hal.inria.fr/hal-02418218

[24] F. LEMARIÉ, G. SAMSON, J.-L. REDELSPERGER, G. MADEC, H. GIORDANI, R. BOURDALLÉ-BADIE.*Toward an improved representation of air-sea interactions in high-resolution global oceanic forecasting systems*, in "IMMERSE 2019 - IMMERSE Kick-Off Meeting", Grenoble, France, January 2019, https://hal.inria.fr/hal-02418229

Other Publications

[25] D. BIGONI, O. ZAHM, A. SPANTINI, Y. MARZOUK.*Greedy inference with layers of lazy maps*, June 2019, https://arxiv.org/abs/1906.00031 - 15 pages, 24 figures, https://hal.inria.fr/hal-02147706

[26] M. BILLAUD FRIESS, A. MACHEREY, A. NOUY, C. PRIEUR.*Stochastic methods for solving high-dimensional partial differential equations*, May 2019, https://arxiv.org/abs/1905.05423 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02140703

[27] M. BRACHET, J.-P. CHEHAB.*Fast and Stable Schemes for Phase Fields Models*, September 2019, https://arxiv.org/abs/1909.13511 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02301006

[28] 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

[29] M. HILT, L. ROBLOU, C. NGUYEN, P. MARCHESIELLO, F. LEMARIÉ, S. JULLIEN, F. DUMAS, L. DEBREU, X. CAPET, L. BORDOIS, R. BENSHILA, F. AUCLAIR.*Numerical Modelling of Hydraulic Control, Solitary Waves and Primary Instabilities in the Strait of Gibraltar*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02418114

[30] B. IOOSS, C. PRIEUR.*Shapley effects for sensitivity analysis with correlated inputs: comparisons with Sobol' indices, numerical estimation and applications*, November 2019, https://arxiv.org/abs/1707.01334 - working paper or preprint, https://hal.inria.fr/hal-01556303

[31] D. LE ROUX, C. ELDRED, M. TAYLOR.*Fourier analyses of continuous and discontinuous Galerkin methods of arbitrary degree of approximation*, May 2019, working paper or preprint, https://hal.inria.fr/hal-02125326

[32] F. LEMARIÉ, H. BURCHARD, K. KLINGBEIL, L. DEBREU.*Challenges and prospects for dynamical cores of oceanic models across all scales*, April 2019, PDEs on the sphere, Poster, https://hal.inria.fr/hal-02418199

[33] L. LI, A. VIDARD, F.-X. LE DIMET, J. MA.*Adaptive Image Assimilation for 2D Velocity Reconstruction*, July 2019, AOGS 2019 - 16th Annual Meeting Asia Oceania Geosciences Society, Poster, https://hal.inria.fr/hal-02263716

[34] K. SMETANA, O. ZAHM.*Randomized residual-based error estimators for the Proper Generalized Decomposition approximation of parametrized problems*, October 2019, https://arxiv.org/abs/1910.11837 - working paper or preprint [*DOI :* 10.11837], https://hal.inria.fr/hal-02335617

[35] V. TRAPPLER, E. ARNAUD, A. VIDARD, L. DEBREU.*Robust calibration of numerical models based on relative regret*, February 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02464780

[36] O. ZAHM, P. CONSTANTINE, C. PRIEUR, Y. MARZOUK.*Gradient-based dimension reduction of multivariate vector-valued functions*, November 2019, https://arxiv.org/abs/1801.07922 - working paper or preprint, https://hal.inria.fr/hal-01701425

References in notes

[37] 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° 2, p. 977-989 [*DOI :* 10.5194/GMD-10-977-2017], https://hal.inria.fr/hal-01406623

[38] 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

[39] 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° 3, p. 1236-1260 [*DOI :* 10.1016/J.SPA.2013.10.008], https://hal.archives-ouvertes.fr/hal-00739136

[40] 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° 1, p. 359-384, https://hal.archives-ouvertes.fr/hal-00877054

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

[42] 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° 1, p. 11-29 [*DOI :* 10.1080/02331888.2016.1259807], https://hal.archives-ouvertes.fr/hal-01405427

[43] 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

[44] F. GAMBOA, A. JANON, T. KLEIN, A. LAGNOUX, ET AL. .*Sensitivity analysis for multidimensional and functional outputs*, in "Electronic Journal of Statistics", 2014, vol. 8, n° 1, p. 575–603

[45] 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

[46] 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

[47] D. P. KINGMA, J. BA.*Adam: A Method for Stochastic Optimization*, 2014

[48] M. LAMBONI, H. MONOD, D. MAKOWSKI.*Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models*, in "Reliability Engineering & System Safety", 2011, vol. 96, n° 4, p. 450–459

[49] 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

[50] 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° RT-0464, https://hal.inria.fr/hal-01184711

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

[52] 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

[53] 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° 2, p. 812-833 [*DOI :* 10.1002/2017JC013294], https://hal.inria.fr/hal-01661645

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

[55] 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° 1, p. 986–1002 [*DOI :* 10.1137/16M1097717], https://hal.archives-ouvertes.fr/hal-01379188

[56] C. PELLETIER, F. LEMARIÉ, É. 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

[57] C. PELLETIER, F. LEMARIÉ, É. 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

[58] E. PLISCHKE, E. BORGONOVO, C. L. SMITH.*Global sensitivity measures from given data*, in "European Journal of Operational Research", 2013, vol. 226, n° 3, p. 536–550

[59] E. PLISCHKE.*An effective algorithm for computing global sensitivity indices (EASI)*, in "Reliability Engineering & System Safety", 2010, vol. 95, n° 4, p. 354–360

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

[61] S. THÉRY.*Algorithmes de Schwarz et conditions absorbantes pour le couplage océan-atmosphère*, in "CANUM 2018 - 44e Congrès National d'Analyse Numérique", Cap d'Agde, France, May 2018, https://hal.inria.fr/hal-01947885

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]
Nathalie Revol [Inria, Researcher]
Warwick Tucker [Inria, International Chair, Advanced Research Position]
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

Rikki Amit Inder Deo [Inria, Engineer, from Nov 2019]
Joris Picot [École Normale Supérieure de Lyon, part-time Engineer]

PhD Students

Florent Bréhard [École Normale Supérieure de Lyon, PhD Student, until Aug 2019]
Qian Chen [École normale supérieure de Rennes, PhD Student, until Aug 2019]
Adel Hamdi [Orange Labs, PhD Student, granted by CIFRE]
Huyen Nguyen [École Normale Supérieure de Lyon, PhD Student]
Alice Pellet–Mary [École Normale Supérieure de Lyon, PhD Student, until Oct 2019]
Miruna Rosca [Bitdefender, PhD Student]
Radu Titiu [Bitdefender, PhD Student]
Ida Tucker [École Normale Supérieure de Lyon, PhD Student]

Post-Doctoral Fellows

Chitchanok Chuengsatiansup [Inria, Post-Doctoral Fellow, until May 2019]

Alonso Gonzalez [École Normale Supérieure de Lyon, Post-Doctoral Fellow]
Dingding Jia [École Normale Supérieure de Lyon, Post-Doctoral Fellow, from Aug 2019]
Elena Kirshanova [École Normale Supérieure de Lyon, Post-Doctoral Fellow, until Jun 2019]
Changmin Lee [Univ de Lyon, Post-Doctoral Fellow]
Hervé Tale Kalachi [Inria, Post-Doctoral Fellow, from Apr 2019]
Anastasiia Volkova Lozanova [Inria, Post-Doctoral Fellow, until Feb 2019]

Administrative Assistants

Nelly Amsellem [École Normale Supérieure de Lyon, Administrative Assistant, until Sep 2019]
Virginie Bouyer [École Normale Supérieure de Lyon, Administrative Assistant, from Oct 2019]
Octavie Paris [École Normale Supérieure de Lyon, Administrative Assistant, from May 2019]

2. Overall Objectives

2.1. Overall Objectives

A major challenge in modeling and scientific computing is the simultaneous mastery of hardware capabilities, software design, and mathematical algorithms for the efficiency and reliability of the computation. In this context, the overall objective of AriC is to improve computing at large, in terms of performance, efficiency, and reliability. We work on the fine structure of floating-point arithmetic, on controlled approximation schemes, on algebraic algorithms and on new cryptographic applications, most of these themes being pursued in their interactions. Our approach combines fundamental studies, practical performance and qualitative aspects, with a shared strategy going from high-level problem specifications and standardization actions, to computer arithmetic and the lowest-level details of implementations.

This makes AriC the right place for drawing 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 cryptography 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 arithmetic 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 cryptography. 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.

- **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 cryptography (§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 and certified approximation methods

3.1.1. Safe numerical approximations

The last twenty years have seen the advent of computer-aided proofs in mathematics and this trend is getting more and more important. They request: fast and stable numerical computations; numerical results with a guarantee on the error; formal proofs of these computations or computations with a proof assistant. One of our main long-term objectives is to develop a platform where one can study a computational problem on all (or any) of these three levels of rigor. At this stage, most of the necessary routines are not easily available (or do not even exist) and one needs to develop *ad hoc* tools to complete the proof. We plan to provide more and more algorithms and routines to address such questions. Possible applications lie in the study of mathematical conjectures where exact mathematical results are required (e.g., stability of dynamical systems); or in more applied questions, such as the automatic generation of efficient and reliable numerical software for function evaluation. On a complementary viewpoint, numerical safety is also critical in robust space mission design, where guidance and control algorithms become more complex in the context of increased satellite autonomy. We will pursue our collaboration with specialists of that area whose questions bring us interesting focus on relevant issues.

3.1.2. Floating-point computing

Floating-point arithmetic is currently undergoing a major evolution, in particular with the recent advent of a greater diversity of available precisions on a same system (from 8 to 128 bits) and of coarser-grained floating-point hardware instructions. This new arithmetic landscape raises important issues at the various levels of computing, that we will address along the following three directions.

3.1.2.1. Floating-point algorithms, properties, and standardization

One of our targets is the design of building blocks of computing (e.g., algorithms for the basic operations and functions, and algorithms for complex or double-word arithmetic). Establishing properties of these building blocks (e.g., the absence of "spurious" underflows/overflows) is also important. The IEEE 754 standard on floating-point arithmetic (whose next version, a rather minor revision, will be released soon) will have to undergo a major revision within a few years: first because advances in technology or new needs make some of its features obsolete, and because new features need standardization. We aim at playing a leading role in the preparation of the next standard.

3.1.2.2. Error bounds

We will pursue our studies in rounding error analysis, in particular for the "low precision–high dimension" regime, where traditional analyses become ineffective and where improved bounds are thus most needed. For this, the structure of both the data and the errors themselves will have to be exploited. We will also investigate the impact of mixed-precision and coarser-grained instructions (such as small matrix products) on accuracy analyses.

3.1.2.3. High performance kernels

Most directions in the team are concerned with optimized and high performance implementations. We will pursue our efforts concerning the implementation of well optimized floating-point kernels, with an emphasis on numerical quality, and taking into account the current evolution in computer architectures (the increasing width of SIMD registers, and the availability of low precision formats). We will focus on computing kernels used within other axes in the team such as, for example, extended precision linear algebra routines within the FPLLL and HPLLL libraries.

3.2. Lattices: algorithms and cryptology

We intend to strengthen our assessment of the cryptographic relevance of problems over lattices, and to broaden our studies in two main (complementary) directions: hardness foundations and advanced functionalities.

3.2.1. Hardness foundations

Recent advances in cryptography have broaden the scope of encryption functionalities (e.g., encryption schemes allowing to compute over encrypted data or to delegate partial decryption keys). While simple variants (e.g., identity-based encryption) are already practical, the more advanced ones still lack efficiency. Towards reaching practicality, we plan to investigate simpler constructions of the fundamental building blocks (e.g., pseudorandom functions) involved in these advanced protocols. We aim at simplifying known constructions based on standard hardness assumptions, but also at identifying new sources of hardness from which simple constructions that are naturally suited for the aforementioned advanced applications could be obtained (e.g., constructions that minimize critical complexity measures such as the depth of evaluation). Understanding the core source of hardness of today's standard hard algorithmic problems is an interesting direction as it could lead to new hardness assumptions (e.g., tweaked version of standard ones) from which we could derive much more efficient constructions. Furthermore, it could open the way to completely different constructions of advanced primitives based on new hardness assumptions.

3.2.2. Cryptanalysis

Lattice-based cryptography has come much closer to maturity in the recent past. In particular, NIST has started a standardization process for post-quantum cryptography, and lattice-based proposals are numerous and competitive. This dramatically increases the need for cryptanalysis: Do the underlying hard problems suffer from structural weaknesses? Are some of the problems used easy to solve, e.g., asymptotically? Are the chosen concrete parameters meaningful for concrete cryptanalysis? In particular, how secure would they be if all the known algorithms and implementations thereof were pushed to their limits? How would these concrete performances change in case (full-fledged) quantum computers get built?

On another front, the cryptographic functionalities reachable under lattice hardness assumptions seem to get closer to an intrinsic ceiling. For instance, to obtain cryptographic multilinear maps, functional encryption and indistinguishability obfuscation, new assumptions have been introduced. They often have a lattice flavour, but are far from standard. Assessing the validity of these assumptions will be one of our priorities in the mid-term.

3.2.3. Advanced cryptographic primitives

In the design of cryptographic schemes, we will pursue our investigations on functional encryption. Despite recent advances, efficient solutions are only available for restricted function families. Indeed, solutions for general functions are either way too inefficient for pratical use or they rely on uncertain security foundations like the existence of circuit obfuscators (or both). We will explore constructions based on well-studied hardness assumptions and which are closer to being usable in real-life applications. In the case of specific functionalities, we will aim at more efficient realizations satisfying stronger security notions.

Another direction we will explore is multi-party computation via a new approach exploiting the rich structure of class groups of quadratic fields. We already showed that such groups have a positive impact in this field by designing new efficient encryption switching protocols from the additively homomorphic encryption we introduced earlier. We want to go deeper in this direction that raises interesting questions such as how to design efficient zero-knowledge proofs for groups of unknown order, how to exploit their structure in the context of 2-party cryptography (such as two-party signing) or how to extend to the multi-party setting.

In the context of the PROMETHEUS H2020 project, we will keep seeking to develop new quantum-resistant privacy-preserving cryptographic primitives (group signatures, anonymous credentials, e-cash systems, etc). This includes the design of more efficient zero-knowledge proof systems that can interact with lattice-based cryptographic primitives.

3.3. Algebraic computing and high performance kernels

The connections between algorithms for structured matrices and for polynomial matrices will continue to be developed, since they have proved to bring progress to fundamental questions with applications throughout computer algebra. The new fast algorithm for the bivariate resultant opens an exciting area of research which should produce improvements to a variety of questions related to polynomial elimination. Obviously, we expect to produce results in that area.

For definite summation and integration, we now have fast algorithms for single integrals of general functions and sequences and for multiple integrals of rational functions. The long-term objective of that part of computer algebra is an efficient and general algorithm for multiple definite integration and summation of general functions and sequences. This is the direction we will take, starting with single definite sums of general functions and sequences (leading in particular to a faster variant of Zeilberger's algorithm). We also plan to investigate geometric issues related to the presence of apparent singularities and how they seem to play a role in the complexity of the current 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

Florent Bréhard, jointly with Mioara Joldes and Jean-Bernard Lasserre (CNRS LAAS) received the Distinguished paper award at ISSAC 2019 for *On Moment Problems with Holonomic Functions*.

Alice Pellet-Mary was an awardee of the L'Oréal-Unesco scholarship for Women and Science.
BEST PAPERS AWARDS :
[16]
F. BRÉHARD, M. JOLDES, J.-B. LASSERRE.*On Moment Problems with Holonomic Functions*, in "ISSAC 2019 - 44th International Symposium on Symbolic and Algebraic Computation", Pékin, China, July 2019, p. 66-73, https://hal.archives-ouvertes.fr/hal-02006645

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. Exchange algorithm for evaluation and approximation error-optimized polynomials

Machine implementation of mathematical functions often relies on polynomial approximations. The particularity is that rounding errors occur both when representing the polynomial coefficients on a finite number of bits, and when evaluating it in finite precision. Hence, for finding the best polynomial (for a given fixed degree, norm and interval), one has to consider both types of errors: approximation and evaluation. While efficient algorithms were already developed for taking into account the approximation error, the evaluation part is usually a posteriori handled, in an ad-hoc manner. In [15], we formulate a semi-infinite linear optimization problem whose solution is the best polynomial with respect to the supremum norm of the sum of both errors. This problem is then solved with an iterative exchange algorithm, which can be seen as an extension of the well-known Remez algorithm. A discussion and comparison of the obtained results on different examples are finally presented.

7.1.2. On Moment Problems with Holonomic Functions

Many reconstruction algorithms from moments of algebraic data were developed in optimization, analysis or statistics. Lasserre and Putinar proposed an exact reconstruction algorithm for the algebraic support of the Lebesgue measure, or of measures with density equal to the exponential of a known polynomial. Their approach relies on linear recurrences for the moments, obtained using Stokes theorem. In [16], we extend this study to measures with holonomic densities and support with real algebraic boundary. In the framework of holonomic distributions (i.e. they satisfy a holonomic system of linear partial or ordinary differential equations with polynomial coefficients), an alternate method to creative telescoping is proposed for computing linear recurrences for the moments. When the coefficients of a polynomial vanishing on the support boundary are given as parameters, the obtained recurrences have the advantage of staying linear with respect to them. This property allows for an efficient reconstruction method. Given a finite number of numerically computed moments for a measure with holonomic density, and assuming a real algebraic boundary for the support, we propose an algorithm for solving the inverse problem of obtaining both the coefficients of a polynomial vanishing on the boundary and those of the polynomials involved in the holonomic operators which annihilate the density.

7.1.3. A certificate-based approach to formally verified approximations

In [17], we present a library to verify rigorous approximations of univariate functions on real numbers, with the Coq proof assistant. Based on interval arithmetic, this library also implements a technique of validation a posteriori based on the Banach fixed-point theorem. We illustrate this technique on the case of operations of division and square root. This library features a collection of abstract structures that organize the specification of rigorous approximations, and modularize the related proofs. Finally, we provide an implementation of verified Chebyshev approximations, and we discuss a few examples of computations.

7.2. Floating-point and validated numerics

7.2.1. Error analysis of some operations involved in the Cooley-Tukey Fast Fourier Transform

We are interested in [4] in obtaining error bounds for the classical Cooley-Tukey 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.

7.2.2. Algorithms for triple-word arithmetic

Triple-word arithmetic consists in representing high-precision numbers as the unevaluated sum of three floating-point numbers (with “nonoverlapping” constraints that are explicited in the paper). We introduce and analyze in [7] various algorithms for manipulating triple-word numbers: rounding a triple-word number to a floating-point number, adding, multiplying, dividing, and computing square-roots of triple-word numbers, etc. We compare our algorithms, implemented in the Campary library, with other solutions of comparable accuracy. It turns out that our new algorithms are significantly faster than what one would obtain by just using the usual floating-point expansion algorithms in the special case of expansions of length 3.

7.2.3. Accurate Complex Multiplication in Floating-Point Arithmetic

We deal in [24] with accurate complex multiplication in binary floating-point arithmetic, with an emphasis on the case where one of the operands in a “double-word” number. We provide an algorithm that returns a complex product with normwise relative error bound close to the best possible one, i.e., the rounding unit u.

7.2.4. Semi-automatic implementation of the complementary error function

The normal and complementary error functions are ubiquitous special functions for any mathematical library. They have a wide range of applications. Practical applications call for customized implementations that have strict accuracy requirements. Accurate numerical implementation of these functions is, however, non-trivial. In particular, the complementary error function erfc for large positive arguments heavily suffers from cancellation, which is largely due to its asymptotic behavior. We provide a semi-automatic code generator for the erfc function which is parameterized by the user-given bound on the relative error. Our solution, presented in [31], exploits the asymptotic expression of erfc and leverages the automatic code generator Metalibm that provides accurate polynomial approximations. A fine-grained a priori error analysis provides a libm developer with the required accuracy for each step of the evaluation. In critical parts, we exploit double-word arithmetic to achieve implementations that are fast, yet accurate up to 50 bits, even for large input arguments. We demonstrate that for high required accuracies the automatically generated code has performance comparable to that of the standard libm and for lower ones our code demonstrated roughly 25% speedup.

7.2.5. Posits: the good, the bad and the ugly

Many properties of the IEEE-754 floating-point number system are taken for granted in modern computers and are deeply embedded in compilers and low-level softare routines such as elementary functions or BLAS. In [32] we review such properties on the recently proposed Posit number system. Some are still true. Some are no longer true, but sensible work-arounds are possible, and even represent exciting challenge for the community. Some, in particular the loss of scale invariance for accuracy, are extremely dangerous if Posits are to replace floating point completely. This study helps framing where Posits are better than floating-point, where they are worse, and what tools are missing in the Posit landscape. For general-purpose computing, using Posits as a storage format only could be a way to reap their benefits without loosing those of classical floating-point. The hardware cost of this alternative is studied.

7.2.6. The relative accuracy of $(x+y)*(x-y)$

We consider in [8] the relative accuracy of evaluating $(x+y)(x-y)$ in IEEE floating-point arithmetic, when x and y are two floating-point numbers and rounding is to nearest. This expression can be used for example as an efficient cancellation-free alternative to x^2-y^2 and is well known to have low relative error, namely, at most about $3u$ with u denoting the unit roundoff. In this paper we complement this traditional analysis with a finer-grained one, aimed at improving and assessing the quality of that bound. Specifically, we show that if the tie-breaking rule is *to away* then the bound $3u$ is asymptotically optimal. In contrast, if the tie-breaking rule is *to even*, we show that asymptotically optimal bounds are now $2.25u$ for base two and $2u$ for larger bases, such as base ten. In each case, asymptotic optimality is obtained by the explicit construction of a certificate, that is, some floating-point input (x,y) parametrized by u and such that the error of the associated result is equivalent to the error bound as $u\to 0$. We conclude with comments on how $(x+y)(x-y)$ compares with x^2 in the presence of floating-point arithmetic, in particular showing cases where the computed value of $(x+y)(x-y)$ exceeds that of x^2.

7.2.7. The MPFI Library: Towards IEEE 1788-2015 Compliance

The IEEE 1788-2015 has standardized interval arithmetic. However, few libraries for interval arithmetic are compliant with this standard. In the first part of [30], the main features of the IEEE 1788-2015 standard are detailed. These features were not present in the libraries developed prior to the elaboration of the standard. MPFI is such a library: it is a C library, based on MPFR, for arbitrary precision interval arithmetic. MPFI is not (yet) compliant with the IEEE 1788-2015 standard for interval arithmetic: the planned modifications are presented.

7.3. Lattices: algorithms and cryptology

7.3.1. Approx-SVP in ideal lattices with pre-processing

In [28], we describe an algorithm to solve the approximate Shortest Vector Problem for lattices corresponding to ideals of the ring of integers of an arbitrary number field K. This algorithm has a pre-processing phase, whose run-time is exponential in $\log|\Delta|$ with Δ the discriminant of K. Importantly, this pre-processing phase depends only on K. The pre-processing phase outputs an advice, whose bit-size is no more than the run-time of the query phase. Given this advice, the query phase of the algorithm takes as input any ideal I of the ring of integers, and outputs an element of I which is at most $\exp(\widetilde{O}((\log|\Delta|)^{\alpha+1}/n))$ times longer than a shortest non-zero element of I (with respect to the Euclidean norm of its canonical embedding). This query phase runs in time and space $\exp(\widetilde{O}((\log|\Delta|)^{\max(2/3,1-2\alpha)}))$ in the classical setting, and $\exp(\widetilde{O}((\log|\Delta|)^{1-2\alpha}))$ in the quantum setting. The parameter α can be chosen arbitrarily in $[0,1/2]$. Both correctness and cost analyses rely on heuristic assumptions, whose validity is consistent with experiments.

The algorithm builds upon the algorithms from Cramer al. [EUROCRYPT 2016] and Cramer et al. [EUROCRYPT 2017]. It relies on the framework from Buchmann [Séminaire de théorie des nombres 1990], which allows to merge them and to extend their applicability from prime-power cyclotomic fields to all number fields. The cost improvements are obtained by allowing precomputations that depend on the field only.

7.3.2. An LLL algorithm for module lattices

The LLL algorithm takes as input a basis of a Euclidean lattice, and, within a polynomial number of operations, it outputs another basis of the same lattice but consisting of rather short vectors. In [23], we provide a generalization to R-modules contained in K^n for arbitrary number fields K and dimension n, with R denoting the ring of integers of K. Concretely, we introduce an algorithm that efficiently finds short vectors in rank-n modules when given access to an oracle that finds short vectors in rank-2 modules, and an algorithm that efficiently finds short vectors in rank-2 modules given access to a Closest Vector Problem oracle for a lattice that depends only on K. The second algorithm relies on quantum computations and its analysis is heuristic.

7.3.3. The general sieve kernel and new records in lattice reduction

In [14], we propose the General Sieve Kernel (G6K), an abstract stateful machine supporting a wide variety of lattice reduction strategies based on sieving algorithms. Using the basic instruction set of this abstract stateful machine, we first give concise formulations of previous sieving strategies from the literature and then propose new ones. We then also give a light variant of BKZ exploiting the features of our abstract stateful machine. This encapsulates several recent suggestions (Ducas at Eurocrypt 2018; Laarhoven and Mariano at PQCrypto 2018) to move beyond treating sieving as a blackbox SVP oracle and to utilise strong lattice reduction as preprocessing for sieving. Furthermore, we propose new tricks to minimise the sieving computation required for a given reduction quality with mechanisms such as recycling vectors between sieves, on-the-fly lifting and flexible insertions akin to Deep LLL and recent variants of Random Sampling Reduction.

Moreover, we provide a highly optimised, multi-threaded and tweakable implementation of this machine which we make open-source. We then illustrate the performance of this implementation of our sieving strategies by applying G6K to various lattice challenges. In particular, our approach allows us to solve previously unsolved instances of the Darmstadt SVP (151, 153, 155) and LWE (e.g. (75, 0.005)) challenges. Our solution for the SVP-151 challenge was found 400 times faster than the time reported for the SVP-150 challenge, the previous record. For exact SVP, we observe a performance crossover between G6K and FPLLL's state of the art implementation of enumeration at dimension 70.

7.3.4. Statistical zeroizing attack: cryptanalysis of candidates of BP obfuscation over GGH15 multilinear map

In [19], we present a new cryptanalytic algorithm on obfuscations based on GGH15 multilinear map. Our algorithm, statistical zeroizing attack, directly distinguishes two distributions from obfuscation while it follows the zeroizing attack paradigm, that is, it uses evaluations of zeros of obfuscated programs.

Our attack breaks the recent indistinguishability obfuscation candidate suggested by Chen et al. (CRYPTO'18) for the optimal parameter settings. More precisely, we show that there are two functionally equivalent branching programs whose CVW obfuscations can be efficiently distinguished by computing the sample variance of evaluations.

This statistical attack gives a new perspective on the security of the indistinguishability obfuscations: we should consider the shape of the distributions of evaluation of obfuscation to ensure security.

In other words, while most of the previous (weak) security proofs have been studied with respect to algebraic attack model or ideal model, our attack shows that this algebraic security is not enough to achieve indistinguishability obfuscation. In particular, we show that the obfuscation scheme suggested by Bartusek et al. (TCC'18) does not achieve the desired security in a certain parameter regime, in which their algebraic security proof still holds.

The correctness of statistical zeroizing attacks holds under a mild assumption on the preimage sampling algorithm with a lattice trapdoor. We experimentally verify this assumption for implemented obfuscation by Halevi et al. (ACM CCS'17).

7.3.5. Cryptanalysis of the CLT13 multilinear map

The reference [6] is the journal version of the Eurocrypt'15 article with the same title and authors.

7.3.6. Multi-Client Functional Encryption for Linear Functions in the Standard Model from LWE

Multi-client functional encryption (MCFE) allows ℓ clients to encrypt ciphertexts $\mathbf{C}_{t,1}, \mathbf{C}_{t,2}, ..., \mathbf{C}_{t,\ell}$ under some label. Each client can encrypt his own data X_i for a label t using a private encryption key ek_i issued by a trusted authority in such a way that, as long as all $\mathbf{C}_{t,i}$ share the same label t, an evaluator endowed with a functional key dk_f can evaluate $f(X_1, X_2, ..., X_\ell)$ without learning anything else on the underlying plaintexts X_i. Functional decryption keys can be derived by the central authority using the master secret key. Under the Decision Diffie-Hellman assumption, Chotard *et al.* (Asiacrypt 2018) recently described an adaptively secure MCFE scheme for the evaluation of linear functions over the integers. They also gave a decentralized variant (DMCFE) of their scheme which does not rely on a centralized authority, but rather allows encryptors to issue functional secret keys in a distributed manner. While efficient, their constructions both rely on random oracles in their security analysis. In [27], we build a standard-model MCFE scheme for the same functionality and prove it fully secure under adaptive corruptions. Our proof relies on the Learning-With-Errors (LWE) assumption and does not require the random oracle model. We also provide a decentralized variant of our scheme, which we prove secure in the static corruption setting (but for adaptively chosen messages) under the LWE assumption.

7.3.7. Zero-Knowledge Elementary Databases with More Expressive Queries

Zero-knowledge elementary databases (ZK-EDBs) are cryptographic schemes that allow a prover to commit to a set D of key-value pairs so as to be able to prove statements such as "x belongs to the support of D and D(x) = y" or "x is not in the support of D". Importantly , proofs should leak no information beyond the proven statement and even the size of D should remain private. Chase et al. (Eurocrypt'05) showed that ZK-EDBs are implied by a special flavor of non-interactive commitment, called mercurial commitment, which enables efficient instantiations based on standard number theoretic assumptions. On the other hand, the resulting ZK-EDBs are only known to support proofs for simple statements like (non-)membership and value assignments. In [25], we show that mercurial commitments actually enable significantly richer queries. We show that, modulo an additional security property met by all known efficient constructions, they actually enable range queries over keys and values-even for ranges of super-polynomial size-as well as membership/non-membership queries over the space of values. Beyond that, we exploit the range queries to realize richer queries such as k-nearest neighbors and revealing the k smallest or largest records within a given range. In addition, we provide a new realization of trapdoor mercurial commitment from standard lattice asssumptions, thus obtaining the most expressive quantum-safe ZK-EDB construction so far.

7.3.8. Lossy Algebraic Filters With Short Tags

Lossy algebraic filters (LAFs) are function families where each function is parametrized by a tag, which determines if the function is injective or lossy. While initially introduced by Hofheinz (Eurocrypt 2013) as a technical tool to build encryption schemes with key-dependent message chosen-ciphertext (KDM-CCA) security, they also find applications in the design of robustly reusable fuzzy extractors. So far, the only known LAF family requires tags comprised of $\Theta(n^2)$ group elements for functions with input space $\mathbb{Z}_p$, where p is the group order. In [26], we describe a new LAF family where the tag size is only linear in n and prove it secure under simple assumptions in asymmetric bilinear groups. Our construction can be used as a drop-in replacement in all applications of the initial LAF system. In particular, it can shorten the ciphertexts of Hofheinz's KDM-CCA-secure public-key encryption scheme by 19 group elements. It also allows substantial space improvements in a recent fuzzy extractor proposed by Wen and Liu (Asiacrypt 2018). As a second contribution , we show how to modify our scheme so as to prove it (almost) tightly secure, meaning that security reductions are not affected by a concrete security loss proportional to the number of adversarial queries.

7.3.9. Shorter Quadratic QA-NIZK Proofs

Despite recent advances in the area of pairing-friendly Non-Interactive Zero-Knowledge proofs, there have not been many efficiency improvements in constructing arguments of satisfiability of quadratic (and larger degree) equations since the publication of the Groth-Sahai proof system (J. of Cryptology 2012). In [20], we address the problem of aggregating such proofs using techniques derived from the interactive setting and recent constructions of SNARKs. For certain types of quadratic equations, this problem was investigated before by González et al. (Asiacrypt'15). Compared to their result, we reduce the proof size by approximately 50

7.3.10. Shorter Pairing-based Arguments under Standard Assumptions

The paper [22] constructs efficient non-interactive arguments for correct evaluation of arithmetic and Boolean circuits with proof size $O(d)$ group elements, where d is the multiplicative depth of the circuit, under falsifiable assumptions. This is achieved by combining techniques from SNARKs and QA-NIZK arguments of membership in linear spaces. The first construction is very efficient (the proof size is $\approx 4d$ group elements and the verification cost is $4d$ pairings and $O(n + n + d)$ exponentiations, where n is the size of the input and n of the output) but one type of attack can only be ruled out assuming the knowledge soundness of QA-NIZK arguments of membership in linear spaces. We give an alternative construction which replaces this assumption with a decisional assumption in bilinear groups at the cost of approximately doubling the proof size. The construction for Boolean circuits can be made zero-knowledge with Groth-Sahai proofs, resulting in a NIZK argument for circuit satisfiability based on falsifiable assumptions in bilinear groups of proof size $O(n + d)$. Our main technical tool is what we call an "argument of knowledge transfer". Given a commitment C_1 and an opening x, such an argument allows to prove that some other commitment C_2 opens to $f(x)$, for some function f, even if C_2 is not extractable. We construct very short, constant-size, pairing-based arguments of knowledge transfer with constant-time verification for any linear function and also for Hadamard products. These allow to transfer the knowledge of the input to lower levels of the circuit.

7.3.11. Shorter Ring Signatures from Standard Assumptions

Ring signatures, introduced by Rivest, Shamir and Tauman (ASIACRYPT 2001), allow to sign a message on behalf of a set of users while guaranteeing authenticity and anonymity. Groth and Kohlweiss (EUROCRYPT 2015) and Libert *et al.* (EUROCRYPT 2016) constructed schemes with signatures of size logarithmic in the number of users. An even shorter ring signature, of size independent from the number of users, was recently proposed by Malavolta and Schroeder (ASIACRYPT 2017). However, all these short signatures are obtained relying on strong and controversial assumptions. Namely, the former schemes are both proven secure in the random oracle model while the later requires non-falsifiable assumptions.

The most efficient construction under mild assumptions remains the construction of Chandran et al. (ICALP 2007) with a signature of size $\Theta(\sqrt{n})$, where n is the number of users, and security is based on the Diffie-Hellman assumption in bilinear groups (the SXDH assumption in asymmetric bilinear groups).

In [21], we construct an asymptotically shorter ring signature from the hardness of the Diffie-Hellman assumption in bilinear groups. Each signature comprises $\Theta(n^{1/3})$ group elements, signing a message requires computing $\Theta(n^{1/3})$ exponentiations, and verifying a signature requires $\Theta(n^{2/3})$ pairing operations.

7.3.12. Two-Party ECDSA from Hash Proof Systems and Efficient Instantiations

ECDSA is a widely adopted digital signature standard. Unfortunately, efficient distributed variants of this primitive are notoriously hard to achieve and known solutions often require expensive zero knowledge proofs to deal with malicious adversaries. For the two party case, Lindell (CRYPTO 2017) recently managed to get an efficient solution which, to achieve simulation-based security, relies on an interactive, non standard, assumption on Paillier's cryptosystem.

In this paper [18] we generalize Lindell's solution using hash proof systems. The main advantage of our generic method is that it results in a simulation-based security proof without resorting to non-standard interactive assumptions.

Moving to concrete constructions, we show how to instantiate our framework using class groups of imaginary quadratic fields. Our implementations show that the practical impact of dropping such interactive assumptions is minimal. Indeed, while for 128-bit security our scheme is marginally slower than Lindell's, for 256-bit security it turns out to be better both in key generation and signing time. Moreover, in terms of communication cost, our implementation significantly reduces both the number of rounds and the transmitted bits without exception.

7.3.13. Algebraic XOR-RKA-Secure Pseudorandom Functions from Post-Zeroizing Multilinear Maps

In [13], we construct the first pseudorandom functions that resist a strong class of attacks where an adversary is able to run the cryptosystem not only with the fixed secret key, but with related keys where bits of its choice of the original keys are flipped. This problem is motivated by practical attacks that have been performed against physical devices. Our construction guarantees that every output of our construction, for the original key or for tampered keys, are pseudorandom, i.e. are computationally hard to distinguish from truly random values. To achieve this, we rely on a recent tool introduced in cryptography and termed multilinear maps. While multilinear maps have been recently attacked by several techniques, we prove that our construction remains secure despite the numerous vulnerabilities of current constructions of multilinear maps.

7.3.14. Unifying Leakage Models on a Rényi Day

Most theoretical models in cryptography suppose that an attacker can only observe the input/output behavior of a cryptosystem and nothing more. Yet, in the real world, cryptosystems run on physical devices and auxiliary information leaks from these devices. This leakage can sometimes be used to attack the system, even though it is proven secure in theory. To circumvent these issues, cryptographers have introduces several new security models in an attempt to encompass the different forms of leakage. Some models are simple, such as the probing model, and simple compilers allow to transform a system into one secure in the probing model, while some more realistic problems such as the noisy-leakage model are very involved. In [29], we show that these models are actually equivalent, proving in particular that the simple compilers are sufficient to guarantee security in realistic environments.

7.4. Algebraic computing and high-performance kernels

7.4.1. 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].

7.4.2. Absolute root separation

The absolute separation of a polynomial is the minimum nonzero difference between the absolute values of its roots. In the case of polynomials with integer coefficients, it can be bounded from below in terms of the degree and the height (the maximum absolute value of the coefficients) of the polynomial. We improve the known bounds for this problem and related ones. Then we report on extensive experiments in low degrees, suggesting that the current bounds are still very pessimistic. [5]

7.4.3. Improving the complexity of block low-rank factorizations with fast matrix arithmetic

We consider in [9] the LU factorization of an $n \times n$ matrix represented as a block low-rank (BLR) matrix: most of its off-diagonal blocks are approximated by matrices of small rank r, which reduces the asymptotic complexity of computing the LU factorization down to $\mathcal{O}(n^2 r)$. Even though lower complexities can be achieved with hierarchical matrices, the BLR format allows for a very simple and efficient implementation. In this article, our aim is to further reduce the BLR complexity without losing its nonhierarchical nature by exploiting fast matrix arithmetic, that is, the ability to multiply two $n \times n$ full-rank matrices together for $\mathcal{O}(n^\omega)$ flops, where $\omega < 3$. We devise a new BLR factorization algorithm whose cost is $\mathcal{O}(n^{(\omega+1)/2} r^{(\omega-1)/2})$, which represents an asymptotic improvement compared with the standard BLR factorization as soon as $\omega < 3$. In particular, for Strassen's algorithm, $\omega \approx 2.81$ yields the cost $\mathcal{O}(n^{1.904} r^{0.904})$. Our numerical experiments are in good agreement with this analysis.

7.4.4. Fast computation of approximant bases in canonical form

In [10] 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.

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 FastRelax Project

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

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.

9.1.2. 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.

9.1.3. RISQ Project

Participants: Chitchanok Chuengsatiansup, Rikki Amit Inder Deo, Hervé Tale Kalachi, 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. PROMETHEUS Project

Participants: Fabien Laguillaumie, Benoît Libert, Octavie Paris, 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.

9.3. International Initiatives

9.3.1. Participation in Other International Programs

9.3.1.1. 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.

9.3.1.2. Inria International Chairs

- **TUCKER Warwick**
- Department of Mathematics - Uppsala University - Sweden
- Title: Attracteur de Hénon et intégrales abéliennes liées aux 16e problème de Hilbert
- 2018 – 2022

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Ron Steinfeld, Monash University (June)
- Amin Sakzad, Monash University (June)
- Shi Bai, Florida Atlantic University (June and July)
- David Wu, University of Virginia (July)
- Olivier Bernard, Université Rennes 1 and Thalès (October and November)
- Gautier Eberhart, Université Rennes 1 (October and November)
- Federico Savasta, Università degli Studi di Catania (October)

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

- Damien Stehlé organized a Winter School on the mathematical foundations of public-key cryptography, in Aussois (France), from March 17 to March 22.
- Bruno Salvy was a co-chair of AofA'2019 (Analysis of Algorithms), in Luminy, France.

10.1.2. Scientific Events: Selection

- Elena Kirshanova was in the program committee of Asiacrypt 2019.

- Benoît Libert was in the program committees of PKC 2019 and PKC 2020.
- Nathalie Revol is in the steering committee of the Arith conference series. She was in the program committee of Arith'26, of Correctness 2019 (workshop of SuperComputing) and of PPAM (Parallel Processing and Applied Mathematics) 2019.
- Bruno Salvy is in the steering committee of AofA. He was in the program committee for FPSAC 2019 and in the scientific committee of OPSFA 2019. He is in the program committee for AofA 2020.
- Damien Stehlé is in the steering committee of the PQCrypto conference series. He was in the program committee PQCrypto 2019 and is in the program committee of PQCrypto 2020. He was in the program committee of CRYPTO 2019.
- Gilles Villard was in the program committee of ISSAC 2019.

10.1.3. Journal

- Benoît Libert was a member of the editorial board of IET Information Security until July 31, 2019.
- Jean-Michel Muller is associate editor in chief of the IEEE Transactions on Emerging Topics in Computing.
- Nathalie Revol is a member of the editorial board of Reliable Computing.
- 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 of the journal *Annals of Combinatorics*.
- Damien Stehlé is a member of the editorial board of Journal of Cryptology.

10.1.4. Invited Talks

- Claude-Pierre Jeannerod gave an invited talk at the workshop *Structured Matrix Days* (Limoges, May 23–24, 2019).
- Benoît Libert gave an invited presentation during the "Workshop on Modern Trends in Cryptography" organized at Nanyang Technological University (Singapore) on June 13-14, 2019.
- Damien Stehlé gave lectures during the "Euclidean lattices: theory and applications" Summer school that was held in Kaliningrad (Russia), from July 15 to July 19.

10.1.5. Leadership within the Scientific Community

- Guillaume Hanrot was a member of selection committees for professors at Université de Lorraine (in CS) and at Université de Nouvelle-Calédonie (in Mathematics). He was also in the hiring committee of the computer science department of École polytechnique. He is a member of the working group on the revision of the CS curriculum in *Classes préparatoires aux grandes écoles*.
- Claude-Pierre Jeannerod is a member of the scientific committee of JNCF (Journées Nationales de Calcul Formel). He is also a member of the recruitment committee for postdocs and sabbaticals at Inria Grenoble–Rhône-Alpes.
- Jean-Michel Muller is co-head of the GDR Informatique Mathématique of the CNRS. He is also a member of the Scientific Council of CERFACS (Toulouse).
- Alain Passelègue is a member of the steering committee of the *Groupe de Travail Codage et Cryptographie* (GT-C2) of the GDR-IM.
- Nathalie Revol is a member of the steering committee of GDR Calcul; she was a member of the hiring committee (Comité de Sélection) for 2 positions at U. Nantes.
- Bruno Salvy is a member of the scientific councils of the CIRM, Luminy and of the GDR Informatique Mathématique of the CNRS.

10.1.6. Scientific Expertise

- Nathalie Revol has been an expert for the European H2020 program.

10.1.7. Research Administration

- Jean-Michel Muller is a member of the *Commission Administrative Paritaire* (CAP) *Directeurs de Recherches* of CNRS.
- 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: Guillaume Hanrot, Computer algebra, 10h, ENS de Lyon, France
- Master: Guillaume Hanrot, Cryptanalysis, 15h, ENS de Lyon, France
- Master: Claude-Pierre Jeannerod, Computer Algebra, 18h, M2Pro ISFA (Institut de Science Financière et d'Assurances), Université Claude Bernard Lyon 1, France
- Master (1&2): Fabien Laguillaumie, Cryptography, 160 h, ISFA, UCBL, France
- Master: Benoît Libert, Advanced Topics in Cryptography, 15h, ENS de Lyon, France
- Master: Nicolas Louvet, Compilers, 22h, M1, UCB Lyon 1, France
- Master: Alain Passelègue, Computer Algebra, 10h, M1, ENS de Lyon, France
- Master: Alain Passelègue, Advanced Topics in Cryptography, 30h, M2, ENS de Lyon, France
- Master: Nathalie Revol, Numerical Algorithms and Reliability, 12h, M2Pro ISFA (Institut de Science Financière et d'Assurances), Université Claude Bernard Lyon 1, France
- Master: Bruno Salvy, Computer Algebra, 6h, ENS de Lyon, France
- Master: Bruno Salvy, Logic and Complexity, 32h, École polytechnique, France
- Master: Damien Stehlé, Cryptanalysis, 15h, ENS de Lyon, France
- Master : Gilles Villard, Computer Algebra, 8h, ENS de Lyon, France
- Bachelor: Guillaume Hanrot, Calculability and complexity, 32h, ENS de Lyon, France
- Bachelor: Nicolas Louvet, Computer Architecture, 27h, L1, UCB Lyon 1, France
- Bachelor: Nicolas Louvet, Operating Systems, 50h, L2, UCB Lyon 1, France
- Bachelor: Nicolas Louvet, Data Structures and Algorithms, 24h, L2, UCB Lyon 1, France
- Bachelor: Nicolas Louvet, Data Structures and Algorithms, 40h, L3, UCB Lyon 1, France
- Bachelor: Nicolas Louvet, Formal Languages, 15h, L3, UCB Lyon 1, France
- Bachelor: Nicolas Louvet, Classical Logic, 15h, L3, UCB Lyon 1, France
- Bachelor: Bruno Salvy, Design and Analysis of Algorithms, 20h, École polytechnique, France

10.2.2. Supervision

- PhD: Florent Bréhard, Certified Numerics in Function Spaces: Polynomial Approximations Meet Computer Algebra and Formal Proof, ENS de Lyon, July 12, Nicolas Brisebarre (co-supervision with Mioara Joldes, CNRS LAAS, and Damien Pous, LIP)
- PhD: Alice Pellet-Mary, On ideal lattices and the GGH13 multilinear map, ENS de Lyon, October 16, Damien Stehlé
- PhD: Chen Qian, Lossy Trapdoor Primitives, Zero-Knowledge Proofs and Applications, IRISA Rennes, October 4, Benoît Libert (co-supervision with Pierre-Alain Fouque, IRISA)
- PhD in progress: Miruna Rosca, The middle-product learning with errors problem, January 2017, Damien Stehlé

- PhD in progress: Huyen Nguyen, Cryptographic aspects of orthogonal lattices, September 2018, Damien Stehlé
- PhD in progress: Radu Titiu, Advanced cryptographic primitives based on standard assumptions, January 2017, Benoît Libert
- PhD in progress: Adel Hamdi, Functional Encryption, December 2017, Fabien Laguillaumie (co-supervised by Sébastien Canard, Orange)
- PhD in progress: Ida Tucker, Advanced cryptographic primitives from homomorphic encryption, October 2017, Fabien Laguillaumie (co-direction with Guilhem Castagnos, Université de Bordeaux)

10.2.3. Juries

- Damien Stehlé was a jury member for the PhD defences of Ilia Iliashenko (K.U. Leuven, Belgium) and Joost Rijneveld (Radboud U., the Netherlands) and for the habilitation defence of Omar Fawzi (ENS de Lyon). He was a reviewer and jury member of the PhD of Thomas Debris-Alazard (Sorbonne U.) and for the habilitation of Guilhem Castagnos (U. Bordeaux).
- Benoît Libert was a reviewer for the PhD theses of Romain Gay (ENS Paris), Jérémy Chotard (Univ. of Limoges). He was a PhD examiner for the thesis of Andrea Cerulli (University College London, United Kingdom). He also chaired the PhD committee of Anca Nitulescu (ENS Paris).
- Fabien Laguillaumie was reviewer and jury member of the PhD of Pauline Bert (Université de Rennes) and of the HDR of Olivier Blazy (Université de Limoges)
- Jean-Michel Muller was reviewer and jury member of the PhD of Clothilde Jeangoudoux (Sorbonne University, Paris)
- Gilles Villard was reviewer for the PhD thesis of Robin Larrieu (Université Paris-Saclay); examiner for the habilitation of Pascal Giorgi (Université de Montpellier) and the PhD thesis of Matías Bender (Sorbonne Université).

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- Nathalie Revol was in the scientific committee for the Journées Scientifiques Inria (Lyon, 5-7 June 2019).
- Nathalie Revol was a member of the editorial committee of interstices; she is the scientific editor of this Web magazine since September 2019.
- Nathalie Revol belonged to the steering committee of MMI (Maison des Mathématiques et de l'Informatique) until July 2019; she is now a member of its prospective committee.
- 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, launched in February 2019.

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.
- Nathalie Revol works with DANE (Délégation Académique au Numérique dans l'Éducation) of Rectorat de Lyon towards educating primary school teachers, by educating educators (e-RUN); she is a member of the Conseil Scientifique du Numérique.
- Nathalie Revol presented activities for teaching computer science at every school level, and in particular activities led by Inria, for a Taiwanese delegation (Grenoble, October 2019).

10.3.4. Interventions

- Nathalie Revol spent 2 days at école Guilloux, with 50 pupils aged 10 (level: CM2), to teach computer science without any computer (so-called "unplugged computer science"): data, algorithms, networks.
- For high-school pupils ($\simeq$ 150 pupils): as an incentive, especially for girls, to choose scientific careers, Nathalie Revol gave talks at Mondial des Métiers (in February 2019), collège Jean Zay (Brignais, March 2019) and Girls Can Code (Lyon, August 2019). With Jérôme Germoni and Natacha Portier, she organized a day *Filles & Info* in November 2019, gathering about 70 high-school girls of 1e and Terminale.
- For a larger audience: she took part to Pop Science, doing magic tricks in the street at La Duchère, Lyon, May 2019; with Florent Masséglia, she introduced 5 important figures of computer science, chosen among the personalities of the "7 families of computer science" playcard, for L'Esprit Sorcier, for la Fête de la Science, Paris, October 2019, cf https://www.youtube.com/watch?v=ypnQe91Pztc
- Nathalie Revol took part to a day of teaching unplugged computer science for whoever was interested, at La Gaîté Lyrique, Paris, June 2019.

10.3.5. Creation of media or tools for science outreach

Nathalie Revol belongs to the working group "Informatique Sans Ordinateur", which creates unplugged activities to teach computer science; this group meets twice a year, usually in Lyon.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] F. BRE´HARD.*Calcul numérique certifié dans les espaces fonctionnels : Un trilogue entre approximations polynomiales rigoureuses, calcul symbolique et preuve formelle*, Université de Lyon, July 2019, https://tel.archives-ouvertes.fr/tel-02337901

[2] A. PELLET-MARY.*On ideal lattices and the GGH13 multilinear map*, Université de Lyon, October 2019, https://tel.archives-ouvertes.fr/tel-02337930

[3] C. QIAN.*Lossy trapdoor primitives, zero-knowledge proofs and applications*, Université de Rennes 1, October 2019, https://hal.archives-ouvertes.fr/tel-02415243

Articles in International Peer-Reviewed Journal

[4] N. BRISEBARRE, M. JOLDES, J.-M. MULLER, A.-M. NANEŞ, J. PICOT.*Error analysis of some operations involved in the Cooley-Tukey Fast Fourier Transform*, in "ACM Transactions on Mathematical Software", 2019, p. 1-34, forthcoming [*DOI :* 10.1145/NNNNNNN.NNNNNNN], https://hal.archives-ouvertes.fr/hal-01949458

[5] Y. BUGEAUD, A. DUJELLA, W. FANG, T. PEJKOVIĆ, B. SALVY.*Absolute root separation*, in "Experimental Mathematics", 2019, 8, https://arxiv.org/abs/1907.01232 , forthcoming [*DOI :* 10.1080/10586458.2019.1699480], https://hal.archives-ouvertes.fr/hal-02185594

[6] J. H. CHEON, K. HAN, C. LEE, H. RYU, D. STEHLÉ.*Cryptanalysis of the CLT13 Multilinear Map*, in "Journal of Cryptology", April 2019, vol. 32, n^{o} 2, p. 547-565 [*DOI :* 10.1007/S00145-018-9307-Y], https://hal.archives-ouvertes.fr/hal-02397396

[7] N. FABIANO, J.-M. MULLER, J. PICOT.*Algorithms for triple-word arithmetic*, in "IEEE Transactions on Computers", November 2019, vol. 68, n^{o} 11, p. 1573-1583 [*DOI :* 10.1109/TC.2019.2918451], https://hal.archives-ouvertes.fr/hal-01869009

[8] C.-P. JEANNEROD.*The relative accuracy of* $(x + y) * (x - y)$, in "Journal of Computational and Applied Mathematics", 2019, p. 1-17, forthcoming, https://hal.inria.fr/hal-02100500

[9] C.-P. JEANNEROD, T. MARY, C. PERNET, D. S. ROCHE.*Improving the Complexity of Block Low-Rank Factorizations with Fast Matrix Arithmetic*, in "SIAM Journal on Matrix Analysis and Applications", November 2019, vol. 40, n^{o} 4, p. 1478-1496 [*DOI :* 10.1137/19M1255628], https://hal.inria.fr/hal-02008666

[10] C.-P. JEANNEROD, V. NEIGER, G. VILLARD.*Fast computation of approximant bases in canonical form*, in "Journal of Symbolic Computation", 2019, p. 1-33, forthcoming [*DOI :* 10.1016/J.JSC.2019.07.011], https://hal-unilim.archives-ouvertes.fr/hal-01683632

[11] B. SALVY.*Linear Differential Equations as a Data-Structure*, in "Foundations of Computational Mathematics", October 2019, vol. 19, n^{o} 5, p. 1071-1112, https://arxiv.org/abs/1811.08616 - Based on an invited talk at FoCM'2017 [*DOI :* 10.1007/S10208-018-09411-X], https://hal.inria.fr/hal-01940078

[12] A. VOLKOVA, T. HILAIRE, C. LAUTER.*Arithmetic approaches for rigorous design of reliable Fixed-Point LTI filters*, in "IEEE Transactions on Computers", 2019, p. 1-14, forthcoming [*DOI :* 10.1109/TC.2019.2950658], https://hal.archives-ouvertes.fr/hal-01918650

International Conferences with Proceedings

[13] M. ABDALLA, F. BENHAMOUDA, A. PASSELÈGUE.*Algebraic XOR-RKA-Secure Pseudorandom Functions from Post-Zeroizing Multilinear Maps*, in "Advances in Cryptology – ASIACRYPT 2019", Kobe, Japan, S. D. GALBRAITH, S. MORIAI (editors), Lecture Notes in Computer Science, November 2019, vol. 11922, p. 386-412 [*DOI :* 10.1007/978-3-030-34621-8_14], https://hal.inria.fr/hal-02375594

[14] M. ALBRECHT, L. DUCAS, G. HEROLD, E. KIRSHANOVA, E. POSTLETHWAITE, M. STEVENS.*The General Sieve Kernel and New Records in Lattice Reduction*, in "EUROCRYPT", Darmstadt, Germany, April 2019, p. 717-746 [*DOI :* 10.1007/978-3-030-17656-3_25], https://hal.archives-ouvertes.fr/hal-02397424

[15] D. ARZELIER, F. BRÉHARD, M. JOLDES.*Exchange algorithm for evaluation and approximation error-optimized polynomials*, in "ARITH 2019 - 26th IEEE Symposium on Computer Arithmetic", Kyoto, Japan, IEEE, June 2019, p. 1-8, https://hal.archives-ouvertes.fr/hal-02006606

[16] *Best Paper*
F. BRÉHARD, M. JOLDES, J.-B. LASSERRE.*On Moment Problems with Holonomic Functions*, in "ISSAC 2019 - 44th International Symposium on Symbolic and Algebraic Computation", Pékin, China, July 2019, p. 66-73, https://hal.archives-ouvertes.fr/hal-02006645.

[17] F. BRÉHARD, A. MAHBOUBI, D. POUS.*A certificate-based approach to formally verified approximations*, in "ITP 2019 - Tenth International Conference on Interactive Theorem Proving", Portland, United States, 2019, p. 1-19 [*DOI :* 10.4230/LIPICS.ITP.2019.8], https://hal.laas.fr/hal-02088529

[18] G. CASTAGNOS, D. CATALANO, F. LAGUILLAUMIE, F. SAVASTA, I. TUCKER.*Two-Party ECDSA from Hash Proof Systems and Efficient Instantiations*, in "CRYPTO 2019 - 39th Annual International Cryptology Conference", Santa Barbara, United States, Advances in Cryptology – CRYPTO 2019, August 2019, vol. LNCS, n^{o} 11694, p. 191-221 [*DOI :* 10.1007/978-3-030-26954-8_7], https://hal.archives-ouvertes.fr/hal-02281931

[19] J. H. CHEON, W. CHO, M. HHAN, J. KIM, C. LEE.*Statistical Zeroizing Attack: Cryptanalysis of Candidates of BP Obfuscation over GGH15 Multilinear Map*, in "CRYPTO", Santa Barbara, United States, August 2019, p. 253-283 [*DOI :* 10.1007/978-3-030-26954-8_9], https://hal.archives-ouvertes.fr/hal-02397408

[20] V. DAZA, A. GONZÁLEZ, Z. PINDADO, C. RÀFOLS, J. SILVA.*Shorter Quadratic QA-NIZK Proofs*, in "PKC 2019 - 22nd IACR International Conference on Practice and Theory of Public-Key Cryptography", Beijing, China, LNCS, Springer, April 2019, vol. 11442, p. 314-343 [*DOI :* 10.1007/978-3-030-17253-4_11], https://hal.inria.fr/hal-02399179

[21] A. GONZÁLEZ.*Shorter Ring Signatures from Standard Assumptions*, in "PKC 2019 - 22nd IACR International Conference on Practice and Theory of Public-Key Cryptography", Beijing, China, Springer, April 2019, p. 99-126 [*DOI :* 10.1007/978-3-030-17253-4_4], https://hal.inria.fr/hal-02399172

[22] A. GONZÁLEZ, C. RÀFOLS.*Shorter Pairing-based Arguments under Standard Assumptions*, in "ASIACRYPT 2019 - 25th International Conference on the Theory and Application of Cryptology and Information Security", Kobe, Japan, LNCS, Springer, November 2019, vol. 11923, p. 728-757 [*DOI :* 10.1007/978-3-030-34618-8_25], https://hal.inria.fr/hal-02401556

[23] C. LEE, A. PELLET-MARY, D. STEHLÉ, A. WALLET.*An LLL Algorithm for Module Lattices*, in "ASIACRYPT", Kobe, Japan, November 2019, p. 59-90 [*DOI :* 10.1007/978-3-030-34621-8_3], https://hal.archives-ouvertes.fr/hal-02397335

[24] V. LEFÈVRE, J.-M. MULLER.*Accurate Complex Multiplication in Floating-Point Arithmetic*, in "ARITH 2019 - 26th IEEE Symposium on Computer Arithmetic", Kyoto, Japan, ARITH 2019 - 26th IEEE Symposium on Computer Arithmetic, IEEE, June 2019, p. 1-7, https://hal.archives-ouvertes.fr/hal-02001080

[25] B. LIBERT, K. NGUYEN, B. H. M. TAN, H. WANG.*Zero-Knowledge Elementary Databases with More Expressive Queries*, in "PKC 2019 - 22nd IACR International Conference on Practice and Theory of Public-Key Cryptography", Beijing, China, LNCS, Springer, April 2019, vol. 11442, p. 255-285 [*DOI :* 10.1007/978-3-030-17253-4_9], https://hal.inria.fr/hal-02151645

[26] B. LIBERT, C. QIAN.*Lossy Algebraic Filters With Short Tags*, in "PKC 2019 - 22nd International Conference on Practice and Theory of Public Key Cryptography", Beijing, China, LNCS, Springer, April 2019, vol. 11442, p. 34–65 [*DOI :* 10.1007/978-3-030-17253-4_2], https://hal.inria.fr/hal-02124968

[27] B. LIBERT, R. TITIU.*Multi-Client Functional Encryption for Linear Functions in the Standard Model from LWE*, in "ASIACRYPT 2019 - 25th International Conference on the Theory and Application of Cryptology and Information Security", Kobe, Japan, December 2019, p. 1-54 [*DOI :* 10.1007/978-3-030-34618-8_18], https://hal.inria.fr/hal-02352139

[28] A. PELLET-MARY, G. HANROT, D. STEHLÉ.*Approx-SVP in Ideal Lattices with Pre-processing*, in "Eurocrypt 2019 - 38th Annual International Conference on the Theory and Applications of Cryptographic Techniques", Darmstadt, Germany, Advances in Cryptology – EUROCRYPT 2019. Lecture Notes in Computer Science,

vol 11477, April 2019, p. 685-716 [*DOI :* 10.1007/978-3-030-17656-3_24], https://hal.archives-ouvertes.fr/hal-02139939

[29] T. PREST, D. GOUDARZI, A. MARTINELLI, A. PASSELÈGUE.*Unifying Leakage Models on a Rényi Day*, in "Crypto 2019 - 39th Annual International Cryptology Conference", Santa Barbara, CA, United States, August 2019, p. 683-712 [*DOI :* 10.1007/978-3-030-26948-7_24], https://hal.inria.fr/hal-02417544

[30] N. REVOL.*The MPFI Library: Towards IEEE 1788-2015 Compliance*, in "PPAM 2019 - 13th International Conference on Parallel Processing and Applied Mathematics", Bialystok, Poland, September 2019, p. 1-10, https://hal.inria.fr/hal-02162346

[31] A. VOLKOVA, J.-M. MULLER.*Semi-automatic implementation of the complementary error function*, in "ARITH 2019 - 26th IEEE Symposium on Computer Arithmetic", Kyoto, Japan, ARITH 2019 - 26th IEEE Symposium on Computer Arithmetic, IEEE, June 2019, p. 1-8, https://hal.archives-ouvertes.fr/hal-02002315

[32] F. DE DINECHIN, L. FORGET, J.-M. MULLER, Y. UGUEN.*Posits: the good, the bad and the ugly*, in "CoNGA 2019 - Conference on Next-Generation Arithmetic", Singapore, Singapore, ACM Press, March 2019, p. 1-10 [*DOI :* 10.1145/3316279.3316285], https://hal.inria.fr/hal-01959581

Software

[33] T. GIVARO GROUP.*Givaro*, June 2019, Version : 4.1.1
[SWH-ID : swh:1:dir:df65912bd1e5ea4b96b935de95f6638eb6d9472d], Software, https://hal.archives-ouvertes.fr/hal-02130729

[34] T. LINBOX GROUP.*LinBox*, June 2019, Version : 1.6.3
[SWH-ID : swh:1:dir:393b611a1424f032e83569bf6762502371cfcf65], Software, https://hal.archives-ouvertes.fr/hal-02130801

Other Publications

[35] S. BOLDO, C. Q. LAUTER, J.-M. MULLER.*Emulating round-to-nearest-ties-to-zero "augmented" floating-point operations using round-to-nearest-ties-to-even arithmetic*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02137968

[36] A. BOSTAN, T. RIVOAL, B. SALVY.*Explicit degree bounds for right factors of linear differential operators*, July 2019, https://arxiv.org/abs/1906.05529 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02154679

[37] F. BRÉHARD, N. BRISEBARRE, M. JOLDES, W. TUCKER.*A New Lower Bound on the Hilbert Number for Quartic Systems*, March 2019, working paper or preprint, https://hal.laas.fr/hal-02085895

[38] S. MELCZER, B. SALVY.*Effective Coefficient Asymptotics of Multivariate Rational Functions via Semi-Numerical Algorithms for Polynomial Systems*, July 2019, https://arxiv.org/abs/1905.04187 - 47 pages, https://hal.archives-ouvertes.fr/hal-02185586

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
A1.1.2. - Hardware accelerators (GPGPU, FPGA, etc.)
A1.1.4. - High performance computing
A1.1.5. - Exascale
A1.1.13. - Virtualization
A1.3.2. - Mobile distributed systems
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. - Infrastructure software
A2.6.1. - Operating systems
A2.6.2. - Middleware
A2.6.3. - Virtual machines
A2.6.4. - Ressource management
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]
Thierry Gautier [Inria, Researcher]
Laurent Lefèvre [Inria, Researcher, HDR]
Frédéric Suter [CNRS, Senior Researcher, HDR]
Marcos Dias de Assunção [Inria, Starting Research Position]
Farah Ait Salaht [Inria, Starting Research Position]

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]
Alain Tchana [École Normale Supérieure de Lyon, Professor, from Sep 2019, HDR]

Technical Staff

Simon Delamare [CNRS, Engineer]
Marie Durand [Inria, Engineer]
Matthieu Imbert [Inria, Engineer]
David Loup [Inria, Engineer]
Jean-Christophe Mignot [CNRS, Engineer]
Dominique Ponsard [CNRS, Engineer]
Cyril Seguin [Inria, Engineer, from Jul 2019]
Philippe Virouleau [Inria, Engineer, from May 2019 to Sept 2019]

PhD Students

Aurélie Kong-Win-Chang [Univ Claude Bernard, PhD Student]
Dorra Boughzala [Inria, PhD Student, until Jul 2019]
Vo Quoc Bao Bui [INPT, PhD Student, from Sep 2019]
Arthur Chevalier [Orange Labs, PhD Student]
Alexandre Da Silva Veith [Univ de Lyon, PhD Student, until Oct 2019]
Jad Darrous [Inria, PhD Student, until Sep 2019]
Felipe Rodrigo de Souza [École Normale Supérieure de Lyon, PhD Student]
Zeina Houmani [École Normale Supérieure de Lyon, PhD Student]
Barbe Thystere Mvondo Djob [Univ Grenoble Alpes, PhD Student, from Sep 2019]
Celestine Stella Ndonga Bitchebe [CNRS, PhD Student, from Sep 2019]
Laurent Turpin [Inria, PhD Student, from Oct 2019]
Patrick Lavoisier Wapet [INPT, PhD Student, from Sep 2019]
Maverick Chardet [Inria, PhD Student]
Idriss Daoudi [Inria, PhD Student]

Post-Doctoral Fellow

Cyril Seguin [Inria, Post-Doctoral Fellow, until Mar 2019, and MILYON, Post-Doctoral Felloc, from May to June]

Visiting Scientists

Jean Philippe Aboumou [SAHAM Life Insurance, from Oct 2019]

Carlos Henrique Cardonha [IBM Research Brazil, from Jun 2019 until Jul 2019]

Administrative Assistants

Evelyne Blesle [Inria, Administrative Assistant]
Solene Audoux [Inria, Administrative Assistant, until Apr 2019]
Virginie Bouyer [ENS Lyon, Administrative Assistant, from Apr 2019 until Aug 2019]

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, Leco, and Silecs 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 some approaches [24] 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 [27], phase detection for specific HPC applications [31], *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 [26] exposes the number of nodes but hides the complexity of network topology behind a set of collective operations; OpenMP [30] simplifies the management of threads on top of a shared memory machine while OpenACC [29] 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 [25]. 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. Parallel languages (UPC, Fortress, X10, *etc.*) is a first piece of the solution. However, they 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 [32] 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. Software Asset Management

The use of software is generally regulated by licenses, whether they are free or paid and with or without access to their sources. The world of licenses is very vast and unknown (especially in the industrial world). Often only the general public version is known (a software purchase corresponds to a license). For enterprises, the reality is much more complex, especially for main publishers. We work on the OpTISAM software, a software offering tools to perform Software Asset Management (SAM) much more efficiently in order to be able to ensure the full compliance with all contracts from each software and a new type of deployment taking into account these aspects and other additional parameters like energy and performance. This work is built on an Orange™ collaboration.

3.4.4. Cloud deployment and reproducibility

As part of the scientific method, any researcher should be able to reproduce the experimentation in order to not only verify the result but also evaluate and compare this experimentation with other approaches. The need of a standard tool allowing researchers to easily generate, share and reproduce experiments set-up arises. In our research, through a Nokia collaboration, we created SeeDep [10], a framework aiming at being such a standard tool. By associating a generation key to a network experiment set-up, SeeDep allows for reproducing network experiments independently from the used infrastructure.

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 that we are targeting are presented. In addition to highlighting some application needs, they also constitute some of the use cases that will used 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 [28] for ocean modelization), code-coupling applications (*e.g.,* the OASIS coupler [33] 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 can 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 Vera C. Rubin Observatory (https://www.vro.org/) will produce 20 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 des deux Infinis de Lyon* (IP2I) 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 VRO. 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. New Software and Platforms

5.1. 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: - Native Google Drive Support for the data manager - Standardization of internal integer types. - New types (see Changelog for more information)

NEWS OF THE YEAR: New DIET release (DIET 2.11) is avaible since may 2019 Batch's GENCI support to use the GENCI resources

- 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, Yves Caniou and Cyril Seguin
- Partners: CNRS - ENS Lyon - UCBL Lyon 1 - Sysfera
- Contact: Eddy Caron
- URL: http://graal.ens-lyon.fr/diet/

5.2. 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 2019: Python bindings were introduced, SMPI now partially supports some of the MPI/IO functions, a new model for Wifi networks was proposed, and the API for the simulation of storage resources was completely revisited. We also pursued our efforts to improve the documentation of the software, simplified the web site, and made a lot of bug fixing and code refactoring.

- 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/

5.3. 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

5.4. 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

5.5. XKBLAS

KEYWORDS: BLAS - Dense linear algebra - GPU

FUNCTIONAL DESCRIPTION: XKBLAS is yet an other BLAS library (Basic Linear Algebra Subroutines) that targets multi-GPUs architecture thanks to the XKaapi runtime and with block algorithms from PLASMA library. The library offers a wrapper library able to capture calls to BLAS (C or Fortran). The internal API is based on asynchronous invocations in order to enable overlapping between communication by computation and also to better composed sequences of calls to BLAS.

This current version of XKBlas is the first public version and contains only BLAS level 3 algorithms, including XGEMMT:

XGEMM XGEMMT: see MKL GEMMT interface XTRSM XTRMM XSYMM XSYRK XSYR2K XHEMM XHERK XHER2K

For classical precision Z, C, D, S.

RELEASE FUNCTIONAL DESCRIPTION: XKBlas has following limitations:

0.1 versions: calls to BLAS kernels must be initiate by the same thread that initializes the XKBlas library.

- Participants: Thierry Gautier and João Vicente Ferreira Lima
- Contact: Thierry Gautier
- URL: https://gitlab.inria.fr/xkblas/versions

5.6. Concerto

KEYWORDS: Reconfiguration - Distributed Software - Component models - Dynamic software architecture

FUNCTIONAL DESCRIPTION: Concerto is an implementation of the formal model Concerto written in Python. Concerto aloows to : 1. describe the life-cycle and the dependencies of software components, 2. describe a components assembly that forms the overall life-cycle of a distributed software, 3. automatically reconfigure a Concerto assembly of components by using a set of reconfiguration instructions as well as a formal operational semantics.

- Partners: IMT Atlantique - LS2N - LIP
- Contact: Maverick Chardet
- URL: https://gitlab.inria.fr/VeRDi-project/concerto

5.7. 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

5.8. 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 get_cluster_network_equipments - add a FAQ

- Participants: Florent Chuffart, Laurent Pouilloux and Matthieu Imbert
- Contact: Matthieu Imbert
- URL: http://execo.gforge.inria.fr

5.9. Platforms

5.9.1. 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/

5.9.2. 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

5.9.3. Platform: SILECS

Participants: Laurent Lefèvre, Simon Delamare, Christian Perez.

FUNCTIONAL DESCRIPTION

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/

6. New Results

6.1. Energy Efficiency in HPC and Large Scale Distributed Systems

Participants: Laurent Lefèvre, Dorra Boughzala, Thierry Gautier.

6.1.1. Performance and Energy Analysis of OpenMP Runtime Systems with Dense Linear Algebra Algorithms

In the article [4], 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. This paper was first published online in 2018-08-09.

6.1.2. 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. In 2019 we have explored the leverage energy-efficient non-lossy compression for data-intensive applications [9].

6.2. HPC Component Models and Runtimes

Participants: Thierry Gautier, Christian Perez, Laurent Turpin, Marie Durand, Philippe Virouleau.

6.2.1. Fine-Grained MPI+OpenMP Plasma Simulations: Communication Overlap with Dependent Tasks

In the article [15], we demonstrate how OpenMP 4.5 tasks can be used to efficiently overlap computations and MPI communications based on a case-study conducted on multi-core and many-core architectures. The paper focuses on task granularity, dependencies and priorities, and also identifies some limitations of OpenMP. Results on 64 Skylake nodes show that while 64% of the wall-clock time is spent in MPI communications, 60% of the cores are busy in computations, which is a good result. Indeed, the chosen dataset is small enough to be a challenging case in terms of overlap and thus useful to assess worst-case scenarios in future simulations. Two key features were identified: by using task priority we improved the performance by 5.7% (mainly due to an improved overlap), and with recursive tasks we shortened the execution time by 9.7%. We also illustrate the need to have access to tools for task tracing and task visualization. These tools allowed a fine understanding and a performance increase for this task-based OpenMP+MPI code.

6.2.2. Patches to LLVM compiler

We propose two source code patches to LLVM https://reviews.llvm.org/D63196 and https://reviews.llvm.org/D67447 in order to improve performance of application using numerous fine grain tasks such as [15]. Patches were accepted in 2019.

6.3. Modeling and Simulation of Parallel Applications and Distributed Infrastructures

Participants: Eddy Caron, Zeina Houmani, Frédéric Suter.

6.3.1. Bridging Concepts and Practice in eScience via Simulation-driven Engineering

The CyberInfrastructure (CI) has been the object of intensive research and development in the last decade, resulting in a rich set of abstractions and interoperable software implementations that are used in production today for supporting ongoing and breakthrough scientific discoveries. A key challenge is the development of tools and application execution frameworks that are robust in current and emerging CI configurations, and that can anticipate the needs of upcoming CI applications. In [14] we presented WRENCH, a framework that enables simulation-driven engineering for evaluating and developing CI application execution frameworks. WRENCH provides a set of high-level simulation abstractions that serve as building blocks for developing custom simulators. These abstractions rely on the scalable and accurate simulation models that are provided by the SIMGRID simulation framework. Consequently, WRENCH makes it possible to build, with minimum software development effort, simulators that can accurately and scalably simulate a wide spectrum of large and complex CI scenarios. These simulators can then be used to evaluate and/or compare alternate platform, system, and algorithm designs, so as to drive the development of CI solutions for current and emerging applications.

6.3.2. Accurately Simulating Energy Consumption of I/O-intensive Scientific Workflows

While distributed computing infrastructures can provide infrastructure-level techniques for managing energy consumption, application-level energy consumption models have also been developed to support energy-efficient scheduling and resource provisioning algorithms. In [7], we analyze the accuracy of a widely-used application-level model that have been developed and used in the context of scientific workflow executions. To this end, we profile two production scientific workflows on a distributed platform instrumented with power meters. We then conduct an analysis of power and energy consumption measurements. This analysis shows that power consumption is not linearly related to CPU utilization and that I/O operations significantly impact power, and thus energy consumption. We then propose a power consumption model that accounts for I/O operations, including the impact of waiting for these operations to complete, and for concurrent task executions on multi-socket, multi-core compute nodes. We implement our proposed model as part of a simulator that allows us to draw direct comparisons between real-world and modeled power and energy consumption. We find that our model has high accuracy when compared to real-world executions. Furthermore, our model improves accuracy by about two orders of magnitude when compared to the traditional models used in the energy-efficient workflow scheduling literature.

6.4. Cloud Resource Management

Participants: Eddy Caron, Jad Darrous, Christian Perez.

6.4.1. On the Importance of Container Image Placement for Service Provisioning in the Edge

Edge computing promises to extend Clouds by moving computation close to data sources to facilitate short-running and low-latency applications and services. Providing fast and predictable service provisioning time presets a new and mounting challenge, as the scale of Edge-servers grows and the heterogeneity of networks between them increases. Our work [6] is driven by a simple question: can we place container images across Edge-servers in such a way that an image can be retrieved to any Edge-server fast and in a predictable time. To this end, we present KCBP and KCBP-WC, two container image placement algorithms which aim to reduce the maximum retrieval time of container images. KCBP and KCBP-WC are based on k-Center optimization. However, KCBP-WC tries to avoid placing large layers of a container image on the same Edge-server. Evaluations using trace-driven simulations show that KCBP and KCBP-WC can be applied to various network configurations and reduce the maximum retrieval time of container images by 1.1x to 4x compared to state-of-the-art placements (*i.e.,* Best-Fit and Random).

Data-intensive clusters are heavily relying on distributed storage systems to accommodate the unprecedented growth of data. Hadoop distributed file system (HDFS) is the primary storage for data analytic frameworks such as Spark and Hadoop. Traditionally, HDFS operates under replication to ensure data availability and to allow locality-aware task execution of data-intensive applications. Recently, erasure coding (EC) is emerging as an alternative method to replication in storage systems due to the continuous reduction in its computation overhead. We have conducted an extensive experimental study to understand the performance of data-intensive applications under replication and EC [5], [23]. We use representative benchmarks on the Grid'5000 testbed to evaluate how analytic workloads, data persistency, failures, the back-end storage devices, and the network configuration impact their performances. Our study sheds the light not only on the potential benefits of erasure coding in data-intensive clusters but also on the aspects that may help to realize it effectively.

6.5. Data Stream Processing on Edge Computing

Participants: Eddy Caron, Felipe Rodrigo de Souza, Marcos Dias de Assunção, Laurent Lefèvre, Alexandre Da Silva Veith.

6.5.1. Operator Placement for Data Stream Processing on Fog/Edge Computing

DSP (Data Stream Processing) frameworks are often employed to process the large amount of data generated by the increasing number of IoT devices. A DSP application is commonly structured as a directed graph, or dataflow, whose vertices are operators that perform transformations over the incoming data and edges representing the data dependencies between operators. Such applications are often deployed on the Cloud in order to explore the large number of available resources and its pay-as-you-go business model. Fog computing enables offloading operators from the cloud by placing them close to where the data is generated, whereby reducing the time to process data events. However, fog computing resources often have lower capacity than those available in the Cloud. When offloading operators from the Cloud, the scheduler needs to adjust their level of parallelism and hence decides on the number of operator instances to create during placement in order to achieve a given throughput. This gives rise to two interrelated issues, namely deciding the operators parallelism and computing their placement onto available resources [16].

While addressing the placement problem [8], we proposed an approach consisting of a programming model and real-world implementation of an IoT application. The results show that our approach can minimise the end-to-end latency by at least 38% by pushing part of the IoT application to edge computing resources. Meanwhile, the edge-to-cloud data transfers are reduced by at least 38%, and the messaging costs are reduced by at least 50% when using the existing commercial edge cloud cost models.

In addition, we have designed and validated a discrete event simulation for modelling and simulation of DSP applications on edge computing environments [3].

6.5.2. Multi-Objective Reinforcement Learning for Reconfiguring Data Stream Analytics on Edge Computing

As DSP applications are often long-running, their workload and the infrastructure conditions can change over time. When changes occur, the application must be reconfigured. The operator reconfiguration consists of changing the initial placement by reassigning operators to different devices given target performance metrics. We modelled the operator reconfiguration as a Reinforcement Learning (RL) problem and defined a multi-objective reward considering metrics regarding operator reconfiguration, and infrastructure and application improvement [11]. We also use Monte Carlo Tree Search to organise the episodes generated during simulation and training [12]. Experimental results show that reconfiguration algorithms that minimise only end-to-end processing latency can have a substantial impact on WAN traffic and communication cost. The results also demonstrate that when reconfiguring operators, RL algorithms improve by over 50% the performance of the initial placement provided by state-of-the-art approaches.

6.6. An Operational Tool for Software Asset Management Improvement

Participants: Eddy Caron, Arthur Chevalier.

6.6.1. Multi-objective algorithm that guarantees license compliance

We have developed a new feature to OpTISAM, an Orange™ software offering tools to perform Software Asset Management (SAM) much more efficiently in order to be able to ensure the full compliance with all contracts from each software and a new type of deployment taking into account these aspects and other additional parameters like energy and performance. Our new feature is a multi-objective algorithm for deploying services in the Cloud that guarantees license compliance while reducing energy consumption but maintaining reasonable performance. In both cases of use and with a significant set of 5000 servers, we were able to show our approach is close to the best values in each criterion while dropping less than 10% of performance each time while keeping a full compliance.

6.7. Platform

Participants: Thierry Gautier, Christian Perez, Simon Delamare, Laurent Lefèvre.

6.7.1. Gemini cluster based on DGX-1 high density computer

The LECO experimental platform is a new medium size scientific instrument funded by DRRT and Inria to investigate research related to BigData and HPC. It is bi-located in Grenoble as part of the the HPCDA computer managed by UMS GRICAD (deployed in 2018) and in Lyon as part of the Grid5K Gemini cluster. The Gemini cluster is composed of two DGX-1 high density computers for HPC and BigData. Each computers has 8 NVIDIA V100 GPGPU cards with 4 infiniband high speed network cards.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. 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.

7.1.2. MUMPS Technologies

AVALON has a collaboration with MUMPS Technologies. The funding is dedicated for Marie Durand during few months to make experimental validation of the interest of using XKBLAS library to let MUMPS software to gain in performance on multi-GPUs server.

7.2. Bilateral Grants with Industry

7.2.1. 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.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. CPER

Participants: Thierry Gautier, Laurent Lefèvre, Christian Perez.

The LECO experimental platform is a new medium size scientific instrument deployed in Grenoble in 2018 and in Lyon in 2019. It was funded by the CPER 2015-2020 LECO++ to investigate research related to BigData and HPC.

8.1.2. Action Exploratoire Inria: EXODE

Participant: Thierry Gautier.

In biology, the vast majority of systems can be modeled as ordinary differential equations (ODEs). Modeling more finely biological objects leads to increase the number of equations. Simulating ever larger systems also leads to increasing the number of equations. Therefore, we observe a large increase in the size of the ODE systems to be solved. A major lock is the limitation of ODE numerical resolution software (ODE solver) to a few thousand equations due to prohibitive calculation time. The AEx ExODE tackles this lock via 1) the introduction of new numerical methods that will take advantage of the mixed precision that mixes several floating number precisions within numerical methods, 2) the adaptation of these new methods for next generation highly hierarchical and heterogeneous computers composed of a large number of CPUs and GPUs. For the past year, a new approach to Deep Learning has been proposed to replace the Recurrent Neural Network (RNN) with ODE systems. The numerical and parallel methods of ExODE will be evaluated and adapted in this framework in order to improve the performance and accuracy of these new approaches.

8.2. National Initiatives

8.2.1. Inria Large Scale Initiative

8.2.1.1. *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.

8.2.1.2. 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.

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

8.3.1.1. Energy oriented Centre of Excellence for computing applications (EoCoE-II)

Participants: Thierry Gautier, Christian Perez.

Program: H2020 RIA european project, call H2020-INFRAEDI-2018-1

Project acronym: EoCoE-II

Project title: Energy oriented Centre of Excellence for computing applications

Duration: 2018-2021

Coordinator: CEA

Other partners: CEA, FZJ, ENEA, BSC, CNRS, Inria, CERFACS, MPG, FRAUNHOFER, FAU, CNR, UNITN, PSNC, ULB, UBAH, CIEMAT, IFPEN, DDN, RWTH, UNITOV

Abstract: Europe is undergoing a major transition in its energy generation and supply infrastructure. The urgent need to halt carbon dioxide emissions and prevent dangerous global temperature rises has received renewed impetus following the unprecedented international commitment to enforcing the 2016 Paris Agreement on climate change. Rapid adoption of solar and wind power generation by several EU countries has demonstrated that renewable energy can competitively supply significant fractions of local energy needs in favourable conditions. These and other factors have combined to create a set of irresistible environmental, economic and health incentives to phase out power generation by fossil fuels in favour of decarbonized, distributed energy sources. While the potential of renewables can no longer be questioned, ensuring reliability in the absence of constant conventionally powered baseload capacity is still a major challenge.

The EoCoE-II project will build on its unique, established role at the crossroads of HPC and renewable energy to accelerate the adoption of production, storage and distribution of clean electricity. How will we achieve this? In its proof-of-principle phase, the EoCoE consortium developed a comprehensive, structured support pathway for enhancing the HPC capability of energy-oriented numerical models, from simple entry-level parallelism to fully-fledged exascale readiness. At the top end of this scale, promising applications from each energy domain have been selected to form the basis of 5 new Energy Science Challenges in the present successor project EoCoE-II that will be supported by 4 Technical Challenges

8.3.1.2. PRACE 6th Implementation Phase Project (PRACE6-IP)

Participants: Marcos Dias de Assunção, Laurent Lefèvre, Christian Perez.

Program: H2020 RIA european project, call H2020-INFRAEDI-2018-1

Project acronym: PRACE-6IP

Project title: PRACE 6th Implementation Phase Project

Duration: May 2019-Dec 2021

Coordinator: FZJ

Other partners: HLRS, LRZ, GENCI, CEA, CINES, CNRS, IDRIS, Inria, EPCC, BSC, CESGA, CSC, ETH-CSCS, SURFsara, KTH-SNIC, CINECA, PSNC, CYFRONET, WCNS, UiOsingma2, GRNET, UC-LCA, Univ MINHO, ICHEC, UHEM, CASTORCm NCSA, IT4I-VSB, KIFU, UL, CCSAS, CENAERO, Univ Lux, GEANT

Abstract: PRACE, the Partnership for Advanced Computing is the permanent pan-European High Performance Computing service providing world-class systems for world-class science. Systems at the highest performance level (Tier-0) are deployed by Germany, France, Italy, Spain and Switzerland, providing researchers with more than 17 billion core hours of compute time. HPC experts from 25 member states enabled users from academia and industry to ascertain leadership and remain competitive in the Global Race. Currently PRACE is finalizing the transition to PRACE 2, the successor of the initial five year period. The objectives of PRACE-6IP 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 development of PRACE 2; strengthening the internationally recognised PRACE brand; continuing and extend advanced training which so far provided more than 36 400 person·training days; preparing strategies and best practices towards Exascale computing, work on forward-looking SW solutions; coordinating and enhancing the operation of the multi-tier HPC systems and services; and supporting users to exploit massively parallel systems and novel architectures. A high level Service Catalogue is provided. The proven project structure will be used to achieve each of the objectives in 7 dedicated work packages. The activities are designed to increase Europe's research and innovation potential especially through: seamless and efficient Tier-0 services and a pan-European HPC ecosystem including national capabilities; promoting take-up by industry and new communities and special offers to SMEs;

assistance to PRACE 2 development; proposing strategies for deployment of leadership systems; collaborating with the ETP4HPC, CoEs and other European and international organisations on future architectures, training, application support and policies. This will be monitored through a set of KPIs.

8.4. International Initiatives

8.4.1. Inria International Labs

8.4.1.1. Joint Laboratory for Extreme Scale Computing (JLESC) (2014-2023)

Participants: Thierry Gautier, Christian Perez.

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.

Inria@EastCoast

Associate Team involved in the International Lab:

8.4.1.2. SUSTAM

Title: Sustainable Ultra Scale compuTing, dAta and energy Management

International Partner (Institution - Laboratory - Researcher):

Start year: 2017

See also: http://avalon.ens-lyon.fr/sustam

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).

8.5. International Research Visitors

8.5.1. Visits of International Scientists

Carlos Henrique Cardonha, IBM Research Brazil, from Jun 2019 until Jul 2019.

Jean-Philippe Aboumou, SAHAM Life Insurance, from Oct 2019.

8.5.1.1. Internships

Ibrahim Jouwad, M2, *Optimisation de la migration d'un ensemble de machines virtuelles dans un datacentre à l'aide d'un graphe d'états*

Laurent Turpin, M2, *Formalisation de paramètres, evaluation de performance et auto-configuration d'une application HPC en mémoire partagée : application au simulateur Aevol*

Josee Alvine Kouamen, M2, *Prise en main d'une infrastructure cloud et Big data pour l'analyse des fraudes a la simbox*

Zakaria Fraoui, *Distributed Stream Processing in the Edge: The Internet of Things Usecase*

Mohamed Hamnache, PFE, *Optimisation d'un environnement de calculs distribués pour la bio-informatique*

Alice Andres, M1, *Cloud vs Edge: fighting for energy !*

Adrien Berthelot, M1, *Revisiting low tech IT protocols*

Pierre Jacquot, L3, *Analysis of DDFacet/KillMS pipeline*

Marouane Azzouz, IUT, *Mode clients/serveur pour le projet CartomENSia*

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. General Chair, Scientific Chair

Eddy Caron

- served as Co-Chair of SPCLOUD 2019, July 15-19, 2019 Dublin, Ireland.
- served as Local Organizing Committee of SRDS 2019, October 1-4, 2019 Lyon. France.

Laurent Lefèvre was :

- Co-organizer of colloquium : "Atteindre les ODD en combinant leviers technologiques et computationnels, économie circulaire et transdisciplinarité. Mission impossible ?", with Centre Jacques Cartier, UQAM, Inria, GDS CNRS EcoInfo, UQAM, Montreal, Canada, November 4-5, 2019
- Co Special Session Organizer of Special Session on High Performance Computing Benchmarking and Optimization (HPBench 2019), during HPCS conference, Dublin, Ireland, July 15-19, 2019
- Co-organizer of GreenDays@Anglet : How people and machines can reduce their environmental impacts ?, Anglet, France, June 24-25, 2019

Christian Perez served as Awards Co-Chair of HPCS 2019, July 15-19, 2019 Dublin, Ireland.

9.1.1.2. Member of the Organizing Committees

Laurent Lefevre was member of the Organizing Committee of the TILECS Workshop (Grenoble, 3-4 Jul 2019).

Christian Perez was member of the Organizing Committee of the French Journées Calcul Données (Toulouse, 9-11 Oct 2019), and the TILECS Workshop (Grenoble, 3-4 Jul 2019).

9.1.2. Scientific Events: Selection

9.1.2.1. Member of the Conference Program Committees

E. Caron was member of the program committees of CLOSER’2019, Compas’19

C. Perez was member of the program committees of CCGRID’19, HPCS’19, ICPP’10, ParCo’19, SC’19, and Compas’19.

F. Suter served in the program committees of ICPP, CCGrid, e-science, ICA3PP, WORKS, and HeteroPar.

M. Dias de Assuncao was member of the program committees of CCGrid’19, UCC’19, and CloudCom’19.

9.1.3. Journal

9.1.3.1. Reviewer - Reviewing Activities

E. Caron reviewed an article for IEEE Transactions on Sustainable Computing (TSUSC) and an article for IEEE Transactions on Parallel and Distributed Systems (TPDS)

C. Perez reviewed an article for Future Generation Computer Systems.

F. Suter reviewed articles for Future Generation Computer Systems.

M. Dias de Assuncao reviewed articles for Future Generation Computer Systems, IEEE Transactions on Parallel and Distributed Systems, Cluster Computing and Journal of Network and Computer Applications.

9.1.4. Invited Talks

- Laurent Lefèvre gave the following invited talks in 2019 :
  - "Impacts environnementaux du numérique: rien de bon pour la planète ! Comment aller vers plus d'efficacité et de sobriété dans la phase d'usage ?", Entretiens Jacques Cartier, Montreal, Canada, November 4, 2019
  - "Impacts énergétiques et environnementaux du numérique dans les environnements urbains",Journée scientifique sur L'énergie dans les environnements urbains, Montpellier, France, October 10, 2019
  - "Une brève histoire du Green IT (informatique verte) : espoirs, challenges, risques", Café Gourmand du LIP, ENS Lyon, France, October 8, 2019
  - "L'usage des TICs - Consommation, efficacité et proportionnalité des TICs : focus sur usages DCs/Net et quelques propositions venant de la recherche", ENSIMAG, Grenoble, October 4, 2019
  - "Le numérique et ses impacts: C'est bon pour la planète ?", Meetup IBM Ecologie du numérique: facture énergétique de l'IA, Lyon, France, October 1, 2019
  - "Impact du numérique: focus sur les consommations en phase d'usage - Faire plus avec moins: à la recherche de proportionnalité et d'efficacité énergétique en phase d'usage", ANF 2019 EcoInfo, Autrans, France, September 25, 2019
  - "Mesures et efficacité énergétique sur plate-forme expérimentale - The G5K Green Tour", TGIR Visit, Lyon, France, July 5, 2019
  - "Mesures et efficacité énergétique sur plate-forme expérimentale - The G5K Green Tour", with Georges Da Costa, TILECS Workshop, Grenoble, France, July 4, 2019
  - "Eco-design or data centres collapse : - environmental impact of digital - 3 scenarii - GreenIT challenges and role", Colloque "Les limites de la croissance de la smart city:espaces et énergies des infrastructures numériques", Ecole d'architecture de la ville et des territoires, Marne la Vallée, France, June 5, 2019
  - "Numérique et ses impacts: C'est bon pour la planète ? de la grenouille au colibri...", Lycée St Charles, Rillieux la Pape, France, April 5, 2019
  - "Faire plus avec moins: à la recherche de proportionnalité et d'efficacité énergétique en phase d'usage", Séminaire du Département Informatique de l'ENS de Lyon (SIESTE), Lyon, France, March 12, 2019
  - "Les impacts environnementaux du numérique/EcoInfo/Réflexions Grenobloises/EcoInfo++", with Sophie Quinton, Comité des Projets Inria Rhone-Alpes, Montbonnot, February 12, 2019

9.1.5. Scientific Expertise

C. Perez reviewed two projects for PHC programmes.

Olivier Glück is member of the CNU (Conseil National des Universités) section 27 (Computer Science). He participated to the 2019 "Qualifications" session and "Suivi de carriere" session.

9.1.6. Research Administration

Eddy Caron is Deputy Director in charge of call for projects, research transfert and international affairs since September 2017 for the LIP. He is co-leader of the Distributed system and HPC team of the FIL (Fédération Informatique de Lyon).

Olivier Glück is member of the "Conseil Académique" of Lyon 1 University and Lyon University.

Laurent Lefevre is a member of the executive board and the sites committee of the Grid'5000 Scientific Interest Group. He is the scientific leader of the Grid'5000 Lyon site. He is animator and co-chair of the transversal action on "Energy" of the French GDR RSD ("Réseaux et Systèmes Distribués"). He is member of the scientifc advisory board of the Digital League cluster (Région Rhone Alpes). He is elected member in the LIP laboratory council (ENS Lyon).

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.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

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.

Licence: Yves Caniou, Algorithmique programmation impérative initiation, 48h, niveau L1, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Pratique d'Unix, 4h, niveau L1, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Programmation Concurrente, 35h and Responsible of UE, niveau L3, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Projet Informatique, 6h, niveau L3, Université Claude Bernard Lyon 1, France.

Licence: Yves Caniou, Réseaux, 36h, niveau L3, Université Claude Bernard Lyon 1, France.

Licence, Yves Caniou, Responsable mission pédagogique particulière, 3h, 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 Master SRIV (Systèmes, Réseaux et Infrastructures Virtuelles), 30h, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Projet Bibliographie et Certifications, 1.5h and Responsible of UE, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Gestion et supervision d'un parc, 4.5h and Responsible of UE, niveau M2, Université Claude Bernard Lyon 1, France.

Master: Yves Caniou, Sécurité, 27h and Responsible of UE, niveau M2, Université Claude Bernard Lyon 1, France.

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.

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, 42h, niveau M1, Université Claude Bernard Lyon 1, France.

Master: Laurent Lefèvre, Parallélisme, 12h, niveau M1, Université Lyon 1, France.

Master: Laurent Lefèvre, Réseaux avancés, 24h, niveau M2, IGA Casablanca, Morocco.

Licence: Olivier Glück, Licence pedagogical advisor, 30h, niveaux L1, L2, L3, Université Lyon 1, France.

Licence: Olivier Glück, Introduction Réseaux et Web, 54h, niveau L1, Université Lyon 1, France.

Licence: Olivier Glück, Bases de l'architecture pour la programmation, 23h, niveau L1, Université Lyon 1, France.

Licence: Olivier Glück, Algorithmique programmation impérative initiation, 56h, niveau L1, Université Lyon 1, France.

Licence: Olivier Glück, Réseaux, 2x70h, niveau L3, Université Lyon 1, France.

Master: Olivier Glück, Réseaux par la pratique, 10h, 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: Olivier Glück, DIU Enseigner l'Informatique au Lycée, 50h, Formation continue, Université Lyon 1, France.

Licence : Frédéric Suter, Programmation Concurrente, 50.66, L3, Université Claude Bernard Lyon 1, France

Master : Frédéric Suter, DIU Enseigner l'Informatique au Lycée, 30, M2, Université Claude Bernard Lyon 1, France

9.2.2. Supervision

PhD: Alexandre Da Silva Veith: *Quality of Service Aware Mechanisms for (Re)Configuring Data Stream Processing Applications on Highly Distributed Infrastructure*, 23 sept. 2019, Labex MiLyon, Laurent Lefèvre (dir), Marcos Dias de Assunção (co-dir) (2016-2019).

PhD: Jad Darrous, *Geo-distributed storage for distributed Cloud*, 17 dec. 2019, Inria, Gilles Fedak (dir) until Aug. 2017 then Christian Perez (dir), Shadi Ibrahim (co-dir).

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: 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: 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: 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) (since september 2016).

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) (since October 2017).

PhD in progress: Laurent Turpin,*Mastering Code Variation and Architecture Evolution for HPC application*, October 2019, Christian Perez (Inria, AVALON team, dir), Jonathan Rouzaud-Cornabas (INSA, Beagle team, co-dir) and Thierry Gautier (Inria, AVALON team, co-dir).

PhD in progress: Idriss Daoudi,*Simulating OpenMP program*, October 2018, Olivier Aumage (Inria, Storm team, Bordeaux, dir) and Thierry Gautier (Inria, AVALON team, co-dir).

PhD in progress: Vo Quoc Bao Bui, *Extended Para-Virtualization*, 2017, Alain Tchana (dir), Daniel Hagimont (INPT, co-dir)

PhD in progress: Barbe Thystere Mvondo Djob, *Improvement of the privileged domain in virtualized systems*, 1 fev 2018, Alain Tchana (dir), Noel De Palma (UGA, co-dir)

PhD in progress: Celestine Stella Ndonga Bitchebe, *Hardware features for virtualization*, 1 mars 2019, Alain Tchana (dir).

PhD in progress: Patrick Lavoisier Wapet, *Illegitimate app detection in mobile phones* , 1 oct. 2017, Alain Tchana (dir), Daniel Hagimont (INPT, co-dir).

9.2.3. Juries

Thierry Gautier was examiner of the PhD defense committee of Andrés Antón Rey Villaverde, Universidad Complutense de Madrid, Spain (November 22, 2019).

Laurent Lefèvre was

- examiner of the PhD defense of Fatma Ezzahra SALEM : "Management of Advanced Sleep Modes for Energy-Efficient 5G Networks", Institut Polytechnique de Paris, Telecom Paris Sud, December 20, 2019
- reviewer of the PhD of Maroua Haddad : "Sizing and management of a hybrid renewable energy system for data centers supply", University of Franche-Comté, November 28, 2019
- examiner of the PhD defense of Léo Grange : "Datacenter Management for on-site Intermittent and Uncertain Renewable Energy Sources", IRIT, Toulouse, October 3, 2019
- president of the PhD defense committee and examiner of of the PhD defense of Silvina Caino Lores : "On the Convergence of Big Data Analytics and High-Performance Computing: A Novel Approach for Runtime Interoperability", Universidad Carlos III de Madrid, Spain, July 8, 2019
- examiner of the PhD defense of Chaopeng Guo : "Energy-efficient Resource Provisioning for Cloud Database", IRIT, Toulouse, June 14, 2019
- reviewer of the PhD of Christian Heinrich : "Modeling, Prediction and Optimization of Energy Consumption of MPI Applications using SimGrid", Laboratoire Informatique de Grenoble, May 21, 2019

Christian Perez was

- president of the HdR defense committee of Alexandru Costan, ENS Rennes, France, March 14th, 2019.
- reviewer and member of the HdR defense committee of Xavier Etchevers, Université Grenoble Alpes, France, Nov 29th, 2019.
- reviewer and member of the HdR defense committee of Guillaume Mercier, Université de Bordeaux, France, Dec 4th, 2019.
- reviewer and member of the PhD defense committee of Ksander Ejjaaouani, Université de Strasbourg, France, Oct 25th, 2019.
- reviewer and member of the PhD defense committee of Georgios Christodoulis, Université Grenoble Alpes, France, Dec 5th, 2019.
- member of the PhD defense committee of Michael Mercier, Université Grenoble Alpes, France, July 1st, 2019.

- member of the PhD defense committee of Mohammad Mahdi Bazm, Université de Nantes, France, July 8th, 2019.
- member of the PhD defense committee of Cedric Deffo Sikounmo, Université Grenoble Alpes, France, Dec 18th, 2019.
- member of the final PhD committee for 7 PhD students of the University of Pisa (Italy), March 8th, 2019: Giovanna Broccia, Giulio Masetti, Luca Pedrelli, Giulio Ermanno Pibiri, Marco Ponza, Manuele Sabbadin, and Massimo Torquati,

Frédéric Suter was president of the PhD defense committee of Michael Mercier, Université Grenoble Alpes, France, July 1st, 2019.

9.3. Popularization

Yves Caniou est co-fondateur et co-organisateur du Campus du Libre, un événement autour du Libre dont l'objectif est de partager différents aspects du libre et des communs, allant par exemple du logiciel libre (Linux, Firefox, *etc.*) aux espaces communs gérés collaborativement (Wikipedia, OpenStreetMap).
La 2e édition s'est déroulée la journée du samedi 23 novembre 2019.

9.3.1. Articles and contents

- Interviews in order to popularize
  - Laurent Lefèvre was interviewed for the following dissemination media:
    * "En 2025, le numérique pourrait polluer autant que l'automobile", Le guide MAG2lyon du Développement Durable, December 2019
    * Radio show "Le numérique et ses impacts", Radio Fréquence Paris Plurielle, December 20, 2019
    * "L'insoutenable croissance du numérique", Alternatives Economiques, December 19, 2019
    * "Le streaming en expansion, son empreinte écologique aussi", Agence France Presse Dossier, October 28, 2019 - Reprint in Le Soleil, La Presse and le Quotidien (Canada, 28/10/2019), L'écho (France, 28/10/2019), Sciences et Avenir (France, 28/10/2019), L'OBS (France, 28/10/2019), Ouest France (France, 28/10/2019), Le Progrès (France 18/11/2019), DerStandard (Germany, 1/11/2019, "Gratis und ohne Anmeldung: Videostreaming-Dienst Apple TV+ gestartet"), Tierwelt (Germany, 3/11/2019, "STREAMING UND CO2 - Serienjunkies als Klimasunder")
    * "Mail, streaming, cloud... 20 gestes pour réduire sa pollution numérique", Business Insider France, October 29, 2019
    * Radio show "Le téléphone sonne", France Inter, August 16, 2019
    * "Guide Consommation Responsable - 3 gestes pour adopter une consommation numérique responsable", Institut National de la Consommation (INC), May 29, 2019
    * "Stocker trop de photos et de mails crée aussi de la pollution", Journal de Saone et Loire, after the Festival sans Decoder of Dompierre les Ormes, May 21, 2019
    * "Pollution numérique : comment réduire ses effets au quotidien ?", TV5 Monde, January 18, 2019

9.3.2. 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.

- Olivier Glück was in charge of one course in DIU Enseigner l'Informatique au Lycée, 50h, Formation continue, Université Lyon 1, France.

9.3.3. Panels

- Laurent Lefèvre was invited in the following panels in 2019 :
  - The "Empreinte écologique du numérique - Ecological footprint of digital world", Festival sans Decoder, Dompierre les Ormes, May 19, 2019
  - The "Energie numérique: ennemie ou alliée de l'écologie", Ateliers Sciences et Citoyens, Les rencontres du CNRS, Poitiers, April 2, 2019

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] J. DARROUS.*Scalable and Efficient Data Management in Distributed Clouds: Service Provisioning and Data Processing*, Ecole normale supérieure de lyon - ENS LYON, December 2019, https://hal.inria.fr/tel-02501316

[2] A. DA SILVA VEITH.*Quality of Service Aware Mechanisms for (Re)Configuring Data Stream Processing Applications on Highly Distributed Infrastructure*, ENS Lyon, CNRS & Inria ; LIP - Laboratoire de l'Informatique du Parallélisme, September 2019, https://hal.inria.fr/tel-02385744

Articles in International Peer-Reviewed Journal

[3] G. AMARASINGHE, M. DIAS DE ASSUNCAO, A. HARWOOD, S. KARUNASEKERA.*ECSNeT++ : A simulator for distributed stream processing on edge and cloud environments*, in "Future Generation Computer Systems", November 2019, p. 1-18, forthcoming [*DOI :* 10.1016/J.FUTURE.2019.11.014], https://hal.inria.fr/hal-02369500

[4] 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", 2019, vol. 33, n^{o} 3, p. 431-443 [*DOI :* 10.1177/1094342018792079], https://hal.inria.fr/hal-01957220

International Conferences with Proceedings

[5] J. DARROUS, S. IBRAHIM, C. PÉREZ.*Is it time to revisit Erasure Coding in Data-intensive clusters?*, in "MASCOTS 2019 - 27th IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems", Rennes, France, IEEE, October 2019, p. 165-178 [*DOI :* 10.1109/MASCOTS.2019.00026], https://hal.inria.fr/hal-02263116

[6] J. DARROUS, T. LAMBERT, S. IBRAHIM.*On the Importance of Container Image Placement for Service Provisioning in the Edge*, in "ICCCN 2019 - 28th International Conference on Computer Communications and Networks", Valencia, Spain, IEEE, July 2019, p. 1-9 [*DOI :* 10.1109/ICCCN.2019.8846920], https://hal.inria.fr/hal-02134507

[7] R. FERREIRA DA SILVA, A.-C. ORGERIE, H. CASANOVA, R. TANAKA, E. DEELMAN, F. SUTER.*Accurately Simulating Energy Consumption of I/O-intensive Scientific Workflows*, in "ICCS 2019 - International Conference on Computational Science", Faro, Portugal, ICCS 2019 - International Conference on Computational Science, Springer, June 2019, p. 138-152 [*DOI :* 10.1007/978-3-030-22734-0_11], https://hal.archives-ouvertes.fr/hal-02112893

[8] E. GIBERT RENART, A. DA SILVA VEITH, D. BALOUEK-THOMERT, M. DIAS DE ASSUNCAO, L. LEFÈVRE, M. PARASHAR.*Distributed Operator Placement for IoT Data Analytics Across Edge and Cloud Resources*, in "CCGrid 2019 - 19th Annual IEEE/ACM International Symposium in Cluster, Cloud, and Grid Computing", Larnaca, Cyprus, May 2019, p. 1-10 [*DOI :* 10.1109/CCGRID.2019.00060], https://hal.inria.fr/hal-02103942

[9] I. RAÏS, D. BALOUEK-THOMERT, A.-C. ORGERIE, L. LEFÈVRE, M. PARASHAR.*Leveraging energy-efficient non-lossy compression for data-intensive applications*, in "HPCS 2019 - 17th International Conference on High Performance Computing & Simulation", Dublin, Ireland, HPCS 2019 - 17th International Conference on High Performance Computing & Simulation, July 2019, p. 1-7, https://hal.archives-ouvertes.fr/hal-02179621

[10] C. SÉGUIN, E. CARON, S. DUBUS.*SeeDep: Deploying Reproducible Application Topologies on Cloud Platform*, in "CLOSER 2019 - 9th International Conference on Cloud Computing and Services Science", Heraklion, Greece, May 2019, vol. 1, p. 363-370 [*DOI :* 10.5220/0007721103630370], https://hal.archives-ouvertes.fr/hal-02119654

[11] A. DA SILVA VEITH, F. R. DE SOUZA, M. DIAS DE ASSUNCAO, L. LEFÈVRE, J. C. S. DOS ANJOS.*Multi-Objective Reinforcement Learning for Reconfiguring Data Stream Analytics on Edge Computing*, in "ICPP 2019 - 48th International Conference on Parallel Processing", Kyoto, Japan, ACM, August 2019, p. 1-10 [*DOI :* 10.1145/3337821.3337894], https://hal.inria.fr/hal-02140844

[12] A. DA SILVA VEITH, M. DIAS DE ASSUNCAO, L. LEFÈVRE.*Monte-Carlo Tree Search and Reinforcement Learning for Reconfiguring Data Stream Processing on Edge Computing*, in "SBAC-PAD 2019 - International Symposium on Computer Architecture and High Performance Computing", Campo Grande, Brazil, October 2019, p. 1-8, https://hal.inria.fr/hal-02305472

National Conferences with Proceeding

[13] D. DELABROYE, S. DELAMARE, D. LOUP, L. NUSSBAUM.*Remplacer un routeur par un serveur Linux : retour d'expérience des passerelles d'accès à Grid'5000*, in "JRES - Journées Réseaux de l'Enseignement et de la Recherche", Dijon, France, December 2019, https://hal.inria.fr/hal-02401684

Conferences without Proceedings

[14] R. FERREIRA DA SILVA, H. CASANOVA, R. TANAKA, F. SUTER.*Bridging Concepts and Practice in eScience via Simulation-driven Engineering*, in "BC2DC 2019 - Workshop on Bridging from Concepts to Data and Computation for eScience", San Diego, CA, United States, September 2019, p. 1-6, https://hal.archives-ouvertes.fr/hal-02329541

[15] J. RICHARD, G. LATU, J. BIGOT, T. GAUTIER.*Fine-Grained MPI+OpenMP Plasma Simulations: Communication Overlap with Dependent Tasks*, in "Euro-Par 2019: Parallel Processing - 25th International Conference on Parallel and Distributed Computing", Göttingen, Germany, Springer, August 2019, p. 419-433 [*DOI :* 10.1007/978-3-030-29400-7_30], https://hal-cea.archives-ouvertes.fr/cea-02404825

[16] F. RODRIGO DE SOUZA, M. DIAS DE ASSUNCAO, E. CARON.*A Throughput Model for Data Stream Processing on Fog Computing*, in "HPCS 2019 - 17th International Conference on High Performance Computing & Simulation", Dublin, Ireland, July 2019, p. 1-7, https://hal.archives-ouvertes.fr/hal-02140851

Scientific Books (or Scientific Book chapters)

[17] A. OLEKSIAK, L. LEFÈVRE, P. ALONSO, G. DA COSTA, V. DE MAIO, N. FRASHERI, V. GARCIA, J. GUERRERO, S. LAFOND, A. LASTOVETSKY, R. R. MANUMACHU, B. MUITE, A.-C. ORGERIE, W. PIATEK, J.-M. PIERSON, R. PRODAN, P. STOLF, E. SHEME, S. VARRETTE.*Energy aware ultrascale systems*, in "Ultrascale Computing Systems", Institution of Engineering and Technology, January 2019, p. 127-188 [*DOI :* 10.1049/PBPC024E_CH], https://hal.inria.fr/hal-02163289

Research Reports

[18] F. BERTHOUD, P. GUITTON, L. LEFÈVRE, S. QUINTON, A. ROUSSEAU, J. SAINTE-MARIE, C. SERRANO, J.-B. STEFANI, P. STURM, E. TANNIER.*Sciences, Environnements et Sociétés : Rapport long du groupe de travail MakeSEnS d'Inria*, Inria, October 2019, https://hal.inria.fr/hal-02340948

[19] C. DIGUET, F. LOPEZ, L. LEFÈVRE.*L'impact spatial et énergétique des data centers sur les territoires*, ADEME, Direction Villes et territoires durables, 2019, p. 1-141, https://hal.archives-ouvertes.fr/hal-02133607

Software

[20] T. GIVARO GROUP.*Givaro*, June 2019, Version : 4.1.1
[SWH-ID : swh:1:dir:df65912bd1e5ea4b96b935de95f6638eb6d9472d], Software, https://hal.archives-ouvertes.fr/hal-02130729

[21] T. LINBOX GROUP.*LinBox*, June 2019, Version : 1.6.3
[SWH-ID : swh:1:dir:393b611a1424f032e83569bf6762502371cfcf65], Software, https://hal.archives-ouvertes.fr/hal-02130801

Scientific Popularization

[22] C. BONAMY, L. LEFÈVRE, G. P. A. MOREAU.*High performance computing and energy efficiency: focus on OpenFOAM*, in "Congrès JRES : Les Journées Réseaux de l'Enseignement et de la Recherche", Dijon, France, RENATER, December 2019, p. 1-9, https://arxiv.org/abs/1912.03920 - Vidéo https://replay.jres.org/videos/watch/811bb2fe-582e-4996-9643-89401d2c0d2c, https://hal.archives-ouvertes.fr/hal-02387863

Other Publications

[23] J. DARROUS, S. IBRAHIM.*Enabling Data Processing under Erasure Coding in the Fog*, August 2019, 1, ICPP 2019 - 48th International Conference on Parallel Processing, Poster, https://hal.inria.fr/hal-02388835

References in notes

[24] 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° 5, p. 658–671 [*DOI :* 10.1109/TPDS.2009.76], http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4906989

[25] 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

[26] 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

[27] 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

[28] G. MADEC.*NEMO ocean engine*, Institut Pierre-Simon Laplace (IPSL), France, 2008, n^{o} 27, ISSN No 1288-1619

[29] OPENACC.*The OpenACC Application Programming Interface*, November 2011, Version 1.0, http://www.openacc-standard.org

[30] OPENMP ARCHITECTURE REVIEW BOARD.*OpenMP Application Program Interface*, July 2011, Version 3.1, http://www.openmp.org

[31] 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

[32] C. SZYPERSKI.*Component Software - Beyond Object-Oriented Programming*, 2, Addison-Wesley / ACM Press, 2002, 608

[33] 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
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]
Eric Tannier [Inria, Senior Researcher, HDR]
Leonardo Trujillo Lugo [Inria, Advanced Research Position, from Nov 2019]

Faculty Members

Guillaume Beslon [Team leader, INSA Lyon, Full Professor, HDR]
Carole Knibbe [INSA Lyon, 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]

PhD Students

Audrey Denizot [INSA Lyon, PhD Student]
Julie Etienne [INSERM, PhD Student, from Oct 2019]
Marco Foley [Inria, PhD Student, from Nov 2019]
Théotime Grohens [INSA Lyon, PhD Student, from Sep 2019]
Vincent Liard [INSA Lyon, PhD Student]

Post-Doctoral Fellow

Charles Rocabert [Inria, Post-Doctoral Fellow, from May 2019]

Visiting Scientist

Leonardo Trujillo Lugo [INSA Lyon, until Sep 2019]

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. 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.

[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.

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* [27] 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. [28]). 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. Functional and Evolutionary Biology

We do not distinguish our research and its application domains. Our shared idea is that the research is oriented by a scientific question, which is a multidisciplinary one, most often of biological nature. We do not develop methodologies independently from this question and then look for applications. Instead we collectively work with other disciplines to solve a question, with our competencies.

In consequence the application domains are already listed in the description of our projects and goals. They concern functional and evolutionary biology, related to critical social questions as human and plant health.

4.2. Social and Environmental Responsibility (Implication domains)

These last years we have maintained a frequent team discussion on the social and environmental responsibility of researchers. It has become more frequent this year, with the announcements of serious environmental issues by many governmental or non governmental organizations.

We are engaged in many actions regarding this responsibility. A constant ethics questioning, directing our research projects according to our values, teaching and popularization of ethical values. In particular we are engaged in several research projects on health and environment, and one of us has been a member of the institutional workgroup on environmental issues at Inria.

Regarding the functioning of research activities, we attempted a measure of the environmental footprint of our activities, regardless of their aims. It has the shape of a carbon footprint analysis, gathering the carbon footprint of travels, computer usage, computer equipment. We are aware of the incompleteness of this analysis, as well by not including many activities (nutrition, homeplace-workplace trips), and not taking other environmental issues than carbon emissions.

However it is a starting point, that we presented to many colleagues who were interested in reproducing the computation, so we give the headlines here. We used the unitary costs given by a website we constructed: https://ferme.yeswiki.net/Empreinte. The data was collected on an average taken over 3 years, 2016-2018 (we cannot yet at this stage make the analysis for 2019). Travels by members of the team or invited researchers emitted 39.86tCO2. Computing hours on local clusters emitted 17.46tCO2. The acquisition of computers accounts for the emission of 4.53tCO2. The total is 61.85tCO2.

Based on this lower bound of our CO2 emissions, the total per person is around 3 tons per person. Probably this number is highly underevaluated since it accounts for only part of our professional activities.

Whether it is too high or acceptable, and the establishment of a carbon budget is a difficult question. If we refer to the goals of the COP21 conclusions, we should emit less than 2 tons of CO2 per person in 2050 to reach carbon neutrality. This includes professional and personal life, and all the services we benefit from. We did not arrive yet at a consensus on the objective we should reach to consider we have a sutainable activity, but we in majority recognize that we are anyway far above what would be a consensus objective. We are still in a discussion to engage in a reduction.

5. Highlights of the Year

5.1. Highlights of the Year

Last year our highlights were focused on remarkable publications. This year the main events are on organizations and grants applications.

- We have been in charge of organizing the Scientific Days of Inria in July 2019 https://project.inria.fr/journeesscientifiques2019/
- We were awarded two exploratory actions by Inria in 2019, one on high performance computing, the other in agro-ecology
- We were auditioned for an ERC synergy grant call (very last step in the many steps for the grant obtention)
- We organized MMEE https://mmee2019lyon.sciencesconf.org/ in Lyon

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 P. 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. 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 P. Parsons, Eric Tannier and Benoît Morel
- Partner: Laboratoire de Biométrie et Biologie Evolutive (LBBE) - UMR CNRS 5558
- Contact: Eric Tannier

7. New Results

7.1. Computational Glioscience: A book to review the existing mathematical models of the glial cells

[participant: H. Berry]

Over the last two decades, the recognition that astrocytes - the predominant type of cortical glial cells - could sense neighboring neuronal activity and release neuroactive agents, has been instrumental in the uncovering of many roles that these cells could play in brain processing and the storage of information. These findings initiated a conceptual revolution that leads to rethinking how brain communication works since they imply that information travels and is processed not just in the neuronal circuitry but in an expanded neuron-glial network. On the other hand the physiological need for astrocyte signaling in brain information processing and the modes of action of these cells in computational tasks remain largely undefined. This is due, to a large extent, both to the lack of conclusive experimental evidence, and to a substantial lack of a theoretical framework to address modeling and characterization of the many possible astrocyte functions. This book [https://hal.inria.fr/hal-01995842] aims at filling this gap, providing the first systematic computational approach to the complex, wide subject of neuron-glia interactions. The organization of the book is unique insofar as it considers a selection of “hot topics” in glia research that ideally brings together both the novelty of the recent experimental findings in the field and the modelling challenge that they bear. A chapter written by experimentalists, possibly in collaboration with theoreticians, will introduce each topic. The aim of this chapter, that we foresee less technical in its style than in conventional reviews, will be to provide a review as clear as possible, of what is “established” and what remains speculative (i.e. the open questions). Each topic will then be presented in its possible different aspects, by 2-3 chapters by theoreticians. These chapters will be edited in order to provide a “priming” reference for modeling neuron-glia interactions, suitable both for the graduate student and the professional researcher.

7.2. The impact of tracers on lipid digestion kinetics

[Participant: Carole Knibbe]

Dietary fats are present in the diet under different types of structures, such as spread vs emulsions (notably in processed foods and enteral formula), and interest is growing regarding their digestion and intestinal absorption. In clinical trials, there is often a need to add stable isotope-labeled triacylglycerols (TAGs) as tracers to the ingested fat in order to track its intestinal absorption and further metabolic fate. Because most TAG tracers contain saturated fatty acids, they may modify the physicochemical properties of the ingested labeled fat and thereby its digestion. However, the actual impact of tracer addition on fat crystalline properties and lipolysis by digestive lipases still deserves to be explored. In this context, we monitored the thermal and polymorphic behavior of anhydrous milk fat (AMF) enriched in homogeneous TAGs tracers and further compared it with the native AMF using differential scanning calorimetry and power X-ray diffraction. As tracers, we used a mixture of tripalmitin, triolein and tricaprylin at 2 different concentrations (1.5 and 5.7wt%, which have been used in clinical trials). The addition of TAG tracers modified the AMF melting profile, especially at the highest tested concentration (5.7 wt%). Both AMF and AMF enriched with 1.5wt% tracers were completely melted around 37°C, i.e. close to the body temperature, while the AMF enriched with 5.7wt% tracers remained partially crystallized at this temperature. Similar trends were observed in both bulk and emulsified systems. Moreover, the kinetics of AMF polymorphic transformation was modified in the presence of tracers. While only β' form was observed in the native AMF, the β-form was clearly detected in the AMF containing 5.7wt% tracers. We further tested the impact of tracers on the lipolysis of AMF in bulk using a static in vitro model of duodenal digestion. Lipolysis of AMF enriched with 5.7wt% tracers was delayed compared with that of AMF and AMF enriched with 1.5wt% tracers. Therefore, low amounts of TAG tracers including tripalmitin do not have a high impact on fat digestion, but one has to be cautious when using higher amounts of these tracers.

7.3. The control of synaptic plasticity by external factors

[participant: H. Berry]

The dorsal striatum exhibits bidirectional corticostriatal synaptic plasticity, NMDAR- and endocannabinoids-(eCB)-mediated, necessary for the encoding of procedural learning. Therefore, characterizing factors controlling corticostriatal plasticity is of crucial importance. Brain-derived neurotrophic factor (BDNF) and its receptor, the tropomyosine receptor kinase- B (TrkB), shape striatal functions and their dysfunction deeply affects basal ganglia. BDNF/TrkB signaling controls NMDAR-plasticity in various brain structures including the striatum. However, despite cross-talk between BDNF and eCBs, the role of BDNF in eCB- plasticity remains unknown. In https://hal.inria.fr/hal-02076121, we show that BDNF/TrkB signaling promotes eCB-plasticity (LTD and LTP) induced by rate-based (low-frequency stimulation) or spike-timing- based (spike-timing-dependent plasticity, STDP) paradigm in striatum. We show that TrkB activation is required for the expression and the scaling of both eCB-LTD and eCB-LTP. Using two-photon imaging of dendritic spines combined with patch-clamp recordings, we show that TrkB activation prolongs intracellular calcium transients, thus increasing eCB synthesis and release. We provide a mathematical model for the dynamics of the signaling pathways involved in corticostriatal plasticity. Finally, we show that TrkB activation enlarges the domain of expression of eCB-STDP. Our results reveal a novel role for BDNF/TrkB signaling in governing eCB-plasticity expression in striatum, and thus the engram of procedural learning.

7.4. A new model for calcium signals in tiny sub-cellular domains

[participants: A. Denizot, H. Soula, H. Berry]

Astrocytes, a glial cell type of the central nervous system, have emerged as detectors and regulators of neuronal information processing. Astrocyte excitability resides in transient variations of free cytosolic calcium concentration over a range of temporal and spatial scales, from sub-microdomains to waves propagating throughout the cell. Despite extensive experimental approaches, it is not clear how these signals are transmitted to and integrated within an astrocyte. The localization of the main molecular actors and the geometry of the system, including the spatial organization of calcium channels IP3R, are deemed essential. However, as most calcium signals occur in astrocytic ramifications that are too fine to be resolved by conventional light microscopy, most of those spatial data are unknown and computational modeling remains the only methodology to study this issue. In https://hal.inria.fr/hal-02184344v2, we propose an IP3R-mediated calcium signaling model for dynamics in such small sub-cellular volumes. To account for the expected stochasticity and low copy numbers, our model is both spatially explicit and particle-based. Extensive simulations show that spontaneous calcium signals arise in the model via the interplay between excitability and stochasticity. The model reproduces the main forms of calcium signals and indicates that their frequency crucially depends on the spatial organization of the IP3R channels. Importantly, we show that two processes expressing exactly the same calcium channels can display different types of calcium signals depending on the spatial organization of the channels. Our model with realistic process volume and calcium concentrations successfully reproduces spontaneous calcium signals that we measured in calcium micro-domains with confocal microscopy and predicts that local variations of calcium indicators might contribute to the diversity of calcium signals observed in astrocytes. To our knowledge, this model is the first model suited to investigate calcium dynamics in fine astrocytic processes and to propose plausible mechanisms responsible for their variability.

7.5. Evolution of genome size

Using the Aevol software, we investigated the dynamics of genome size under different evolutionary pressures (variation of mutation rates and variation of population sizes). The dynamics of the model enabled us to identify a new mutational pressure on genome size that spontaneously increase the fraction of non-coding sequences. We showed that this mutational pressure interact with the selective pressure for robustness (knibbe et al., 2007), resulting in an equilibrium of genome size and non-coding proportion. Moreover, we showed that this equilibrium can change depending on the size of the population due to the resulting effect on selection intensity. A paper has been published in the proceedings of the ALife 2019 conference (cardes et al, 2019) and an article in in preparation.

7.6. Dynamics of evolutionary innovation

Using a combination of mathematical and computational models (NK-Fitness-Landscapes and Aevol), we investigated the dynamics of innovation in evolving systems. We showed that innovation is often triggered by specific mutational events, typically structural variation of the genome (e.g. duplications, inversions, ...). We further studied this effect and showed that innovation is due to the differences of time scale between the different kinds of mutations: fast mutations (typically point mutations) are rapidly exhausted, resulting in a fitness plateau. However, slow mutations (typically structural variations) can open new evolutionary paths, resulting in the population escaping from the fitness plateau. An article is in preparation in collaboration with Santiago F. Elena (CSIC, Spain).

7.7. Evolution of biological complexity

Using a modified version of the Aevol platform, we studied the evolution of complex features. By evolving population of organisms in conditions where complexity is counter-selected, we showed that complexity accumulates even in these conditions, i.e. even when complex organisms are less fit than simple ones. Moreover we showed that complex organisms are not more robust and not more evolvable than simple ones. This shows that evolution spontaneously initiate a "complexity ratchet" that forces complexity to grow. An article is in press in the Artificial Life Journal (to be published in 2020).

7.8. Dynamics of mutator strains

In a long-lasting collaboration with Utrecht University, we studied the dynamics of mutator strains in constant environments (mutator strains being individuals which mutation rate is increased by several orders of magnitude). Contrary to what is generally admitted, we showed that, although mutators initially suffer from a mutational burden (in coherence with the theory), they are able to quickly recover and avoid the burden.

Moreover, we showed that they do so by contracting their coding genome compartment and expanding their non-coding compartment. This result show that mutators can thrive even in a constant environment (rutten et al., 2019).

7.9. Mutiscale phylogenetics models

[Participant: Eric Tannier]

We progressed in the modeling of multi-scale phylogenetic events: we gave an algorithm to infer gene conversions according to a phylogeny [7], a complexity result and an algorithm for transfers with replacements [6], and we devised a simulation tool integrating extinct species and horizontal inheritance [3].

7.10. Evolution of the Drosophila melanogaster Chromatin Landscape and Its Associated Proteins

[participant: A. Crombach]

In the nucleus of eukaryotic cells, genomic DNA associates with numerous protein complexes and RNAs, forming the chromatin landscape. Through a genome-wide study of chromatin-associated proteins in Drosophila cells, five major chromatin types were identified as a refinement of the traditional binary division into hetero- and euchromatin. These five types were given color names in reference to the Greek word chroma. They are defined by distinct but overlapping combinations of proteins and differ in biological and biochemical properties, including transcriptional activity, replication timing, and histone modifications. We assessed the evolutionary relationships of chromatin-associated proteins and presented an integrated view of the evolution and conservation of the fruit fly Drosophila melanogaster chromatin landscape. We combined homology prediction across a wide range of species with gene age inference methods to determine the origin of each chromatin-associated protein. This provided insight into the evolution of the different chromatin types. Our results indicate that for the euchromatic types, YELLOW and RED, young associated proteins are more specialized than old ones; and for genes found in either chromatin type, intron/exon structure is lineage-specific. Next, we provide evidence that a subset of GREEN-associated proteins is involved in a centromere drive in D. melanogaster. Our results on BLUE chromatin support the hypothesis that the emergence of Polycomb Group proteins is linked to eukaryotic multicellularity. In light of these results, we discuss how the regulatory complexification of chromatin links to the origins of eukaryotic multicellularity.

8. Partnerships and Cooperations

8.1. Regional Initiatives

- CPER LECO++: Parallel HPC architectures evolve and the calculation codes are naturally bound to vary over time. Indeed, the architectures change every 2-3 years while the lifespan of a scientific code is much longer (at least 10 years). Knowing how to control the impacts of these changes in order to automatically adapt the digital simulation codes to maintain a high level of performance is a necessity to guarantee a certain sustainability of the developed code. Currently, these variations are manually managed by programmers which require a high level of expertise as well as time.

  A collaboration between the AVALON teams from LIP and BEAGLE from LIRIS on this subject involved one master trainees this year (funding from Federation Informatique de Lyon – PMSISEE project). More specifically, BEAGLE is interested in designing AEVOL a high performance parallel code for simulating the evolution of a population of bacteria. The different parts of the code have been adapted to the hardware characteristics of current architectures (multicore, vector computing, etc.) for which certain operations have several implementations (CPU or vector) or several parallel variants. Designing the assembly of the right versions and choosing the right parameters remains a difficult problem. In this issue, the AVALON team brings its expertise in the development and exploitation of component models, in parallel programming models and in the expertise of executive supports for HPC.

A PhD thesis between Avalon and Beagle (Laurent Turpin) linked to the CPER LECO++ project (coordinator: T. Gautier, AVALON) has started with the aim of studying the robustness of computer codes on modern parallel architectures and their evolution. Thus, the targeted hardware is that being acquired through the LECO++ project (ARM machine, massively multi-GPU (10)).

The work of this thesis aims to study the methods and approaches allowing to contribute to a solution to the problems of composition, choice of parameters and efficient execution on a parallel architecture in HPC. The problem addressed in the thesis concerns the portability of the performance of a parallel application for managing code variants and variations at runtime. The solutions that will be studied will be those at the interface between a programming model and its exploitation by executive support. In order to exploit the performance of a class of machines in a portable manner, the candidate will propose the necessary adaptations, whether to the existing component-based programming model (typically Comet) and to executive support (OpenMP type or an executive engine with task base). A major constraint of this work is the performance at execution: the evaluation will be based on an experimental methodology with AEVol as target application. The target hardware is that of an HPC computing node of tomorrow: a multi-core server coupled with a large number of hardware accelerators - GPUs - allowing to have a significant computing density (approximately from 30 to 128 TFlops double precision for 4 to 16 GPUs).

8.2. National Initiatives

8.2.1. ANR

- Evoluthon (2019-2022): Artificial Life as a benchmark for evolutionary studies, a 4-year project leaded by E Tannier with 2 partners, Beale Inria and Le Cocon, LBBE.
- 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

- 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).

- Action Exploratoire "Community Garden Book": IPBES's recent report on declining biodiversity calls for generalization of agroecological, productive, biodiversity and environmental friendly methods, oriented towards participatory action research. This exploratory action is a proposal to develop tools from open science, evolution science and algorithmics for the co-construction and use of an agroecological network of interactions between groups, species, varieties found in fields and gardens.
- Action Exploratoire ExODE: In biology, the vast majority of systems can be modeled as ordinary differential equations (ODEs). Modeling more finely biological objects leads to increase the number of equations. Simulating ever larger systems also leads to increasing the number of equations. Therefore, we observe a large increase in the size of the ODE systems to be solved. A major lock is the limitation of ODE numerical resolution so ware (ODE solver) to a few thousand equations due to prohibitive calculation time. The AEx ExODE tackles this lock via 1) the introduction of new numerical methods that will take advantage of the mixed precision that mixes several floating number precisions within numerical methods, 2) the adaptation of these new methods for next generation highly hierarchical and heterogeneous computers composed of a large number of CPUs and GPUs. For the past year, a new approach to Deep Learning has been proposed to replace the Recurrent Neural Network (RNN) with ODE systems. The numerical and parallel methods of ExODE will be evaluated and adapted in this framework in order to improve the performance and accuracy of these new approaches.

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Declared Inria International Partners

- Beagle is a member of the CNRS Laboratoire International Associé "EvoAct" (Evolution in Action). Other members of EvoAct are the TIMC-IMAG (Grenoble) and the Beacon Center (Michigan State University, USA).

8.3.1.2. Informal International Partners

- Collaboration with Alexander Fleischmann at Brown University (USA) on neuro-evo-devo.
- Collaboration with Cedric Chauve, SFU, Vancouver (Canada) on phylogeny and rearrangements.
- Collaboration with Tom Williams, Bristol (UK) on phylogeny.

8.4. International Research Visitors

8.4.1. Visits of International Scientists

- We welcomed Leonardo Trujillo (Venezuela) as a visiting professor from January 2019 to July 2019. Leonardo Trujillo worked on the innovation dynamics in evolution using NK Fitness-Landscapes.
- Corrado Cali, BESE Division, KAUST University, Saudi Arabia, 1 week in november

8.4.1.1. Internships

- Barbara Genocchi (PhD candidate, Tampere University of Technology, Tampere, Finland) visited us for 16 days (Sept 9 - Sept 24).

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. General Chair, Scientific Chair

- Eric Tannier was the organizing committee chair of the Inria scientific days, July 2019.

9.1.1.2. Member of the Organizing Committees

- Guillaume Beslon and Anton Crombach, members of the organizing committee of MMEE 2019 (Mathematical Modeling in Ecology and Evolution, Lyon, July 2019)
- Guillaume Beslon : member of the organizing committee of the EvoLyon Day (Lyon, November 2019)
- Guillaume Beslon : member of the organizing committee of BDA 2019 (Bases de Données Avancées, Lyon, October 2019)
- Anton Crombach was part of the organizing and scientific committees of the winter course "Advanced Lectures in Computational Systems Biology" in Aussois.
- Hugues Berry was a member of the Local Organising Committee for conference MedInfo 2019 (https://www.medinfo-lyon.org/en/)
- Eric Tannier is a member of the organizing committee of "SEMOVI", séminaire de modélisation du vivant, a local scientifica animation in systems biology.

9.1.2. Scientific Events: Selection

9.1.2.1. Chair of Conference Program Committees

- Eric Tannier was the program committee chair of the Inria scientific days, July 2019.

9.1.2.2. Member of the Conference Program Committees

- Anton Crombach, involved in the selection of talks and posters for MMEE (Mathematical models in ecology and evolution, Lyon, July)
- Christophe Rigotti, member of the program committee of the 19th IEEE International Conference on Data Mining (ICDM).
- Christophe Rigotti, member of the program committee of the 34rd ACM Symposium On Applied Computing (SAC).
- Guillaume Beslon and Jonathan Rouzaud-Cornabas, members of the program committee of ALife 2019 (2019 Conference on Artificial Life. Newcastle. United Kingdom)
- Eric Tannier, member of the program committee of Jobim 2020
- Eric Tannier, member of the program committee of RECOMB Comparative Genomics 2019
- Eric Tannier, member of the program committee of ISBM/ECCB 2019

9.1.2.3. Reviewer

- Eric Tannier reviewed for RECOMB 2020

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Hugues Berry is Associate Editor for PLoS Computational Biology
- Eric Tannier is a guest editor for the epijournal DMTCS, and edited a special issue in 2019
- Eric Tannier is an editor of "Peer Community in Evolutionary Biology"
- Eric Tannier is a founding editor of "Peer Community in Mathematical and Computational Biology"

9.1.3.2. Reviewer - Reviewing Activities

- Bioinformatics, Journal of Experimental Zoology Part B...

9.1.4. Invited Talks

- Carole Knibbe, 2019 Gordon Conference on "Organismal, Cellular, Molecular and Theoretical Approaches to Understanding Evolution", Easton, MA, USA, June 9-14 2019.
- Anton Crombach, lab seminar at LBMC
- Anton Crombach, lab seminar at LBBE
- Anton Crombach, invited speaker at the EvoLyon conference in november 2019
- Hugues Berry, Artificial Intelligence and Health: the ambition of the French Plan, Aviesan Meeting, Brussels Belgium, Dec. 2019
- Hugues Berry, An introduction to the theoretical models for passive molecule movements in living cells, Action Nationale de Formation du CNRS, Measurement of Molecular Dynamics in Living Cells, Lille, France, Oct. 2019
- Hugues Berry, Searching for deterministic chaos in biological data,Dynamical Systems and Applications to Biology, CIMPA School, Dhaka, Bangladesh, June 2019
- Hugues Berry, Anomalous diffusion in living cells : bridging the gap between experiments and models through collaborative challenges, Random Walks and Intracellular Transport, School of Mathematics, University of Manchester, Manchester, UK, April 2019
- Hugues Berry, Investigating the effect of the nanoscale architecture of astrocytic processes on the propagation of calcium signals, Tampere University of Technology, Tampere, Finland, April 2019
- Hugues Berry, Inria research at the forefront of digital sciences for medicine, 2019 French-American Innovation Day (FAID2019), Houston, Texas, USA, march 2019
- Carole Knibbe, Eric Tannier, Which social and environmental responsibility for researchers, lab seminar at LBMC
- Eric Tannier, open science as an institutional response to the editorial crisis in research, "open science day" of the research biology federation, Lyon.
- Guillaume Beslon, invited talk at UGA (séminaire "Translation et Innovation en Médecine et Complexité") in November 2019
- Guillaume Beslon, invited talk at the I4ID 2019 conference (Immunotherapies & Innovations for Infectious Diseases Congress, Lyon, December 2019)

9.1.5. Leadership within the Scientific Community

- Anton Crombach is the lead editor of book "Evolutionary Systems Biology, 2nd edition" (ongoing work).
- Hugues Berry organised the 1st Singapore-France Joint Workshop on AI in Health, in association with the international conference MedInfo (https://medinfo-lyon.org/en/), Lyon, France, August 27-29, 2019.
- Hugues Berry co-organized the INSERM Workshop "Intracellular dynamics of molecules: analysis and models" (https://research.pasteur.fr/en/event/intracellular-dynamics-of-molecules-analysis-and-models/). The workshop combined a 3-day critical assessment phase (in Bordeaux, France, June 24-26 2019) and a 3-day technical worskshop (in Lyon, France, July 1-3 2019). This workshop addressed a biologist audience and covered the main experimental methods to quantify the mobility and trajectories of biomolecules in living cells, with an emphasis on the quantification methods for individual trajectories and the interest of computer simulations for analysis and interpretation.
- Carole Knibbe and Guillaume Beslon (together with Dusan Misevic from the Center for Research and Interdisciplinarity in Paris) edited a special issue of the Artificial Life journal (vol25 issue 4, nov 2019), with extended versions of the best papers of ECAL 2017. Link : https://www.mitpressjournals.org/doi/abs/10.1162/artl_a_00298?af=R

9.1.6. Scientific Expertise

- Hugues Berry, Reviewer for the Fond National de la Recherche Scientifique (FNRS, Belgium) for the call "FRQ-FNRS Québec - Communauté française de Belgique"

9.1.7. Research Administration

- Christophe Rigotti, elected member of Insa Scientific board (Conseil scientifique)
- Eric Tannier, elected member of the Inria Administration Council
- Eric Tannier, member of the open science committee of the French Ministry of research
- Eric Tannier, member of the ethics platform of the university of Lyon
- Eric Tannier, member of the "Sciences Environments Societies" workgroup of Inria
- Eric Tannier, member of the Environmental responsibility workgroup of LBBE
- Eric Tannier, scientific responsible of the "cours colloques" committee of Inria
- Eric Tannier, member of the scientific committee of the science shop, univ Lyon
- Hugues Berry, Deputy scientific director of Inria for digital biology and health
- Guillaume Beslon, member of the ANR CE45 committee (Mathématiques et sciences du numérique pour la biologie et la santé)
- Guillaume Beslon, member of the IRD CSS5 committee (Science des données et des modèles)

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

License: Jonathan Rouzaud-Cornabas, Object-Oriented Programming, 120h, L3, Computer Science Department, INSA de Lyon

Master: Jonathan Rouzaud-Cornabas, System Programming, 60h, M1, Computer Science Department, INSA de Lyon

Master: Jonathan Rouzaud-Cornabas, Network Programming, 20h, M1, Computer Science Department, INSA de Lyon

Master: Jonathan Rouzaud-Cornabas, High Performance Computing, 60h, M2, Computer Science Department, INSA de Lyon

Master: Jonathan Rouzaud-Cornabas, High Performance Computing, 12h, M2, Bioinformatics and Modeling Department, INSA de Lyon

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

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, Comparative Genomics, 8h, M1, Bioinformatics and Modeling Department of INSA-Lyon.

Master: Eric Tannier, String algorithmics, 12h, M1, Bioinformatics and Modeling Department of INSA-Lyon.

Master: Eric Tannier, String algorithmics, 12h, M1, Bioinformatics UCBL.

Master: Eric Tannier, Research Ethics, 6h, M2, Bioinformatics UCBL

Doctorat: Eric Tannier, Environmental responsibility of research activities, NGS formation, Lyon

Doctorat: Eric Tannier, Research Ethics, 12h, all specialities, Université de Lyon

Doctorat: Eric Tannier, Research Ethics, 6h, Inria PhD students

Master: Audey Denizot, 64h de cours en enzymologie (M1, INSA Lyon) et biologie cellulaire (M2, ENS Lyon).

E-learning

MOOC: Eric Tannier, Research Ethics, FUN, Ph-D candidates, 3000 registered participants

9.2.2. Supervision

PhD in progress: Julie Etienne, "Modélisation et simulation du flux de triglycérides alimentaires, de l'absorption entérocytaire à la sécrétion des chylomicrons", INSA-Lyon, co-supervised by Carole Knibbe and Marie-Caroline michalski (CarMeN laboratory), started in October 2019.

M2 : Julie Etienne, "Analysing and modelling the traffic of triglycerides through enterocytes", co-supervised by Carole Knibbe (80

M2 : Ella Beaumann, "Modélisation compartimentale de cinétiques postprandiales cliniques", co-supervised by Carole Knibbe (70

M2: Hugues Berry and Christophe Rigotti where co-supervisors of the M2 internship of Tymofii Prokopenko (co-supervision with Olivier Pascual of team SynatAc INSERM/CNRS) from 28/01/2019 to 14/06/2019. Title: "Classification and clustering of activity data of microglial cells".

M2 Sujin Hyun, 1 feb 2019 to 31 jul 2019 (6 months), title "Re-assessing the link between splicing and mRNA quality control" (supervised by A Crombach)

PhD: Audrey Denizot, "Simulation de la signalisation calcique dans les prolongements fins astrocytaires", INSA Lyon, 08 Nov. 2019 (co-supervision H. Soula and H. Berry)

PhD: Hugues Berry participated to the PhD Advisory Committee of Barbara Genocchi (BioMediTech, Tampere University of Technology, Finland)

PhD: Eric Tannier participated to the PhD Advisory Committee of Raphael Forquet (UCBLyon)

PhD in progress: Alexandre Laverré, "influence croisée de l'organisation spatiale et des mutations structurales dans les génomes", encadré par Anamaria Necsulea et Eric Tannier, début en septembre 2018

PhD in progress: Théo Tricou, "Détection d'espèces éteintes avec les transferts horizontaux", encadré par Damien de Vienne et Eric Tannier, début en septembre 2018

PhD in progress: Hugo Menet, "Phylogénie multi-échelles des holobiontes", encadré par Eric Tannier et Vincent Daubin, début en septembre 2019

M2: Hugo Menet, "Phylogénie multi-échelles des holobiontes", encadrée par Eric Tannier, 2019

9.2.3. Juries

- HDR: Jérémie Roux,"Integrative single-cell approaches to understanding cancer drug response heterogeneity in tumor cell dynamics" Université Côte d'Azur, Nice, France, September 2019 (Hugues Berry, reviewer)
- HDR: Jean-Baptiste Masson, "Probabilistic induction and physics modelling to probe biological decision making In situ", Sorbonne University, Paris, France, May 2019 (Hugues Berry, reviewer)
- HDR: Celine Scornavacca, "Phylogenetic reconciliations and networks", Univ Montpellier (Eric Tannier, examinateur)
- PhD: Audrey Denizot (Ph-D, INSA-Lyon, October 2019) (Guillaume Beslon was a member of the PhD Committee)
- PhD: Andreas Odorico (Ph-D, Gif-sur-Yvette, December 2019) Guillaume Beslon was reviewer
- PhD: Adrien Legrand (University of Picardie Jules Verne, Amiens, November 2019) Christophe Rigotti was reviewer

9.3. Popularization

9.3.1. Articles and contents

- Hugues Berry participated in popularization activities with the edition of a book, written in French, on the relationships between sex (i.e. a biology concept) and Gender (a sociology approach) (Abou, B. and Berry, H., eds., (2019). Sexe et Genre: de la Biologie a la Sociologie. Editions Materiologiques, Paris, https://materiologiques.com/sciences-philosophie-2275-9948/282-sexe-et-genre-de-la-biologie-a-la-sociologie.html) In recent years, debates about sex and gender concepts between biology and the humanities and social sciences have become more intense, both for scientific reasons and for their societal impacts. Many biologists reject the questioning made by researchers in humanities and social sciences, of what biology considers fundamental, such as sex binarity or sex differences. Humanities and social scientists, on the other hand, often see biology as an academic and institutional source of naturalistic arguments used to oppose gender studies. They denounce biases in the interpretation of biologists as resulting, precisely, from gender-related biases. However, scientific exchanges at the interface between biology and the humanities and social sciences are undoubtedly necessary to overcome these antagonisms. The objective of this book is to implement a dialogue between the two fields, trying to overcome misunderstandings and misconceptions that have gone on for long. The book consists in 12 chapters stemming from biologists and sociologists, organized so as to try and foster dialogue between both communities. The target readership of the book is composed of non specialist readers (in particular non biologists nor sociologists) with a pre-existing interest on gender issues and their relations to biology. That is why the book was written in French and includes chapters intended to newcomers in the field (e.g. "The determination of sex in the human species and the rest of life" or "Why Gender?") to more advanced viewpoints ("What we learn from transidentities"). The objective of the publisher was to publish a reference book in French on the relations between sex and gender.
- "Le mouvement est devenu massif", an interview of Hugues Berry to the French journal L'Usine Nouvelle, Special Edition "A year of simulation", number 3608, April 25, 2019. https://www.usinenouvelle.com/editorial/le-mouvement-est-devenu-massif.N834680
- Eric Tannier is a co-author of "Quand les branches de l'arbre du vivant s'entremêlent", in *Pour la science*, 2019.

9.3.2. Education

- We have an ongoing collaboration with the Grenoble Rectorat. The aim is to use our computational models to train school professor to evolutionary concepts.

9.3.3. Interventions

- Audrey Denizot participated to activities in the mediation association DéMesures:

- "Cosmograff" project. This project aims to present the solar system and its scales to the general public, in collaboration with the Musée des Confluences and the collective of street artists Superposition https://bit.ly/35gtYpI. Creation of audio guides in French and English for reuse at the 2019 EWASS international astronomy conference. https://bit.ly/2rQlB5y
- "ArtScience" project, which aims to present the work of scientists to artists. This project culminated in an artistic exhibition at the ENS Lyon, from April to July 2019, in collaboration with the Taverne Gutenberg https://bit.ly/2OldNR3. Creation of a series of photographs presenting the links between science and art that emerged from this project https://bit.ly/2XqjaCz.

- Eric Tannier co-organized a citizen science event at the Théâtre Nouvelle Génération, Vaise, december 2019
- Eric Tannier organizes a workshop on environmental responibility of researchers in january 2020
- Eric Tannier participated to a "pint of science" evening on the social and environmental responsibility of researchers, May 20, 2019

9.3.4. Internal action

- Eric Tannier and Carole Knibbe gave a lecture on environmental responsibility at the Inria Bio days, 2019
- Eric Tannier gave a training to PhD students on research ethics, 6h, Inria Grenoble Rhône-Alpes, January 2019

9.3.5. Creation of media or tools for science outreach

- We developed the GreenMice game, a small video-game designed to teach the basis of evolutionary biology to young children

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. DENIZOT. *Simulating calcium signaling in fine astrocytic processes*, Université de Lyon, November 2019, https://tel.archives-ouvertes.fr/tel-02406699

Articles in International Peer-Reviewed Journal

[2] S. DANTHINE, C. VORS, D. AGOPIAN, A. DURAND, R. GUYON, F. CARRIÈRE, C. KNIBBE, M. LÉTISSE, M.-C. MICHALSKI. *Homogeneous triacylglycerol tracers have an impact on the thermal and structural properties of dietary fat and its lipolysis rate under simulated physiological conditions*, in "Chemistry and Physics of Lipids", December 2019, vol. 225, 104815 [*DOI :* 10.1016/J.CHEMPHYSLIP.2019.104815], https://hal.archives-ouvertes.fr/hal-02326893

[3] A. A. DAVIN, T. TRICOU, E. TANNIER, D. M. DE VIENNE, G. J. SZÖLLOSI. *Zombi: A phylogenetic simulator of trees, genomes and sequences that accounts for dead lineages*, in "Bioinformatics", 2019 [*DOI :* 10.1093/BIOINFORMATICS/BTZ710], https://hal.archives-ouvertes.fr/hal-02302010

[4] A. DENIZOT, M. ARIZONO, V. U. NÄGERL, H. SOULA, H. BERRY. *Simulation of calcium signaling in fine astrocytic processes: effect of spatial properties on spontaneous activity*, in "PLoS Computational Biology", August 2019, vol. 15, n^{o} 8, e1006795, p. 1-33 [*DOI :* 10.1371/JOURNAL.PCBI.1006795], https://hal.inria.fr/hal-02184344

[5] G. GANGAROSSA, S. PEREZ, Y. DEMBITSKAYA, I. PROKIN, H. BERRY, L. VENANCE.*BDNF controls bidirectional endocannabinoid-plasticity at corticostriatal synapses*, in "Cerebral Cortex", April 2019, vol. bhz081, p. 1-18 [*DOI :* 10.1093/CERCOR/BHZ081], https://hal.inria.fr/hal-02076121

[6] D. HASIC, E. TANNIER.*Gene tree reconciliation including transfers with replacement is NP-hard and FPT*, in "Journal of Combinatorial Optimization", 2019, vol. 38, n^o 2, p. 502–544 [*DOI :* 10.1007/S10878-019-00396-Z], https://hal.archives-ouvertes.fr/hal-02301454

[7] D. HASIC, E. TANNIER.*Gene tree species tree reconciliation with gene conversion*, in "Journal of Mathematical Biology", 2019, vol. 78, n^o 6, p. 1981–2014, https://arxiv.org/abs/1703.08950 [*DOI :* 10.1007/S00285-019-01331-W], https://hal.archives-ouvertes.fr/hal-01495707

[8] C. KNIBBE, G. BESLON, D. MISEVIC.*Introduction to the ECAL 2017 Special Issue*, in "Artificial Life", November 2019, vol. 25, n^o 4, p. 313-314 [*DOI :* 10.1162/ARTL_A_00298], https://hal.inria.fr/hal-02428658

[9] 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", January 2019, vol. 33, n^o 1, p. 131-167 [*DOI :* 10.1007/S10618-018-0591-9], https://hal.archives-ouvertes.fr/hal-01898015

[10] E. PAREY, A. CROMBACH.*Evolution of the Drosophila melanogaster Chromatin Landscape and Its Associated Proteins*, in "Genome Biology and Evolution", January 2019, vol. 11, n^o 3, p. 660-677 [*DOI :* 10.1093/GBE/EVZ019], https://hal.inria.fr/hal-02064719

[11] S. PEIGNIER, C. RIGOTTI, G. BESLON.*Evolutionary Subspace Clustering Using Variable Genome Length*, in "Computational Intelligence", 2020, p. 1-39, forthcoming, https://hal.archives-ouvertes.fr/hal-02405598

[12] J. P. RUTTEN, P. HOGEWEG, G. BESLON.*Adapting the engine to the fuel: mutator populations can reduce the mutational load by reorganizing their genome structure*, in "BMC Evolutionary Biology", 2019, vol. 19, n^o 1, p. 1-17 [*DOI :* 10.1186/S12862-019-1507-Z], https://hal.archives-ouvertes.fr/hal-02350040

International Conferences with Proceedings

[13] Q. CARDE, M. FOLEY, C. KNIBBE, D. P. PARSONS, J. ROUZAUD-CORNABAS, G. BESLON.*How to reduce a genome? ALife as a tool to teach the scientific method to school pupils*, in "ALIFE 2019 - Conference on Artificial Life", Newcastle, United Kingdom, MIT Press, July 2019, p. 497-504 [*DOI :* 10.1162/ISAL_A_00211], https://hal.inria.fr/hal-02285718

[14] L. FOULON, S. FENET, C. RIGOTTI, D. JOUVIN.*Scoring Message Stream Anomalies in Railway Communication Systems*, in "LMID 2019 - IEEE Workshop on Learning and Mining with Industrial Data", Beijing, China, November 2019, p. 1-8, https://hal.archives-ouvertes.fr/hal-02357924

[15] T. NGUYEN, N. MÉGER, C. RIGOTTI, C. POTHIER, N. GOURMELEN, E. TROUVÉ.*A pattern-based mining system for exploring Displacement Field Time Series*, in "19th IEEE International Conference on Data Mining (ICDM) Demo", Beijing, China, Proceedings of the IEEE International Conference on Data Mining (ICDM) Demo, IEEE, November 2019, p. 1110-1113, https://hal.archives-ouvertes.fr/hal-02361793

National Conferences with Proceeding

[16] L. FOULON, C. RIGOTTI, S. FENET, D. JOUVIN.*Approximation du score CFOF de détection d'anomalie dans un arbre d'indexation iSAX : Application au contexte SI de la SNCF*, in "EGC 2019 - 19ème Conférence francophone sur l'Extraction et la Gestion des Connaissances", Metz, France, 2019, p. 1-12, https://hal.inria.fr/hal-02019035

Scientific Books (or Scientific Book chapters)

[17] M. DE PITTÀ, H. BERRY.*Computational Glioscience*, Springer Series in Computational Neuroscience, Springer, 2019, p. 1-505 [*DOI :* 10.1007/978-3-030-00817-8], https://hal.inria.fr/hal-01995842

[18] M. DE PITTÀ, E. BEN-JACOB, H. BERRY.*G protein-coupled receptor-mediated calcium signaling in astrocytes*, in "Computational Glioscience", M. DE PITTÀ, H. BERRY (editors), Springer Series in Computational Neuroscience, Springer, January 2019, p. 115-150 [*DOI :* 10.1007/978-3-030-00817-8_5], https://hal.inria.fr/hal-01995850

[19] M. DE PITTÀ, H. BERRY.*A Neuron-Glial Perspective for Computational Neuroscience*, in "Computational Glioscience", M. DE PITTÀ, H. BERRY (editors), Springer Series in Computational Neuroscience, Spinger, January 2019, p. 3-35 [*DOI :* 10.1007/978-3-030-00817-8_1], https://hal.inria.fr/hal-01995849

[20] A. DENIZOT, H. BERRY, S. VENUGOPAL.*Intracellular Ca 2+ Dynamics in Astrocytes: Modeling the Underlying Spatiotemporal Diversity*, in "Encyclopedia of Computational Neuroscience", 2019, p. 1-19, forthcoming, https://hal.inria.fr/hal-02419317

[21] J. LALLOUETTE, M. DE PITTÀ, H. BERRY.*Astrocyte networks and intercellular calcium propagation*, in "Computational Glioscience", M. DE PITTÀ, H. BERRY (editors), Springer Series in Computational Neuroscience, January 2019, p. 177-210 [*DOI :* 10.1007/978-3-030-00817-8_7], https://hal.inria.fr/hal-01995852

[22] M. STIMBERG, D. F. M. GOODMAN, R. BRETTE, M. DE PITTÀ.*Modeling Neuron–Glia Interactions with the Brian 2 Simulator*, in "Computational Glioscience", M. D. PITTÀ, H. BERRY (editors), January 2019, p. 471-505 [*DOI :* 10.1007/978-3-030-00817-8_18], https://hal.sorbonne-universite.fr/hal-02325277

Research Reports

[23] F. BERTHOUD, P. GUITTON, L. LEFÈVRE, S. QUINTON, A. ROUSSEAU, J. SAINTE-MARIE, C. SERRANO, J.-B. STEFANI, P. STURM, E. TANNIER.*Sciences, Environnements et Sociétés : Rapport long du groupe de travail MakeSEnS d'Inria*, Inria, October 2019, https://hal.inria.fr/hal-02340948

Scientific Popularization

[24] E. TANNIER, B. BOUSSAU, V. DAUBIN.*Quand les branches de l'arbre du vivant s'entremêlent*, in "Pour la science", 2019, https://hal.archives-ouvertes.fr/hal-02414245

Other Publications

[25] L. FOULON, S. FENET, C. RIGOTTI, D. JOUVIN.*Detecting Anomalies over Message Streams in Railway Communication Systems*, September 2019, 1, AALTD@ECML/PKDD 2019 - 4th Workshop on Advanced Analytics and Learning on Temporal Data. Poster, Poster, https://hal.archives-ouvertes.fr/hal-02357927

[26] C. ROCABERT, G. BESLON, C. KNIBBE, S. BERNARD.*Phenotypic Noise and the Cost of Complexity*, November 2019, EvoLyon 2019, Poster, https://hal.archives-ouvertes.fr/hal-02402443

References in notes

[27] P. GALISON.*Image and Logic: A Material Culture of Microphysics*, University Of Chicago Press, 1997

[28] 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

Project-Team CASH

Compilation and Analyses for Software and Hardware

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
Architecture, Languages and Compilation

Table of contents

Project-Team CASH

Creation of the Team: 2018 April 01, updated into Project-Team: 2019 June 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
A2.5.3. - Empirical Software Engineering
A2.5.4. - Software Maintenance & Evolution

Other Research Topics and Application Domains:

B9.5.1. - Computer science

1. Team, Visitors, External Collaborators

Research Scientists

Christophe Alias [Inria, Researcher, HDR]
Ludovic Henrio [CNRS, Researcher, 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 [École Normale Supérieure de Lyon, PhD Student]
Paul Iannetta [École Normale Supérieure de Lyon, PhD Student]
Amaury Maille [École Normale Supérieure de Lyon, PhD Student, from Oct 2019]

Administrative Assistants

Kadiatou Bangoura [Inria, Administrative Assistant, until Aug 2019]
Sophie Gerard [Inria, Administrative Assistant, from Sep 2019]

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 [33] 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, the CASH team works 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 works 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 is 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 focuses 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 directions 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 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 relies 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 ourselves to a strict definition of dataflow languages: 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 [58] 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 [42] or Streamit [55], where the scheduling is fully static, and general communicating networks like KPNs [45] or RVC-Cal [25] 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 is 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 can express parallelism in a dataflow way instead of the classical communication-oriented dependencies. The higher-level more declarative point of view makes 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 [32] that nowadays introduce more and more dataflow primitives to enable data-driven interactions between agents, particularly with *futures* [30], [37].

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 [22], [23], [38] and theoretical [43] 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 [50], [46], as well as tailoring static analyses for specialized applications in compilation [36], [54], Synchronous Dataflow scheduling [53], and extending the polyhedral model to irregular applications [21]. 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 [39]. 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 [47], 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 [41] 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 [26] 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 [31] 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 [56] and rely directly on array dataflow analysis [35]. 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 [34]. 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 [59] 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) [23]. 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 [40]. 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 ($\approx$ 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 [38]. 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) [23]. 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 [59] 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 [52] 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 [23], [22]) 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 [51].

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 [24]. 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 opens 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 [60]. Hence, we expect a *unique* optimal set of parameters. Thus, we need to 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) [20] in SystemC [19] 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 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 [28]. 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 [57], 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 started working on the visualization of the sequence of steps leading to a deadlock when the situation occurs, and will 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 [57] from CEA-LIST, some compute-intensive parts delegated to sc-during [49] from Matthieu Moy, and a co-simulation with external physical solvers using SystemC-MDVP [27] 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 gives 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. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Awards

In January 2019, the paper "Static Analysis Of Binary Code With Memory Indirections Using Polyhedra" resulting from a collaboration with colleagues from Lille University, has received a best paper award of the VMCAI 2019 conference.

The paper "Godot: All the Benefits of Implicit and Explicit Futures" received the distinguished artefact at ECOOP'19.

BEST PAPERS AWARDS :

[5]
C. BALLABRIGA, J. FORGET, L. GONNORD, G. LIPARI, J. RUIZ.*Static Analysis Of Binary Code With Memory Indirections Using Polyhedra*, in "VMCAI'19 - International Conference on Verification, Model Checking, and Abstract Interpretation", Cascais, Portugal, LNCS, Springer, January 2019, vol. 11388, p. 114-135 [*DOI :* 10.1007/978-3-030-11245-5_6], https://hal.archives-ouvertes.fr/hal-01939659

[8]
A. CHARIF, G. BUSNOT, R. MAMEESH, T. SASSOLAS, N. VENTROUX.*Fast Virtual Prototyping for Embedded Computing Systems Design and Exploration*, in "RAPIDO2019 - 11th Workshop on Rapid Simulation and Performance Evaluation: Methods and Tools", Valence, Spain, January 2019, p. 1-8 [*DOI :* 10.1145/3300189.3300192], https://hal.archives-ouvertes.fr/hal-02023805

5. New Software and Platforms

5.1. 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

5.2. 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

5.3. 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/

6. New Results

6.1. Dataflow-explicit futures

Participants: Ludovic Henrio, Matthieu Moy, Amaury Maillé.

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 [43], 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:

- With collaborators from University of Uppsala and University of Oslo we worked on the design of programming constructs mixing implicit and dataflow-explicit futures (DeF). This has been published in ECOOP 2019 [10].
- Amaury Maillé did his internship in the Cash team (advised by Matthieu Moy and Ludovic Henrio), he worked on an implementation of DeF in the Encore language. This raised a difficulty regarding the interaction of DeF with generic types that has been partially solved. Now we need to generalize our approach to completely solve the issue.

6.2. Distributed futures

Participant: Ludovic Henrio.

We proposed the definition of *distributed futures*, a construct that provides at the same time a data container similar to a distributed vector, and a single synchronization entity that behaves similarly to a standard future. This simple construct makes it easy to program a composition, in a task-parallel way, of several massively data-parallel tasks. This work will be presented in Sac 2020 (we are currently working on the final version of the paper). This work is realised in collaboration with Pierre Leca and Wijnand Suijlen (Huawei Technologies), and Françoise Baude (Université Côte d'Azur, CNRS, I3S).

6.3. 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. After 2 meetings in 2019, A journal article is still being written but the visit of Reiner Hähnle in the Cash team during two months (as invited professor) in Spring 2019 should allow us to make faster progress on the topic.

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).

6.4. 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:

- we extended the work to support the manipulation of overlapping array. This was accepted as an extended version of our conference paper (presented at 4PAD 2018). It will be published in the JLAMP journal in 2020 [3].

6.5. 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. This work is realized with Eric Madelaine (Inria Sophia-Antipolis) and Rabéa Ameur Boulifa (Telecom ParisTech). A journal article has been written and will be submitted in January 2020.
- A translation from BIP model to open pNets has being formalized and encoded, this work is done in collaboration with Simon Bliudze (Inria Lille). More precisely, we extend the theory of architectures developed previously for the BIP framework with the elements necessary for handling data: definition and operations on data domains, syntax and semantics of composition operators involving data transfer. To verify that individual architectures do enforce their associated properties , we provide an encoding into open pNets, an intermediate model that supports SMT-based verification. This work has been published in Coordination 2019 [6].

These works are under progress and should be continued in 2020.

6.6. 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 have been published in TACAS 2019 [11].

6.7. A Survey on Verified Reconfiguration

Participant: Ludovic Henrio.

We are conducting a survey on the use of formal methods to ensure safety of reconfiguration of distributed system, that is to say the runtime adaptation of a deployed distributed software system. The survey article is written together with Hélène Coullon and Simon Robillard (IMT Atlantique, Inria, LS2N, UBL), and Frédéric Loulergue (Northern Arizona University). Hélène Coullon is the coordinator and we expect the article to be submitted in 2020.

6.8. A Survey on Parallelism and Determinacy

Participants: Ludovic Henrio, Laure Gonnord, Matthieu Moy, Christophe Alias.

We have started to investigate on the solutions that exist to ensure complete or partial determinacy in parallel programs. The objective of this work is to provide a survey based on the different kinds of solutions that exist to ensure determinism or at least limit data-races in concurrent execution of programs. The study will cover language-based, compilation-based and also runtime-based solutions. We started the bibliographic studies in 2019. The objective of this work is to write and submit a survey article in 2020.

This work, coordinated by Laure Gonnord and Ludovic Henrio, also involves contributors outside the CASH team. For the moment Gabriel Radanne (Inria Paris) and Lionel Morel (CEA).

6.9. Pipeline-aware Scheduling of Polyhedral Process Networks

Participants: Christophe Alias, Julien Rudeau.

The polyhedral model is a well known framework to develop accurate and optimal automatic parallelizers for high-performance computing kernels. It is progressively migrating to high-level synthesis through polyhedral process networks (PPN), a dataflow model of computation which serves as intermediate representation for high-level synthesis. Many locks must be overcome before having a fully working polyhedral HLS tool, both from a front-end (C $\rightarrow$ PPN) and back-end (PPN $\rightarrow$ FPGA) perspective. In this work [15], we propose a front-end scheduling algorithm which reorganizes the computation of processes to maximize the pipeline efficiency of the processes' arithmetic operators. We show that our approach improve significantly the overall latency as well as the pipeline efficiency.

6.10. A Compiler Algorithm to Guide Runtime Scheduling

Participants: Christophe Alias, Samuel Thibault, Laure Gonnord.

Task-level parallelism is usually exploited by a runtime scheduler, after tasks are mapped to processing units by a compiler. In this report, we propose a compilation-centric runtime scheduling strategy. We propose a complete compilation algorithm to split the tasks in three parts, whose properties are intended to help the scheduler to take the right decisions [16]. In particular, we show how the polyhedral model may provide a precious help to compute tricky scheduling and parallelism informations. Our compiler is available and may be tried online at http://foobar.ens-lyon.fr/kut.

This is a joint work with University of Bordeaux, which will be continued next year.

6.11. fkcc: the Farkas Calculator

Participant: Christophe Alias.

We propose a new domain-specific language and a tool, FKCC, to prototype program analyses and transformations exploiting the affine form of Farkas lemma. Our language is general enough to prototype in a few lines sophisticated termination and scheduling algorithms. The tool is freely available and may be tried online via a web interface. We believe that FKCC is the missing chain to accelerate the development of program analyses and transformations exploiting the affine form of Farkas lemma.

This work has been presented in the TAPAS'19 workshop [13] and will be presented at the IMPACT'20 workshop [13].

6.12. Standard-compliant Parallel SystemC simulation of Loosely-Timed Transaction Level Models

Participant: Matthieu Moy.

To face the growing complexity of System-on-Chips (SoCs) and their tight time-tomarket constraints, Virtual Prototyping (VP) tools based on SystemC/TLM must get faster while keeping accuracy. However, the Accellera SystemC reference implementation remains sequential and cannot leverage the multiple cores of modern workstations. In this paper, we present a new implementation of a parallel and standard-compliant SystemC kernel, reaching unprecedented performances. By coupling a parallel SystemC kernel and memory access monitoring, we are able to keep SystemC atomic thread evaluation while leveraging the available host cores. Evaluations show a ×19 speed-up compared to the Accellera SystemC kernel using 33 host cores reaching speeds above 2000 Million simulated Instructions Per Second (MIPS).

This work will be published at the ASP-DAC 2020 conference.

6.13. Response time analysis of dataflow applications on a many-core processor with shared-memory and network-on-chip

Participant: Matthieu Moy.

We consider hard real-time applications running on many-core processor containing several clusters of cores linked by a Network-on-Chip (NoC). Communications are done via shared memory within a cluster and through the NoC for inter-cluster communication. We adopt the time-triggered paradigm, which is well-suited for hard real-time applications, and we consider data-flow applications, where communications are explicit.

We extend the AER (Acquisition/Execution/Restitution) execution model to account for all delays and interferences linked to communications, including the interference between the NoC interface and the memory. Indeed, for NoC communications, data is first read from the initiator's local memory, then sent over the NoC, and finally written to the local memory of the target cluster. Read and write accesses to transfer data between local memories may interfere with shared-memory communication inside a cluster, and, as far as we know, previous work did not take these interferences into account.

Building on previous work on deterministic network calculus and shared memory interference analysis, our method computes a static, time-triggered schedule for an application mapped on several clusters. This schedule guarantees that deadlines are met, and therefore provides a safe upper bound to the global worst-case response time.

This work was published at RTNS 2019 [14].

6.14. Smart placement of dynamically allocated objects for heterogeneous memory

Participant: Matthieu Moy.

As part of a partnership with the CITI laboratory (Tristan Delizy's PhD, co-supervised with Guillaume Salagnac and Tanguy Risset), we worked on dynamic memory memory allocation for embedded systems with heterogeneous memory. Unlike cache-based systems, our target architecture exposes several memory banks with different performance characteristics directly to the software, without any hardware mechanism like a cache or an MMU for memory management. The software needs to chose which memory bank to use at allocation time, and cannot change this choice afterwards. We proposed a profiling-based placement policy that is shown to be near-optimal for several applications, and performs much better than naive placement policies especially for systems with a small fraction of fast memory.

This work documented as part of Tristan Delizy's Ph.D manuscript, and we plan to submit it for a journal publication in 2020.

6.15. 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 in 2018 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 presented to VMCAI'19 on Model Checking and Abstract Interpretation [5] and has received the best paper award of the conference.

6.16. 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), and rephrase a complete value and validity

In 2019, the formalisation has been rephrased in a simplier way and extented to deal with more llvm-related semantics (undefined behavior, poisoned values) and we started a collaboration with David Monniaux, Verimag, on this topic.

The paper is being rewritten and we are also writing a project on which we would extend our method to mode complex polyhedral transformations in a context of formally verified tools.

6.17. Decision results for solving Horn Clauses with arrays

Participants: Laure Gonnord, Julien Braine.

Many approaches exist for verifying programs operating on Boolean and integer values (e.g. abstract interpretation, counterexample-guided abstraction refinement using interpolants), but transposing them to array properties has been fraught with difficulties. In the context of the Phd of Julien Braine, we propose to work directly on horn clauses, because we think that it is a suitable intermediate representation for verifying programs.

Currently, two techniques strike out to infer very precise quantified invariants on arrays using Horn clauses: a quantifier instantiation method [1] and a cell abstraction method that can be rephrased on Horn clauses. However, the quantifier instantiation method is parametrized by an heuristic and finding a good heuristic is a major challenge, and the cell abstraction method uses an abstract interpretation to completely remove arrays and is limited to linear Horn clauses. We combine these two techniques. We provide an heuristic for the quantifier instantiation method of [29] by using the ideas from the cell abstraction method of [48] and discover a requirement such that, when met, the heuristic is complete, that is, there is no loss of information by using that heuristic. Furthermore, we prove that Horn clauses that come from program semantic translation verify the requirement and therefore, we have an optimal instantiation technique for program analysis.

This work is done in collaboration with David Monniaux (Verimag), coadvisor of the PhD of Julien Braine. A journal paper is currently being written for submission early 2020.

6.18. Scheduling Trees

Participants: Laure Gonnord, Paul Iannetta.

As a first step to schedule non polyhedral computation kernels, we investigated the tree datastructure. A large bibliography on tree algorithmics and complexity leds us to chose to work on balanced binary trees, for which we have designed algorithms to change their memory layout into adjacent arrays. We rephrased the classical algorithms (construction, search, destruction ...) in this setting, and implemented them in C.

The conclusion of this study is unfortunately negative : the locality gain in transforming trees into linear structures is not contrabalanced by a better cache usage, all our codes have been slowed down in the process. Our experiments are still in progress, but our hypothesis is that our trees are too sparse to be more clever that the *malloc* implementation.

A research paper will be published early 2020. This work is done in collaboration with Lionel Morel (CEA Grenoble), coadvisor of the PhD of Paul Iannetta.

6.19. Formalisation of the Polyhedral Model

Participants: Laure Gonnord, Paul Iannetta.

Last year, together with Lionel Morel (Insa/CEA) and Tomofumi Yuki (Inria, Rennes), we revisited the polyhedral model's key analysis, dependency analysis, published in a research report [44]. This year we pursued in this direction. We have now a better formalisation, and a better understanding of the expressivity and applicability.

We still have one step to study in order to be able to have a full semantic polyhedral model: properly formalise code scheduling and code generation within our semantic model.

This work is made in collaboration with Lionel Morel (CEA Grenoble) who coadvise Paul Iannetta.

6.20. Semantics diffs in LLVM

Participants: Laure Gonnord, Matthieu Moy.

Laure Gonnord and Matthieu Moy have coadvised a Master research Project ("TER") early in 2019 , whose objective was to study the LLVM LLVM compiler infrastructure with software engineering techniques in order to characterise how sequences of code analyses and transformations behave. The project has lead to a sequence of tools to evaluate experimentally how a sequence of passes influence performance.

Laure Gonnord and Matthieu Moy have, together with Sebastien Mosser, coadvised a second internship at UQAM for three months, between May and July 2019. During his internship, Sebastien Michelland has demonstrated that textual diffs are not sufficient to fully characterise the behaviours of code transformation inside compilers. He analysed llvm-diff, a tool of the distribution that makes an analysis at the intermediate representation level, and gives first hints to define a proper notion of semantic diff for this application.

For these interships two research reports have been produced.

This work was done in the context of an ongoing collaboration with Sebastien Mosser, previously in Nice, and now at UQAM. An Inria associate team was proposed for 2020-2023 on similar topics.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

CIFRE Ph.D of Julien Emmanuel with Bull/Atos, hosted by Inria. 2020-2023.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. ANR

- 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

- Laure Gonnord has regular collaborations with Fernando Pereira from UFMG, Brasil (5 publications in total, last in 2017). End of 2019 they have restarted discussions with Gabriel Radanne about proving termination properties of linux kernel BPF programs. These programs must be always terminating, and we hope to be able to prove these properties in a scalable way with the termite analyser.
- In 2018 Laure Gonnord has began a collaboration with Tobias Grösser, from ETH Zurich, and in end of 2019 this collaboration has been extended to involved more people of Verimag (David Monniaux) and CASH, in the contexte of a europeean project proposal around certified polyhedral optimisation.
- In 2019, Laure Gonnord has pursued her collaboration with Sebastien Mosser, who moved from univ Nice to UQAM (Quebec, Canada). This collaboration has led to shared interns and a "inria associate team" proposal late in october 2019, which got accepted in January 2019.
- Ludovic Henrio has regular collaborations with: University of Oslo and University of Bergen in Norway (Cristal C. Din, Einar B. Johnsen, and Silvia Lizeth. Tapia Tarifa, Violet K.I. Pun); Reiner Hähnle (TU Darmstadt), Wolfgang Ahrendt (Chalmers); Kiko Fernandez-Reyes, Dave Clarke, and Tobiaas Wrigstad (Univ Uppsala); Christoph Kessler and Ahmed Rezine (Univ of Linköping).

8.3. International Research Visitors

8.3.1. Visits of International Scientists

8.3.1.1. Internships

- Amaury Maillé, M2: from Dec 2018 to Aug 2019 (6 months in total), "Dataflow explicit futures: Formalisation and/or experimentation".
- Julien Rudeau, INSA 4A, from to Apr 2019 to Aug 2019, "Ordonnancement sous contrainte de pipeline", supervised by Christophe Alias.
- Julien Philippon, EPITECH 1A, from to Jul 2019 to Dec 2019, "Compiling dataflow models to circuits", supervised by Christophe Alias and Matthieu Moy.
- Mohamed Hadjoudj, ENS Paris-Saclay 1A, from Jun 2019 to Jul 2019, "Parallélisation sous contrainte de ressources", supervised by Christophe Alias.
- Julian Bruyat, Lyon 1 M1, part-time from January 2019 to May 2019, "Outillage pour l'étude de l'impact de l'ordre des passes de LLVM", supervised by Laure Gonnord and Matthieu Moy.
- Sebastien Michelland, ENS de Lyon M1, abroad co-supervision by Laure Gonnord and Matthieu Moy with main supervision Sebastien Mosser at UQAM (Canada), from May 2019 to July 2019 "Exploration et cartographie des passes de LLVM".

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. Member of the Organizing Committees

- Laure Gonnord animates the french compilation community since 2010 (http://compilfr.ens-lyon.fr.

9.1.2. Scientific Events: Selection

9.1.2.1. Chair of Conference Program Committees

- Laure Gonnord was PC chair and organizer of the 8th Numerical and Symbolic Abstract Domain (NSAD, satellites workshop of SAS 2019, FM2019, in Porto.)
- Ludovic Henrio was chair of the ICE'19 workshop.

9.1.2.2. Member of the Conference Program Committees

- Laure Gonnord is a PC member of **CAV** 2019 Conference on Computer-Aided Verification. She will be PC member of SBLP'20 (Brazilian Symposium on Programming Languages).
- Ludovic Henrio has been a PC member of FASE 2019, ACSD 2019, Coordination 2019, HLPP 2019, AGERE 2019, and iFM 2019
- Christophe Alias has been a PC member of COMPAS'19.

9.1.2.3. Reviewer

- Christophe Alias was reviewer for MCSoC'19.
- Matthieu Moy was reviewer for CAV'19, IFM'19, MCSOC'19

9.1.3. Journal

9.1.3.1. Reviewer - Reviewing Activities

- Ludovic Henrio was reviewer for the HLPP'19 special issue in IJPP. He is also guest editor for the special issues associated to the ICE workshop of the last 4 years.
- Christophe Alias was reviewer for ACM TOPC.
- Matthieu Moy was reviewer for Journal of Electronics.

9.1.4. Invited Talks

- Matthieu Moy presented a talk “La génération de code temps réel sur architecture many-coeur” at the annual colloquium of GDR SOC2.

9.1.5. Research Administration

- Laure Gonnord is elected member of the LIP council and the Fédération d'Informatique de Lyon council.
- Laure Gonnord is elected member of the Inria National “Commission d'Evaluation” from september 2019.

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, TD+TP, 42h, L2, UCBL
- Laure Gonnord, Operating Systems, TD+TP, 26h, L2, UCBL

- Matthieu Moy, Concurrent Programming, CM+TD+TP, 57h, L3, UCBL.
- Matthieu Moy, Recursive Programming, TD+TP, 48h, L1, UCBL.
- Matthieu Moy, Git, CM+TP: 12h, L3, UCBL.
- Amaury Maillé, Concurrent Programming, 8h TD, 14 h 30 TP, L3, UCBL.
- Paul Iannetta, ACM, TD, 32h, L3, ENS de Lyon.
- Paul Iannetta, Projet 1, TD, 36h, L3, ENS de Lyon.
- Paul Iannetta, Colloquium L3, TD, 2h, L3, ENS de Lyon
- Paul Iannetta, Jury de stage, TD, 4h, L3, ENS de Lyon
- Julien Braine, ASR, TP et TD, 32h, L3, ENS de Lyon.

Master:

- Christophe Alias, Compiler optimizations for embedded applications, CM+TD, 27h, 4A, INSA Centre Val de Loire.
- Laure Gonnord, Compilation and Program Analysis, CM, 10h, TP 8h, 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, Distributed Systems, TD 9h, M1, UCBL.
- Laure Gonnord, Graphs, Complexity, Algorithmics, M1 MEEF (CAPES Maths, prepa), CM+TD+TP+oral training, 18h, UCBL.
- Laure Gonnord, Algorithmics, Systems M1 MEEF (CAPES NSI, prepa), CM+TD, 15h, UCBL.
- Laure Gonnord, Algorithmics, Architecture, Systems, DIU EIL, CM+TD+TP, 60h, UCBL.
- Matthieu Moy, Software Engineering, CM+TD+TP, 25h, M1, UCBL.
- Matthieu Moy, Compilation and Program Analysis, TP, 16h, M1, ENS de Lyon.
- Matthieu Moy, Compilation and program transformations, TD+TP, 25h, M1, UCBL.
- Ludovic Henrio, Compilation and Program Analysis, CM, 10h; TP, 6h, M1, ENS de Lyon.
- Ludovic Henrio, Distributed Systems: an algorithmic approach, CM+TD, 3h, M2 Specialite IFI (Ingénierie et Fondements de l'Informatique), parcours CSSR, and UBINET, Université de Nice Sophia-Antipolis.
- Amaury Maillé, Networks, 2h TD, M1, UCBL.
- Paul Iannetta, Projet Intégré, TD, 14h, M1, ENS de Lyon.
- Paul Iannetta, Jury de stage, TD, 4h, M1, ENS de Lyon
- Julien Braine, APPD, TP et TD, 28h, M1, ENS de Lyon.

9.2.2. Supervision

- 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: 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).
- PhD in progress: Amaury Maillé, "Programming model to assemble compute kernels safely and efficiently: Future- based synchronization for arrays and matrices", ENS Lyon, supervised by Matthieu Moy and Ludovic Henrio. Started in October 2019, supervised by Gaëtan Hains (Huawei), Wijnand Suijlen (Huawei), Françoise Baude (UNS./I3S), Ludovic Henrio (LIP).

9.2.3. Juries

- Laure Gonnord was external jury member for the PhD of Yohan Uguen, "High-level synthesis and arithmetic optimizations" (INSA de Lyon).
- Ludovic Henrio was part of the reading commitee of Keyvan Azadbakht (Universiteit Leiden).
- Christophe Alias was a reviewer and member of the PhD committee of Hang YU (Université Grenoble-Alpes).
- Christophe Alias was *correcteur* and jury member for *concours d'admission X/ENS*.
- Laure Gonnord was jury member for the *C*oncours d'admission de l'Agrégation de Sciences Industrielles, spécialité Informatique Industrielle, in June 2019.
- Matthieu Moy was reviewer for the Ph.D of Hamza Deroui, "Étude et implantation d'algorithmes pour l'ordonnancement d'applications dataflow" (INSA Rennes).

9.3. Popularization

9.3.1. Articles and contents

- "Pourquoi créer des nouveaux langages de programmation?" Ludovic Henrio. Interstices. Janvier 2019.

9.3.2. Education

- Laure Gonnord has participated to the construction of the diploma "teaching computer science in high schools" in Lyon, and also the new preparatory cursus for the "CAPES-NSI" at Lyon1 university.

9.3.3. Interventions

- Laure Gonnord was invited by the "Association des profs de Prepa" to give her expertise about teaching computer science in the first universitary cycle. The context is the creation of a new preparatory class with a computer science minor "MPI" (Maths, Physics, Computer Science).
- Matthieu Moy presented a talk "Se lancer dans le logiciel libre quand on est étudiant (ou pas)" at "Campus du libre", Lyon.

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] C. ALIAS.*Contributions to Program Optimization and High-Level Synthesis*, ENS de Lyon, May 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02151877

Articles in International Peer-Reviewed Journal

[2] W. AHRENDT, L. HENRIO, W. OORTWIJN.*Who is to Blame? Runtime Verification of Distributed Objects with Active Monitors*, in "Electronic Proceedings in Theoretical Computer Science", August 2019, vol. 302, p. 32-46, https://arxiv.org/abs/1908.10042 - In Proceedings VORTEX 2018, arXiv:1908.09302 [*DOI :* 10.4204/EPTCS.302.3], https://hal.archives-ouvertes.fr/hal-02303148

[3] L. HENRIO, C. KESSLER, L. LI.*Leveraging access mode declarations in a model for memory consistency in heterogeneous systems*, in "Journal of Logical and Algebraic Methods in Programming", January 2020, vol. 110, p. 1-17, 100498, https://arxiv.org/abs/1910.11110 [*DOI :* 10.1016/J.JLAMP.2019.100498], https://hal.archives-ouvertes.fr/hal-02331964

International Conferences with Proceedings

[4] C. ALIAS.*fkcc: the Farkas Calculator*, in "10th Workshop on Tools for Automatic Program Analysis", Porto, Portugal, Lecture Notes in Computer Science, Springer, December 2019, https://hal.inria.fr/hal-02414224

[5] *Best Paper*
C. BALLABRIGA, J. FORGET, L. GONNORD, G. LIPARI, J. RUIZ.*Static Analysis Of Binary Code With Memory Indirections Using Polyhedra*, in "VMCAI'19 - International Conference on Verification, Model Checking, and Abstract Interpretation", Cascais, Portugal, LNCS, Springer, January 2019, vol. 11388, p. 114-135 [*DOI :* 10.1007/978-3-030-11245-5_6], https://hal.archives-ouvertes.fr/hal-01939659.

[6] S. BLIUDZE, L. HENRIO, E. MADELAINE.*Verification of concurrent design patterns with data*, in "21th International Conference on Coordination Languages and Models (COORDINATION)", Kongens Lyngby, Denmark, H. R. NIELSON, E. TUOSTO (editors), Coordination Models and Languages, Springer International Publishing, 2019, vol. LNCS-11533, p. 161-181, t Part 4: Coordination Patterns [*DOI :* 10.1007/978-3-030-22397-7_10], https://hal.archives-ouvertes.fr/hal-02143782

[7] G. BUSNOT, T. SASSOLAS, N. VENTROUX, M. MOY.*Standard-compliant Parallel SystemC simulation of Loosely-Timed Transaction Level Models*, in "25th Asia and South Pacific Design Automation Conference (ASP-DAC 2020)", Beijing, China, January 2020, https://hal.archives-ouvertes.fr/hal-02416253

[8] *Best Paper*
A. CHARIF, G. BUSNOT, R. MAMEESH, T. SASSOLAS, N. VENTROUX.*Fast Virtual Prototyping for Embedded Computing Systems Design and Exploration*, in "RAPIDO2019 - 11th Workshop on Rapid Simulation and Performance Evaluation: Methods and Tools", Valence, Spain, January 2019, p. 1-8 [*DOI :* 10.1145/3300189.3300192], https://hal.archives-ouvertes.fr/hal-02023805.

[9] M. DUPONT DE DINECHIN, M. SCHUH, M. MOY, C. MAÏZA.*Scaling Up the Memory Interference Analysis for Hard Real-Time Many-Core Systems*, in "Design, Automation and Test in Europe Conference (DATE)", Grenoble, France, March 2020, https://hal.archives-ouvertes.fr/hal-02431273

[10] K. FERNANDEZ-REYES, D. CLARKE, L. HENRIO, E. BROCH JOHNSEN, T. WRIGSTAD.*Godot: All the Benefits of Implicit and Explicit Futures*, in "ECOOP 2019 - 33rd European Conference on Object-Oriented Programming", London, Royume-Uni, Leibniz International Proceedings in Informatics (LIPIcs), 2019, p. 1-28, https://hal.archives-ouvertes.fr/hal-02302214

[11] Z. GANJEI, A. REZINE, L. HENRIO, P. ELES, Z. PENG.*On Reachability in Parameterized Phaser Programs*, in "TACAS 2019 - 25th International Conference on Tools and Algorithms for the Construction and Analysis of Systems", Prague, Czech Republic, LNCS, Springer, April 2019, vol. 11427, p. 299-315, https://arxiv.org/abs/1811.07142 [*DOI :* 10.1007/978-3-030-17462-0_17], https://hal.archives-ouvertes.fr/hal-02061520

[12] P. LECA, L. HENRIO, F. BAUDE, W. SUIJLEN.*Distributed futures for efficient data transfer between parallel processes*, in "The 35th ACM/SIGAPP Symposium On Applied Computing", Brno, Czech Republic, March 2020 [*DOI :* 10.1145/3341105.3373932], https://hal.archives-ouvertes.fr/hal-02417953

Conferences without Proceedings

[13] C. ALIAS.*Farkas Lemma made easy*, in "IMPACT 2020 - 10th International Workshop on Polyhedral Compilation Techniques", Bologna, Italy, December 2019, p. 1-6, https://hal.inria.fr/hal-02422033

[14] A. GRAILLAT, C. MAIZA, M. MOY, P. RAYMOND, B. DUPONT DE DINECHIN.*Response Time Analysis of Dataflow Applications on a Many-Core Processor with Shared-Memory and Network-on-Chip*, in "RTNS 2019 - 27th International Conference on Real-Time Networks and Systems", Toulouse, France, ACM, November 2019, p. 61-69 [*DOI :* 10.1145/3356401.3356416], https://hal.archives-ouvertes.fr/hal-02320463

Research Reports

[15] C. ALIAS, J. RUDEAU.*Pipeline-aware Scheduling of Polyhedral Process Networks*, Inria Grenoble - Rhone-Alpes, December 2019, n^{o} RR-9314, https://hal.inria.fr/hal-02414340

[16] C. ALIAS, S. THIBAULT, L. GONNORD.*A Compiler Algorithm to Guide Runtime Scheduling*, Inria Grenoble ; Inria Bordeaux - Sud-Ouest, December 2019, n^{o} RR-9315, https://hal.inria.fr/hal-02421327

Scientific Popularization

[17] L. HENRIO.*Pourquoi créer des nouveaux langages de programmation ?*, in "Interstices", January 2019, https://hal.inria.fr/hal-02008111

Other Publications

[18] C. ALIAS.*On Channel Restructuring for Complete FIFO Recovery*, November 2019, ICCD 2019 - 37th IEEE International Conference on Computer Design, Poster, https://hal.inria.fr/hal-02433318

References in notes

[19] *IEEE 1666 Standard: SystemC Language Reference Manual*, Open SystemC Initiative, 2011, http://www.accellera.org/

[20] *OSCI TLM-2.0 Language Reference Manual*, Open SystemC Initiative (OSCI), June 2008, http://www.accellera.org/downloads/standards

[21] 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

[22] C. ALIAS, A. PLESCO.*Method of Automatic Synthesis of Circuits, Device and Computer Program associated therewith*, April 2014, Patent FR1453308

[23] C. ALIAS, A. PLESCO.*Data-aware Process Networks*, Inria - Research Centre Grenoble – Rhône-Alpes, June 2015, n° RR-8735, 32, https://hal.inria.fr/hal-01158726

[24] C. ALIAS, A. PLESCO.*Optimizing Affine Control with Semantic Factorizations*, in "ACM Transactions on Architecture and Code Optimization (TACO) ", December 2017, vol. 14, n° 4, 27

[25] 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° 6, p. 113–123, http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5230810

[26] S. ANANIAN.*The Static Single Information Form*, MIT, September 1999

[27] 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

[28] 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° 2, 22 [*DOI :* 10.3390/ELECTRONICS5020022], https://hal.archives-ouvertes.fr/hal-01321055

[29] N. BJØRNER, K. MCMILLAN, A. RYBALCHENKO.*On Solving Universally Quantified Horn Clauses*, in "Static Analysis: 20th International Symposium, SAS 2013, Seattle, WA, USA, June 20-22, 2013. Proceedings", Berlin, Heidelberg, F. LOGOZZO, M. FÄHNDRICH (editors), Springer Berlin Heidelberg, 2013, p. 105–125, http://dx.doi.org/10.1007/978-3-642-38856-9_8

[30] D. CAROMEL, L. HENRIO.*A Theory of Distributed Objects*, Springer-Verlag, 2004

[31] 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

[32] 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° 5, p. 76:1–76:39, http://doi.acm.org/10.1145/3122848

[33] M. DURANTON, D. BLACK-SCHAFFER, K. DE BOSSCHERE, J. MAEBE.*The HIPEAC VISION FOR ADVANCED COMPUTING IN HORIZON 2020, https://www.hipeac.net/v13*, 2013, https://www.hipeac.net/v13

[34] P. FEAUTRIER.*Scalable and Structured Scheduling*, in "International Journal of Parallel Programming", October 2006, vol. 34, n° 5, p. 459–487

[35] P. FEAUTRIER.*Dataflow analysis of array and scalar references*, in "International Journal of Parallel Programming", 1991, vol. 20, n° 1, p. 23–53

[36] 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

[37] K. FERNANDEZ-REYES, D. CLARKE, E. CASTEGREN, H.-P. VO.*Forward to a Promising Future*, in "Conference proceedings COORDINATION 2018", 2018

[38] 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

[39] 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

[40] 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

[41] 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

[42] 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

[43] L. HENRIO.*Data-flow Explicit Futures*, I3S, Université Côte d'Azur, April 2018, https://hal.archives-ouvertes.fr/hal-01758734

[44] P. IANNETTA, L. GONNORD, L. MOREL, T. YUKI.*Semantic Array Dataflow Analysis*, Inria Grenoble Rhône-Alpes, December 2018, n^o RR-9232, p. 1-22, https://hal.archives-ouvertes.fr/hal-01954396

[45] G. KAHN.*The semantics of a simple language for parallel programming*, in "Information processing", North-Holland, 1974

[46] M. MAALEJ, 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

[47] M. MAALEJ KAMMOUN.*Low-cost memory analyses for efficient compilers*, Université Lyon 1, 2017, Thèse de doctorat, Université Lyon1, http://www.theses.fr/2017LYSE1167

[48] D. MONNIAUX, L. GONNORD.*Cell morphing: from array programs to array-free Horn clauses*, in "23rd Static Analysis Symposium (SAS 2016)", Edimbourg, United Kingdom, X. RIVAL (editor), Static Analysis Symposium, September 2016, https://hal.archives-ouvertes.fr/hal-01206882

[49] 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

[50] V. PAISANTE, M. MAALEJ, 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

[51] 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

[52] P. QUINTON.*Automatic synthesis of systolic arrays from uniform recurrent equations*, in "ACM SIGARCH Computer Architecture News", 1984, vol. 12, n° 3, p. 208–214

[53] H. RIHANI, M. MOY, C. MAÏZA, 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

[54] 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

[55] W. THIES.*Language and compiler support for stream programs*, Massachusetts Institute of Technology, 2009

[56] A. TURJAN.*Compiling Nested Loop Programs to Process Networks*, Universiteit Leiden, 2007

[57] 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

[58] S. VERDOOLAEGE.*Polyhedral Process Networks*, Handbook of Signal Processing Systems, Springer, 2010, p. 931–965

[59] S. WILLIAMS, A. WATERMAN, D. PATTERSON.*Roofline: an insightful visual performance model for multicore architectures*, in "Communications of the ACM", 2009, vol. 52, n° 4, p. 65–76

[60] 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]
Ozgur Erkent [Inria, Starting Research Position, from Aug 2019]
David Sierra Gonzalez [Inria, Starting Research Position, from Jun 2019]

Faculty Members

Olivier Simonin [Team leader, INSA Lyon, Professor, HDR]
Jilles Dibangoye [INSA Lyon, Associate Professor]
Anne Spalanzani [Univ Grenoble Alpes, Associate Professor, HDR]
Christian Wolf [INSA Lyon, Associate Professor, delegation until Aug 2019, HDR]
Laetitia Matignon [Univ Claude Bernard, Associate Professor, delegation until Aug 2019]
Samba Ndojh Ndiaye [Univ Claude Bernard, Associate Professor, delegation from Sep 2019]

External Collaborators

Fabrice Jumel [École supérieure de chimie physique électronique de Lyon]
Jacques Saraydaryan [École supérieure de chimie physique électronique de Lyon]

Technical Staff

Rabbia Asghar [Inria, Engineer]
Stephane d'Alu [INSA Lyon, Engineer]
Jean-Alix David [Inria, Engineer]
Thomas Genevois [Inria, Engineer]
Vincent Le Doze [Inria, Engineer]
Jerome Lussereau [Inria, Engineer]
Anshul Paigwar [Inria, Engineer]
Alessandro Renzaglia [Inria, Engineer]
Lukas Rummelhard [Inria, Engineer]
David Sierra Gonzalez [Inria, Engineer, until May 2019]

PhD Students

Edward Beeching [Inria, PhD Student]
Alexandre Bonnefond [Inria, PhD Student, from Nov 2019]
Guillaume Bono [INSA Lyon, PhD Student]
Remi Cambuzat [Région, PhD Student, until Jun 2019]
Niranjan Deshpande [Inria, PhD Student]
Maria Kabtoul [Inria, PhD Student]
Mihai-Ioan Popescu [INSA Lyon, PhD Student]
Manon Predhumeau [Univ Grenoble Alpes, PhD Student]
Benoit Renault [INSA Lyon, PhD Student]
Luiz Serafim Guardini [Renault, PhD Student, granted by CIFRE]
Pavan Vasishta [Inria, PhD Student, until Jun 2019]

Post-Doctoral Fellows

Lotfi Abdi [Inria, Post-Doctoral Fellow, from Mar 2019]
Mario Andrei Garzon Oviedo [Inria, Post-Doctoral Fellow, until Oct 2019]

Andres Gomez Hernandez [Inria, Post-Doctoral Fellow]
Mohamad Hobballah [INSA Lyon, Post-Doctoral Fellow]

Administrative Assistants

Sophie Karsai [Inria, Administrative Assistant, until Jul 2019]
Anouchka Ronceray [Inria, Administrative Assistant, from Jul 2019]

2. Overall Objectives

2.1. Origin of the project

Chroma is a bi-localized project-team at Inria Grenoble Rhône-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] 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, HDR) 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. 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 in 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 Section 9.3.

[0] Karlsruhe Institut fur Technologie
[0] International Center of Excellence in Intelligent Robotics and Automation Research.

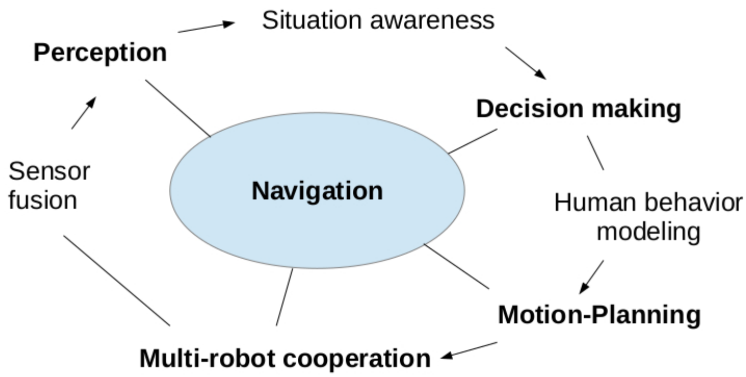


Figure 1. Research themes of the team and their relation

3.2. Perception and Situation Awareness

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 [65] 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 [89] 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 [96]) 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 [84]. 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 [76] 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 short-comings 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 [78] (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 [80], which extends a heuristic solution derived by the team of Prof. Antonio Bicchi [60] 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 [70]. Recently, this sensor fusion problem has been successfully addressed by enforcing observability constraints [74] and by using optimization-based approaches [77]. 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 [79] and [81]. 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

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 [69], the MIT [57], the Autonomous Intelligent Systems lab in Freiburg [61], or the LAAS [85]. 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.

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 [93] (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).

[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.
[0] known environment without uncertainty

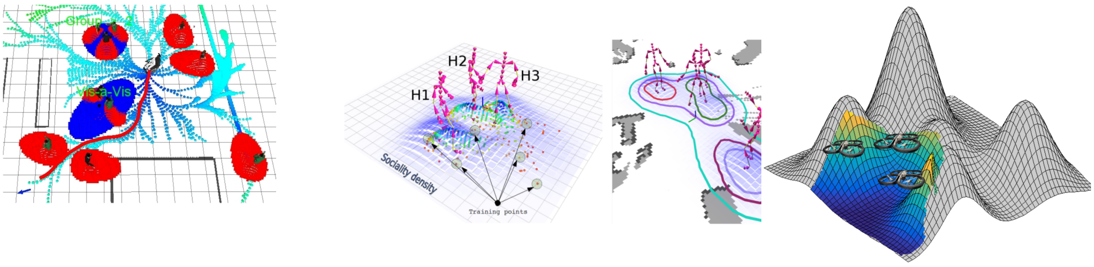


Figure 3. Illustrations of a. the Risk-RRT planning b. The human interaction space model c. Multi-UAV 3D coverage and exploration.

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 7.5.1).

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 [71], [93], [100]). 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.

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.

[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.

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 [92], 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) [67]. 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 [97] the interest of such an approach for driverless vehicle traffic optimization.

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 OR [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. For instance we explore online stochastic-optimization planning to deal with multi-robot coverage/exploration of 3D environments, see Fig 3.c and [42].

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 tens to thousands 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.

[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.

[0]Operations Research

[0]see for instance the first International Workshop on Cloud and Robotics, 2016.

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 [86]).

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 [86] (wireless communication protocols) and with Dynamid team [63] (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.

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).

[0] Advanced Driver Assistance Systems

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

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 (Dynaflock Inria/DGA project, Inria ADT CORDES [0]) or mobile robots (COMODYS FIL project [82]).

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 (see Figure 17 for illustration). 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

- Success for European H2020 ICT Robotics project application 'BugWright2' (9M€), led by C. Pradalier (CNRS, GeorgiaTech Metz). O. Simonin leads the multi-robot systems Work-Package (funding for Chroma & Agora teams 600K€). Domain : Autonomous Robotic Inspection and Maintenance on Ship Hulls and Storage Tanks.
- Success for several ANR project applications in the field of Artificial Intelligence :
  - ANR JCJC 'PLASMA' led by J. Dibangoye (250K€)

[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
[0] Coordination d'une Flotte de Drones Connectés pour la Cartographie 3D d'édifices, led by O. Simonin.
[0] Personnaly assisted Living

  - ANR ’DELICIO’ led by C. Wolf (510 K€), Chroma is partner.
  - AI Chair led by C. Wolf (520 K€), Chroma is partner (O. Simonin, J. Dibangoye).
- Success for several project applications in the field of Autonomous Vehicles : 2 multi-annual R&D projects with Toyota Motor Europe, a PSPC project ES3CAP leaded by Kalray (3 years), and an EU ECSEL project CPS4EU (3 years).
- Our team LyonTech obtained the 3rd place at the Robocup@Home Pepper league in the 2019 RoboCup competition organized in Sydney (July).
- O. Simonin co-chaired with F. Charpillet (Inria Nancy) the JNRR’2019 bi-annual conference, gathering the French Robotic community (GDR Robotique) ($\sim$ 200 pers.).
- New book by A. Martinelli : "Observability: A new theory based on the group of invariance". To be edited by SIAM on year 2020.
- Exploitation Licenses of CMCDOT have respectively been sold to Toyota and to a French company in the field autonomous vehicles (confidential), with an engineer support for the related transfer of technology.

5.1.1. Awards

BEST PAPERS AWARDS :

[43]

J. SARAYDARYAN, R. LEBER, F. JUMEL.*People management framework using a 2D camera for human-robot social interactions*, in "RoboCup 2019 - 23rd Annual RoboCup International Symposium", Sydney, Australia, Robocup 2019: Robot World Cup XXIII, July 2019, p. 1-13, https://hal.archives-ouvertes.fr/hal-02318916

6. New Software and Platforms

6.1. 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

6.2. 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

6.3. 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

6.4. 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

6.5. 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

6.6. 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/

6.7. 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/

6.8. 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

6.9. 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

6.10. cuda_US_grid

KEYWORDS: Robotics - Environment perception

FUNCTIONAL DESCRIPTION: This module generates occupation grids from data generated by an ultrasonic range sensor. 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: Christian Laugier and Thomas Genevois
- Partner: CEA
- Contact: Christian Laugier

6.11. Embedded Perception

FUNCTIONAL DESCRIPTION: The method for computing occupancy grids from a stereoscopic sensor, developped in the e-motion team, has been implemented on GPU, using NVIDIA CUDA. This allows a real time implementation and an online processing within the Lexus experimental platform.

- Participants: Amaury Nègre, Christian Laugier and Mathias Perrollaz
- Contact: Christian Laugier

6.12. spank

Swarm Protocol And Navigation Kontrol

KEYWORD: Protocoles

FUNCTIONAL DESCRIPTION: Communication and distance measurement in an uav swarm

- Contact: Stéphane d'Alu
- URL: https://gitlab.inria.fr/dalu/spank

6.13. S-NAMO-SIM

S-NAMO Simulator

KEYWORDS: Simulation - Navigation - Robotics - Planning

FUNCTIONAL DESCRIPTION: 2D Simulator of NAMO algorithms (Navigation Among Movable Obstacles) ROS compatible

RELEASE FUNCTIONAL DESCRIPTION: Creation

- Contact: Benoit Renault

7. New Results

7.1. 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.

7.1.1. Visual-inertial structure from motion

Participant: Agostino Martinelli.

We have continued our study on the visual inertial sensor fusion problem in the cooperative case, with a special focus on the case of two agents. During this year, we have carried out an exhaustive analysis of all the singularities and minimal cases of this cooperative sensor fusion problem. As in the case of a single agent and in the case of other computer vision problems, the key of the analysis is the establishment of an equivalence between the cooperative visual-inertial sensor fusion problem and a Polynomial Equation System (PES). In the case of a single agent, the PES consists of linear equations and a single polynomial of second degree. In the case of two agents, the number of second degree equations becomes three and, also in this case, a complete analytic solution can be obtained [19], [20]. The power of the analytic solution is twofold. From one side, it allows us to determine the state without the need of an initialization. From another side, it provides fundamental insights into all the structural properties of the problem. The research of this year has focused on this latter issue. Specifically, we have obtained all the minimal cases and singularities depending on the number of camera images and the relative trajectory between the agents. The problem, when non singular, can have up to eight distinct solutions. The usefulness of this analysis has also been illustrated with simulations. In particular, we have quantitatively obtained how the performance of the state estimation worsens near a singularity. The results of this research will be published by the Robotics and Automation Letter (RA-L) journal [18].

7.1.2. 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. So far, I only published the solution for systems characterized by driftless dynamics. In particular, this solution was published as a full paper on the IEEE Transaction on Automatic Control [17]. In December 2018, I was invited by the Society for Industrial and Applied Mathematics (SIAM) to write a book with the general solution. This has been the main work of this year. Since this general solution is based on tensorial calculus (Ricci algebra) and many mathematics procedures and tricks borrowed from theoretical physics, the scope of book has gone much more beyond the presentation of the solution. Basically, by writing this book, I've obtained a new theory of observability.

The current theory of nonlinear observability, does not capture/exploit the key features that are intimately related to the concept of observability. This results in two important limitations:

- The theory, although simple and based on elementary mathematics, can be sometimes burdensome with the risk of easily loosing the meaning of the results and losing the meaning of their assumptions.
- More complex observability problems (e.g., the unknown input observability problem to which this book provides the complete analytic solution) remained unsolved for half a century.

The key to overcome the two above limitations, consists in building a new theory of observability that accounts for the **group of invariance that is inherent to the concept of observability**. This is the typical manner the research in physics has always proceeded. To this regard, I wish to emphasize that the derivation of the basic equations of any physics theory (e.g., the General Relativity, the Yang Mills theory, the Quantum Chromodynamics) starts precisely from the characterization of the group of invariance of the theory.

One of the major novelties introduced by this book is the characterization of the group of invariance of observability and, regarding the case of unknown inputs, the characterization of a subgroup that was called the *Simultaneous Unknown Input Output transformations' group*.

In summary, the book provides several novelties with respect to the existing literature in control theory. Specifically, the reader will learn the following:

- The solution of two open problems in control theory (the book provides separately the solution and the derivation), which are:
  - The extension of the observability rank condition to nonlinear systems driven by also unknown inputs.
  - The extension of the observability rank condition to nonlinear, time-variant systems (both in presence and in absence of unknown inputs)
- A new and more palatable derivation of the existing results in nonlinear observability.
- A new manner of approaching scientific and technological problems, borrowed from theoretical physics (a chapter summarizes in a very intuitive and quick manner the basic mathematics, which includes tensorial calculus).
- A new manner of dealing with the variable *time* in system theory, which is obtained by introducing a new framework, which was called the *chronospace*.

I believe this book could be an opportunity for control and information theory communities to borrow basic mathematics, tricks, types of reasoning from theoretical physics to revisit many aspects of control and information theory.

7.2. Bayesian Perception

Participants: Christian Laugier, Lukas Rummelhard, Jean-Alix David, Jerome Lussereau, Thomas Genevois, Nicolas Turro [SED], Rabbia Asghar, Mario Garzon.

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 and the associated risks of collision, 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. In 2019, this framework, and the corresponding software, has continued to be the core of many important industrial partnerships and academic contributions, and to be the subject of important developments, both in terms of research and engineering. Some of those recent evolutions are detailed below.

In 2019, the new results have been presented in several invited talks given in some of the major international conferences of the domain [30], [28], [26], [29], [27].

7.2.1. Conditional Monte Carlo Dense Occupancy Tracker (CMCDOT) Framework

Participants: Lukas Rummelhard, Jerome Lussereau, Jean-Alix David, Thomas Genevois, Christian Laugier, Nicolas Turro [SED].

Important developments in the CMCDOT (Fig. 5), in terms of calculation methods and fundamental equations, were introduced and tested. These developments are currently being patented, and will then be used for academic publications. These changes lead to a much higher update frequency, greater flexibility in the management of transitions between states (and therefore a better system reactivity), as well as to the management of a high variability in sensor frequencies (for each sensor over time, and in the set of sensors). The changes include:

- Grid fusion: a new fusion of occupancy grids, enhanced with “unknown” variables, has been developed and implemented. The role of unknown variables has also been enlarged. Currently being patented, it should be the subject of an upcoming paper.
- Ground Estimator: a new method of occupancy grid generation, more accurately taking into account the height of each laser beam, has been developed. Currently being patented, it should be the subject of an upcoming paper.
- 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 2019, on the optimization of the software and methods to be embedded on low energy consumption embedded boards (Nvidia Jetson TX1, TX2, AGX Xavier).

Figure 5. CMCDOT results

7.2.2. Multimodal Bayesian perception

Participants: Thomas Genevois, Christian Laugier.

The objective is to extend the concept of Bayesian Perception to the fusion of multiple sensing modalities (including raw data provided by low cost sensors). In 2019, we have developed and implemented a Bayesian model dedicated to ultrasonic range sensors. For any given measurement provided by the sensor, the model computes the occupancy probability in a 2 dimensional grid around the sensor. This computation takes into account the accuracy and the possibility to “miss” an object. Thanks to various parameters, this model has been applied to the sensors of our Renault Zoe demonstrator and to the low cost sensors of our light vehicle demonstrator (flycar).

Fig. 6.a shows an example, developed and implemented on our light vehicle demonstrator. In this example, the perception is relying on 1 lidar and 5 ultrasonic range sensors. An occupancy grid is generated for each sensor. Then they are fused in a single occupancy grid which is filtered using the CMCDOT approach.

7.2.3. Embedding deep learning for semantics

Participants: Thomas Genevois, Christian Laugier.

The objective is to improve embedded Bayesian Perception outputs in our experimental vehicle platforms (Renault Zoe and Flycar), by adding semantics obtained using RGB images and embedded deep learning approaches. In 2019, we have tested several networks for road scene semantic segmentation and implemented two of them in our vehicle platforms:

- LaneNet is a network that provides lane markings detection in road scenarios [83]
- KittiSeg is a network that performs the segmentation of roads [95]

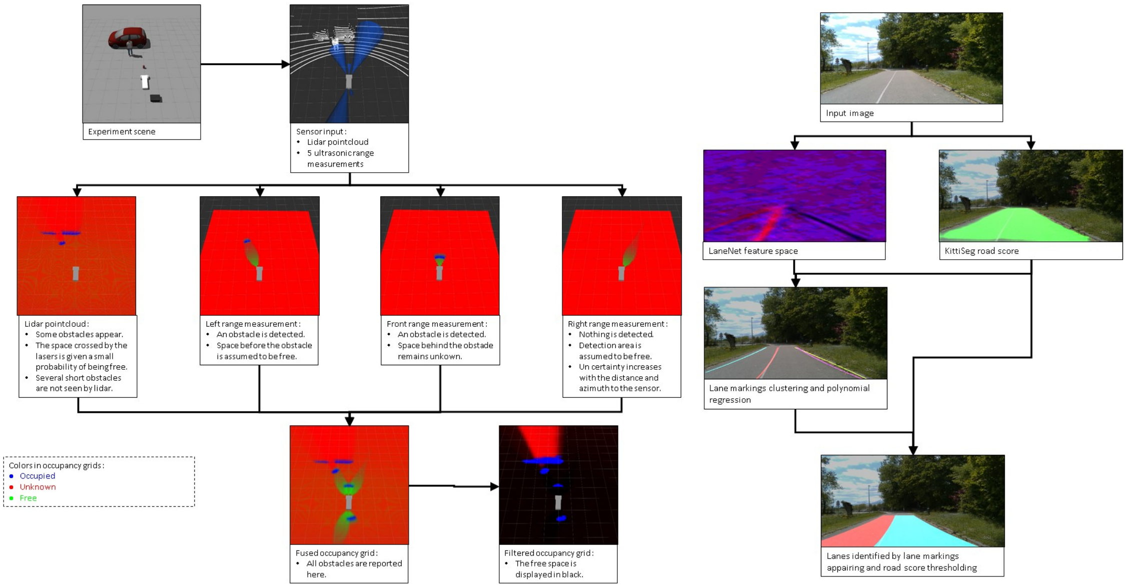


Figure 6. a. Example of multimodal perception, implemented both in simulation and on an actual vehicle demonstrator. b. Combining LaneNet and KittiSeg into a common lane recognition tool.

Therefore, Kittiseg is used to identify the shape of the road within an RGB image and LaneNet is used to identify the lane markings that divide the road into lanes. Upon this, we have developed a post-processing technique based on filtering, clustering and regression (Fig. 6.b). This post-processing technique makes the whole system far more robust and allows to express the lanes in a simple way (polynomial curves in the vehicle's base frame).

Since the objective is to embed semantic segmentation tools on our vehicle platforms, an emphasis has been put on the related embedded constraints (in particular strong real time constraints and appropriate light hardware such as the NVIDIA Jetson TX2). However, the networks LaneNet and KittiSeg have not been optimized neither for real-time inference nor for inference on light hardware. This is why we had to propose an approach for adapting these networks to our strong embedded constraints. This approach relies on the following three main steps: Reducing the resolution of the input image, Removing all computations not needed at inference (some parts of the networks are only needed in the learning phase), Adapting the network's shape to the hardware.

These optimization steps have been followed for KittiSeg and LaneNet networks. The improvement is obvious. Namely, for the network LaneNet the initial inference needed 334 operations while, after optimization, it needs only 10 operations. The inference initially runs at 0.3Hz on our board NVIDIA Jetson TX2 while, after optimization, it runs at 10Hz. Also the memory needed for inference is divided by two due to the optimization.

7.2.4. Online map-relative localization

Participants: Rabbia Asghar, Mario Garzon, Jerome Lussereau, Christian Laugier.

Localization is one of the key components of the system architecture of autonomous driving and Advanced Driver Assistance Systems (ADAS). Accurate localization is crucial to reliable vehicle navigation and acts as a prerequisite for the planning and control of autonomous vehicles. Offline digital maps are readily available especially in urban scenarios and they play an important role in the field of autonomous vehicles and ADAS.

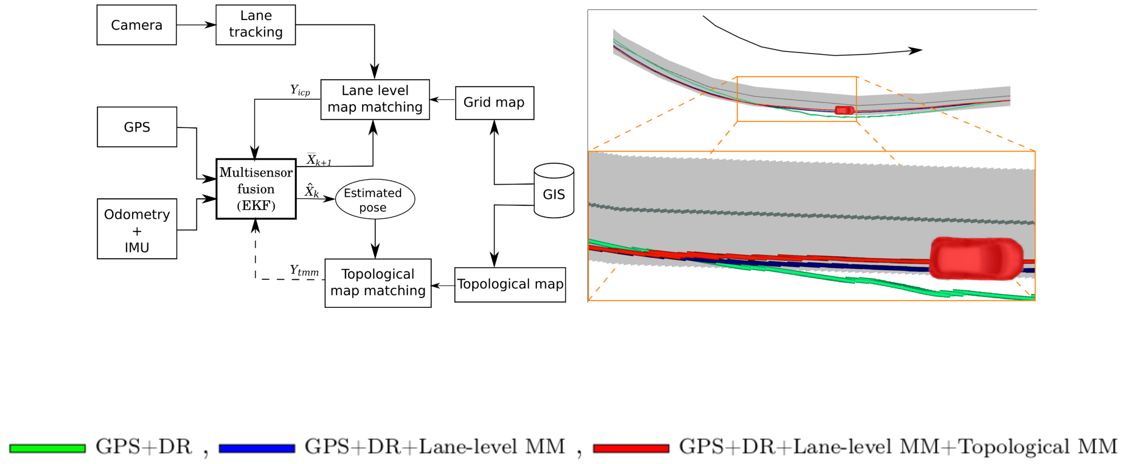


Figure 7. (a) Overview of the map relative localization approach. (b) Estimated pose of the vehicle using three different localization approaches on a curved section of road. The vehicle is provided as a reference where the estimate vehicle pose is just at the curb of the road. Black arrow represents direction of travel.

In this framework, we have developed a novel approach for online vehicle localization in a digital map. Two distinct map matching algorithms are proposed:

- Iterative Closest Point (ICP) based lane level map matching (Ll.MM) is performed with visual lane tracker and grid map.
- Decision-rule (DR) based approach is used to perform topological map matching (T. MM).

Results of both map matching algorithms are fused together with GPS and dead reckoning using Extended Kalman Filter to estimate the vehicle's pose relative to the map (see Fig. 7). The approach has been validated on real life conditions on a road-equipped vehicle using a readily available, open source map. Detailed analysis of the experimental results show improved localization using the two aforementioned map matching algorithms (see [50] for more details).
This research work has been carried out in the scope of Project Tornado. A paper on this work was submitted to ICRA2020 and is awaiting review.

7.2.5. System Validation using Simulation and Formal Methods

Participants: Alessandro Renzaglia, Anshul Paigwar, Mathieu Barbier, Philippe Ledent [Chroma/Convecs], Radu Mateescu [Convecs], Christian Laugier, Eduard Baranov [Tamis], Axel Legay [Tamis].

Since 2017, we are working on novel approaches, tools and experimental methodologies with the objective of validating probabilistic perception-based algorithms in the context of autonomous driving. To achieve this goal, a first approach based on Statistical Model Checking (SMC) has been mainly studied in the scope of the European project Enable-S3 and in collaboration with the Inria team Tamis. In this work, we studied the behavior of specifically defined Key Performance Indicators (KPIs), expressed as temporal properties depending on a set of identified metrics, during a large number of simulations via a statistical model checker. As a result, we obtained an evaluation of the probability for the system to meet the KPIs. In particular, we show how this method can be applied to two different subsystems of an autonomous vehicle: a perception system and a decision-making approach for intersection crossing [31]. A more detailed description of the validation scheme for the decision-making approach has been also presented in [49]. This work has been developed in the framework of M. Barbier's PhD thesis, which has been defended in December 2019 [11]. In parallel, in [38], we also proposed a methodology based on a combination of simulation, formal verification, and statistical analysis to validate the collision-risk assessment generated by the Conditional Monte Carlo Dense Occupancy Tracker (CMCDOT), a probabilistic perception system developed in the team. This second work is in collaboration with the Inria team Convecs.

In both cases, the validation methodology relies on the simulation of realistic scenarios generated by using the CARLA simulator [0]. CARLA simulation environment consists of complex urban layouts, buildings and vehicles rendered in high quality, allowing for a realistic representation of real-world scenarios. The ego-vehicle and its sensors, as well as other moving vehicles can be so configured in the simulation to match with the actual system. In order to be able to efficiently generate a large number of execution traces, we have perfected a parameter-based approach which streamlines the process through which the dimensions and initial position and velocity of non-ego vehicles are specified.

We also collected several traces in real experiments by imitating the collision of the ego-vehicle (equipped Renault Zoe) with a pedestrian (by using a mannequin) and with another vehicle (by throwing a big ball). Since it is unfeasible to generate with real experiments a statistically significant number of traces, we focused our analysis on studying how close the simulation traces are to these real experiments by comparing analogous scenarios. These results have been recently submitted to ICRA and are currently under review [0].

7.2.6. Industrial partners and technological transfer

Participants: Christian Laugier, Lukas Rummelhard, Jerome Lussereau, Jean-Alix David, Thomas Genevois.

In 2019, a significant amount of work has been done with the objective to transfer our Bayesian Perception technologies to industrial companies. In a first step, we have developed a new version of CMCDOT based on a clear split of ROS middle-ware code and of GroundEstimator/CMCDOT CUDA code. This allowed us to develop a new version of CMCDOT using the RTMAPS middleware for Toyota Motor Europe. It also allowed us to transfer the CMCDOT technology to some other industrial partners (confidential), in the scope of the project "Security of Autonomous Vehicle" of IRT Nanoelec. Within the IRT Nanoelec framework, we also developed a new "light urban autonomous vehicle" operating using an appropriate version of the CMCDOT and having the capability to navigate with low cost sensors. A first demo of the prototype of this light vehicle has been shown in December 2019, and a start-up project (named Starlink) is currently in incubation.

7.2.7. Autonomous vehicle demonstrations

Participants: Lukas Rummelhard, Jean-Alix David, Thomas Genevois, Jerome Lussereau, Christian Laugier.

In 2019, Chroma has participated to two main public demonstrations:

Figure 8. Demonstration at the IV2019 conference : a) track b) demonstration event.

[0]http://carla.org/

[0]A. Paigwar, E. Baranov, A. Renzaglia, C. Laugier and A. Legay, "Probabilistic Collision Risk Estimation for Autonomous Driving: Validation via Statistical Model Checking", *submitted to* IEEE ICRA20.

- **IEEE IV 2019 Conference** (Versailles Satory, June 2019): A one day public demonstration of our Autonomous Vehicle Embedded Perception System has been done using our Renault Zoe platform. Fig. 8.a and 8.b show, respectively, the demonstration track (yellow track) and our booth & demonstration vehicle. During the day, we regularly drove people in our Zoe platform for demonstrating how the perception system was working in various situations.
- **FUI Tornado mid-project event** (Rambouillet, September 2019): This one week event included public demonstrations and several open-road tests. During this week, we tested the technologies developed in the scope of the project and we made public and official (for persons from the French Ministries) demonstrations with our Renault Zoe vehicle.

7.3. Situation Awareness & Decision-making for Autonomous Vehicles

Participants: Ozgur Erkent, Christian Wolf, Christian Laugier, Olivier Simonin, Mathieu Barbier, David Sierra-Gonzalez, Jilles Dibangoye, Mario Garzon, Anshul Paigwar, Manuel Alejandro Diaz-Zapata, Victor Romero-Cano [Universidad Autónoma de Occidente, Cali, Colombia], Andrés E. Gómez H., Luiz Serafim-Guardini.

In this section, we include all the novel results in the domains of perception, motion prediction and decision-making for autonomous vehicles. In 2019, these results have also been presented in several invited talks given in some of the major international conferences of the domain [30], [28], [26], [29], [27].

7.3.1. End-to-End Learning of Semantic Grid Estimation Deep Neural Network with Occupancy Grids

Participants: Özgür Erkent, Christian Wolf, Christian Laugier.

Semantic grid is a spatial 2D map of the environment around an autonomous vehicle consisting of cells which represent the semantic information of the corresponding region such as *car, road, vegetation, bikes, etc.*. It consists of an integration of an occupancy grid, which computes the grid states with a Bayesian filter approach, and semantic segmentation information from monocular RGB images, which is obtained with a deep neural network. The network fuses the information and can be trained in an end-to-end manner. The output of the neural network is refined with a conditional random field [15]. The contributions of the study are:

- An end-to-end trainable deep learning method to obtain the semantic grids by integrating the occupancy grids obtained by a Bayesian filter approach and the semantically segmented images by using the monocular RGB images of the environment.
- Grid refinement with conditional random fields (CRFs) on the output of the deep network.
- A comparison of the performances of three different semantic segmentation network architectures in the proposed end-to-end trainable setting.

The proposed method is tested in various datasets (KITTI dataset, Inria-Chroma dataset and SYNTHIA) and different deep neural network architectures are compared (Fig. 9).

7.3.2. Attentional PointNet for 3D object detection in Point Cloud

Participants: Anshul Paigwar, Özgür Erkent, Christian Wolf, Christian Laugier.

Accurate detection of objects in 3D point clouds is a central problem for autonomous navigation. Approaches like PointNet [87] that directly operate on sparse point data have shown good accuracy in the classification of single 3D objects. However, LiDAR sensors on Autonomous Vehicles generate a large scale point cloud. Real-time object detection in such a cluttered environment still remains a challenge. In this study, we propose Attentional PointNet, which is a novel end-to-end trainable deep architecture for object detection in point clouds (Fig. 10). We extend the theory of visual attention mechanisms to 3D point clouds and introduce a new recurrent 3D Localization Network module. Rather than processing the whole point cloud, the network learns where to look (finding regions of interest), which significantly reduces the number of points to be processed and inference time. Evaluation on KITTI [72] car detection benchmark shows that our Attentional PointNet achieves comparable results with the *state-of-the-art* LiDAR-based 3D detection methods in detection (Fig. 11) and speed.

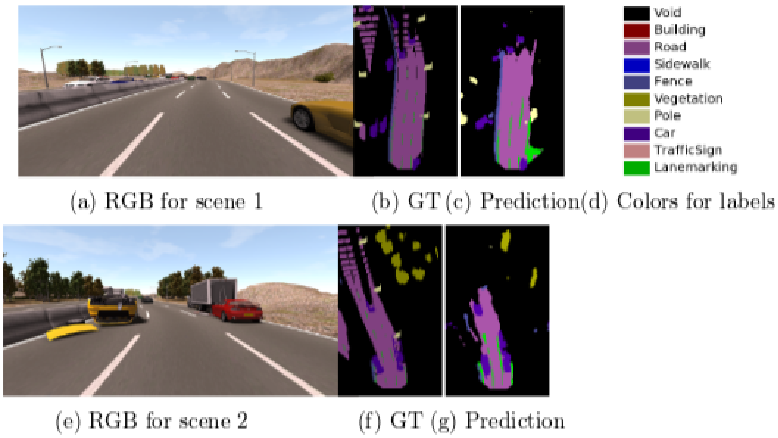


Figure 9. Two scenes with RGB image, ground truth (GT), semantic and segmentation predictions from SYNTHIA dataset.

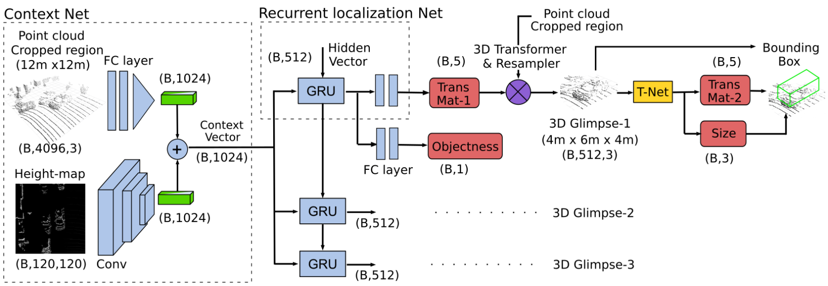


*Figure 10. **Attentional PointNet Architecture**: Given the point cloud and the corresponding height map, network sequentially regress parameters of a 3D Transformation matrix representing pose of a fixed size 3D glimpse. A modified PointNet (T-Net) then estimates another 3D transformation matrix and size representing the 3D bounding box of the object inside the glimpse. Where B is the batch size.*

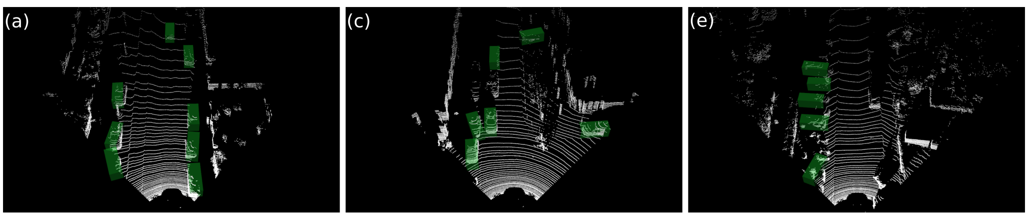


*Figure 11. Visualizations of **Attentional PointNet** results on KITTI dataset for the car category shows model's ability to detect multiple objects in cluttered environments*

This work has been published in CVPR 2019 - Workshop for Autonomous Driving, Long Beach, California, USA [39].

7.3.3. Panoptic Segmentation

Participants: Manuel Alejandro Diaz-Zapata, Victor Romero-Cano [Universidad Autónoma de Occidente, Cali, Colombia], Özgür Erkent, Christian Laugier.

This work has been accomplished during the internship of Manuel Alejandro Diaz Zapata at Inria-Rhone Alpes under supervision of Ozgur Erkent, Victor Romero-Cano and Christian Laugier at Chroma Project Team. Manuel Alejandro Diaz Zapata was a student of Mechatronic Engineering at Universidad Autónoma de Occidente, Colombia during his internship [52].

Semantic segmentation labels an image at the pixel level, where amorphous regions of similar texture or material such as grass, sky or road are given a label depending on the class. Instance segmentation focuses on countable objects such as people, cars or animals by delimiting them in the image using bounding boxes or a segmentation mask. To reduce the gap between the methods used to detect uncountable objects, and things or countable objects, panoptic segmentation has been proposed [75].

We propose a model consisting of three modules: the semantic segmentation module, the instance segmentation module and the panoptic head (Fig.12). Here the semantic segmentation is done by the MobileNetV2 [90] and the instance segmentation is done by Mask R-CNN [73]. The outputs of both networks are joint by the Panoptic Head. The results are provided on two different datasets.

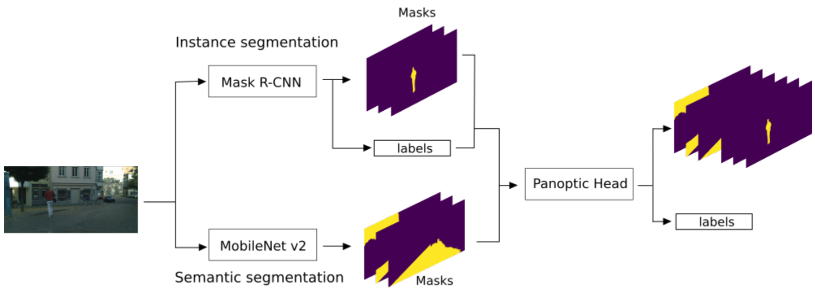


Figure 12. Proposed model for panoptic segmentation.

7.3.4. Recognition of dynamic objects for risk assessment

Participants: Andrés E. Gómez H., Özgür Erkent, Christian Laugier.

The Conditional Monte Carlo Dense Occupancy Tracker (*CMCDOT*) framework has proved its accuracy in describing *2D* spatial maps for the Zoe platform. However, this method nowadays cannot recognize the objects in the surrounding. Specifically, the identification of dynamical objects will let us consider different methodologies of risk assessment. This procedure can be possible, through the fusion of RGB and dynamical occupancy grids information.

In the fusion process development, we took into consideration the following steps: *i)* selection of a deep-learning approach, *ii)* development of the projective transformations and *iii)* joining the sub-results. In each step, we used real data from the Zoe platform. In the first step, the *YoloV3* was the deep-learning approach chosen for its accuracy and time performance. In the second step, the projective transformations let us compute the representation of the dynamical points obtained from the occupancy grid plane (i.e., *CMCDOT* framework) in the image plane. Finally, in the third step, we compare the result obtained between the last two-step to identify the dynamic objects around the Zoe platform.

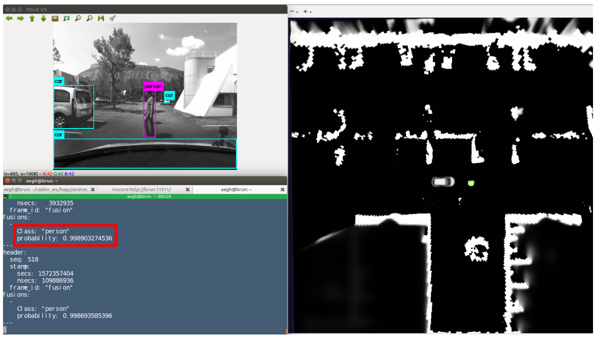


Figure 13. Identification of a pedestrian moving in front of the Zoe platform using the fusion process proposed.

Figure 13 lets us observe the inputs needed for the fusion process and its result.

The work described in this section was done during 2019, inside the activities developed for the Star project. The future work in our project aims to consider the velocity and direction of the dynamic points to define and implement risk behavior functions.

7.3.5. Driving behavior assessment and anomaly detection for intelligent vehicles

Participants: Chule Yang [Nanyang Technological University], Alessandro Renzaglia, Anshul Paigwar, Christian Laugier, Danwei Wang [Nanyang Technological University].

Ensuring safety of both traffic participants and passengers is an important challenge for rapidly growing autonomous vehicle technology. To this purpose, intelligent vehicles not only have to drive safe but must be able to safeguard themselves from other abnormally driving vehicles and avoid potential collisions [56]. Anomaly detection is one of the essential abilities in behavior analysis, which can be used to infer the moving intention of other vehicles and provide evidence for collision risk assessment. In this work, we propose a behavior analysis method based on Hidden Markov Model (HMM) to assess the driving behavior of vehicles on the road and detect anomalous moments. The algorithm uses the real-time velocity and position of the surrounding vehicles provided by the Conditional Monte Carlo Dense Occupancy Tracker (CMCDOT) [89] framework. The movement of each vehicle can be classified into several observation states, namely, Approaching, Braking, Lane Changing, and Lane Keeping. Finally, by chaining these observation states using a Markov model, the abnormality of driving behavior can be inferred into Normal, Attention, and Risk. We perform experiments using CARLA simulator environment to simulate abnormal driving behaviors as shown in Fig. 14, and we provide results showing the successful detection of abnormal situations.

This work has been published in IEEE CIS-RAM 2019, Bangkok [45].

7.3.6. Human-Like Decision-Making for Automated Driving in Highways

Participants: David Sierra-Gonzalez, Mario Garzon, Jilles Dibangoye, Christian Laugier.

Sharing the road with humans constitutes, along with the need for robust perception systems, one of the major challenges holding back the large-scale deployment of automated driving technology. The actions taken by human drivers are determined by a complex set of interdependent factors, which are very hard to model (e.g. intentions, perception, emotions). As a consequence, any prediction of human behavior will always be inherently uncertain, and becomes even more so as the prediction horizon increases. Fully automated vehicles are thus required to make navigation decisions based on the uncertain states and intentions of surrounding vehicles. Building upon previous work, where we showed how to estimate the states and maneuver intentions of surrounding drivers [91], we developed a decision-making system for automated vehicles in highway environments. The task is modeled as a Partially Observable Markov Decision Process and solved in an online

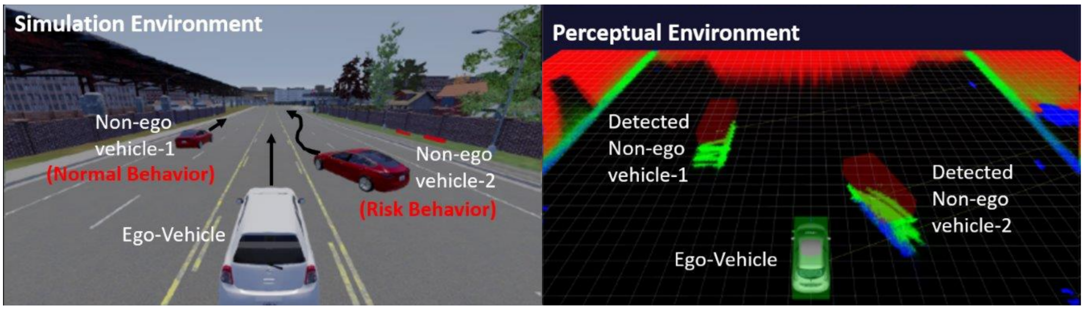


Figure 14. (Left) Simulation environment with CARLA simulator. The white vehicle is the ego-vehicle and two non-ego vehicles are simulated to perform anomaly movements. (Right) Perceptual environment with CMCDOT framework. By analyzing the real-time velocity and position of vehicles, the state and behavior of vehicles can be inferred.

fashion using Monte Carlo tree search. At each decision step, a search tree of beliefs is incrementally built and explored in order to find the current best action for the ego-vehicle. The beliefs represent the predicted state of the world as a response to the actions of the ego-vehicle and are updated using an interaction- and intention-aware probabilistic model. To estimate the long-term consequences of any action, we rely on a lightweight model-based prediction of the scene that assumes risk-averse behavior for all agents. We refer to the proposed decision-making approach as human-like, since it mimics the human abilities of anticipating the intentions of surrounding drivers and of considering the long-term consequences of their actions based on an approximate, common-sense, prediction of the scene. We evaluated the proposed approach in two different simulated navigational tasks: lane change planning and longitudinal control. The results obtained demonstrated the ability of the proposed approach to make foresighted decisions and to leverage the uncertain intention estimations of surrounding drivers.

This work was published in ITSC 2019 [44]. It constitutes the last contribution of the PhD dissertation of David Sierra González, which was defended in April 2019 [12].

7.3.7. Contextualized Emergency Trajectory Planning using severity curves

Participants: Luiz Serafim Guardini, Anne Spalanzani, Christian Laugier, Philippe Martinet.

Perception and interpretation of the surroundings is essential for human drivers as well as for (semi-)autonomous vehicles navigation. To improve such interpretation, a lot of effort has been put in place, for example predicting the behavior of pedestrians and other drivers. Nevertheless, to date, cost maps still have considered simple contextualized objects (for instance, binary allowed/forbidden zones or a fixed weight to each type of object). In this work, the risk of injury issued by accidentology is employed to each class of object present in the scene. The scene is analyzed according to dynamic characteristics related to the Ego vehicle and enclosing objects. The aim is to have a better assessment of the surroundings by creating a navigation cost map and to get an improvement on the understanding of the collision severity in the scene. During the first year of his PhD, Luiz Serafim Gaurdini focused on the development of a probabilistic costmap that expresses the Probability of Collision with Injury Risk (PCIR) (see an example on Figure 15). On top of the information gathered by sensors, it includes the severity of injury in the event of a collision between ego and the objects in the scene. This cost map provides enhanced information to perform vehicle motion planning.

7.3.8. Game theoretic decision making for autonomous vehicles' merge manoeuvre in high traffic scenarios

Participants: Mario Garzon, Anne Spalanzani.

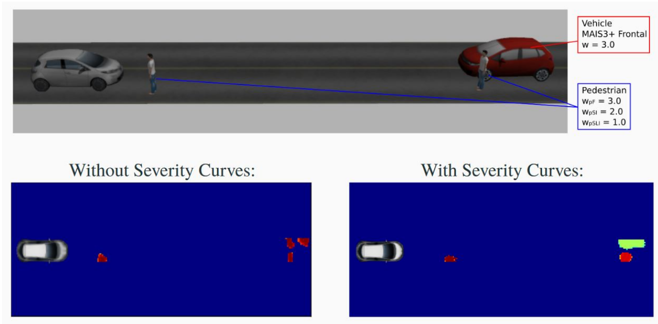


Figure 15. Illustration of the Probabilistic Costmap including the notion of Injury Risk

The goal of this work is to provide a solution for a very challenging task: the merge manoeuvre in high traffic scenarios (see Figure 16). Unlike previous approaches, the proposed solution does not rely on vehicle-to-vehicle communication or any specific coordination, moreover, it is capable of anticipating both the actions of other players and their reactions to the autonomous vehicle's movements. The game used is an iterative, multi-player level-k model,which uses cognitive hierarchy reasoning for decision making and has been proved to correctly model human decisions in uncertain situations. This model uses reinforcement learning to obtain a near-optimal policy, and since it is an iterative model, it is possible to define a goal state so that the policy tries to reach it. To test the decision making process, a kinematic simulation was implemented. The resulting policy was compared with a rule-based approach. The experiments show that the decision making system is capable of correctly performing the merge manoeuvre, by taking actions that require reactions of the other players to be successfully completed. This work was published in [48].

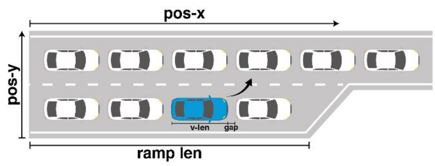


Figure 16. Typical scenario of changing lane in high traffic

7.4. Motion-planning in dense pedestrian environments

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). The works of year 2019 are presented below.

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 cars or platoons (PhD work of P. Vasishta). We first 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 [98]. We then developed an extension to the Growing Hidden Markov Model (GHMM) method that 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. 17.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 [99] and won the **best student paper** of the conference. In 2019, Pavan Vasishta defended his PhD on this topic.

Figure 17. 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 the Principle of Proactive Navigation.

7.4.2. Proactive Navigation for navigating dense human populated environments

Participants: Maria Kabtoul, Anne Spalanzani, Philippe Martinet.

Developing autonomous vehicles capable of navigating safely and socially around pedestrians is a major challenge in intelligent transportation. This challenge cannot be met without understanding pedestrians' behavioral response to an autonomous vehicle, and the task of building a clear and quantitative description of the pedestrian to vehicle interaction remains a key milestone in autonomous navigation research. As a step towards safe proactive navigation in a spaceshared with pedestrians, we start to introduce in 2018 a pedestrian-vehicle interaction behavioral model. The model estimates the pedestrian's cooperation with the vehicle in an interaction scenario by a quantitative time-varying function. Using this cooperation estimation the pedestrian's trajectory is predicted by a cooperation-based trajectory planning model (see Figure 17.b). Both parts of the model are tested and validated using real-life recorded scenarios of pedestrian-vehicle interaction. The model is capable of describing and predicting agents' behaviors when interacting with a vehicle in both lateral and frontal crossing scenarios.

7.4.3. Modelling crowds and autonomous vehicles using Extended Social Force Models

Participants: Manon Predhumeau, Anne Spalanzani, Julie Dugdale.

The focus of this work has been on the realistic simulation of crowds in shared spaces. We have developed a simulator, based on empirical studies and the state of the art, using PED-SIM software. The simulator takes into account the density of crowds, different social group structures in different contexts, inter and intra group forces, and collision avoidance strategies of pedestrians. The Social Force Model (SFM) successfully reproduces many collective phenomena in evacuations or dense crowds. However, pedestrians behaviour is context dependent and the SFM has some limitations when simulating crowds in an open environment under normal conditions. Specifically, in an urban public square pedestrians tend to expand their personal space and try to avoid dense areas to reduce the risk of collision. Based on the SFM, the proposed model splits the perception of pedestrians into a large perception zone and a restricted frontal zone to which they pay more attention. Through their perceptions, the agents estimate the crowd density and dynamically adapt their personal space. Finally, the original social force is tuned to reflect pedestrians preference of avoiding dense areas by turning rather than slowing down as long as there is enough space. Simulation results show that in the considered context the proposed approach produces more realistic behaviours than the original SFM. The simulated crowd is less dense with the same number of pedestrians and less collisions occur, which better fits the observations of sparse crowds in an open place under normal condition [40].

7.4.4. Deep Reinforcement Learning based Vehicle Navigation amongst pedestrians

Participants: Niranjan Deshpande, Anne Spalanzani, Dominique Vaufreydaz.

The objective of this work is to develop a navigation system for an autonomous vehicle in urban environments. The urban environment would consist of other road users as well including other vehicles and pedestrians. Specifically, the focus is on the decision making (behaviour planning) aspect of navigation. In this work, we propose to use Deep Reinforcement Learning as a method to learn decision making. We have developed a Deep Q-Network based agent for decision making amongst pedestrians using the SUMO simulator. This Deep Q-Network based agent is trained for a typical intersection crossing setup amongst pedestrians (see Figure 18). We propose a grid based representation as a state space input to the learning agent. With this grid based representation and our reward function the agent learns a policy capable of driving safely around pedestrians and also follow the traffic rule. This work was published in [35].

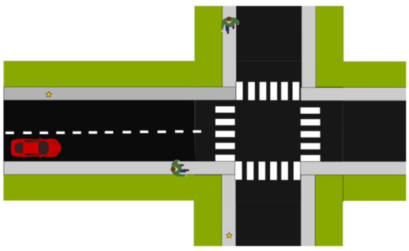


Figure 18. Typical intersection crossing used for training the behavior of the autonomous vehicle

7.5. Learning robot high-level behaviors

7.5.1. 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 complex 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. find all occurences of a given object in the environment, or recognize the current activities of all the actors in this environment) or which can be given in an encoded form as additional input, like text. Addressing these problems requires competences in computer vision, machine learning and AI, and robotics (navigation and paths planning).

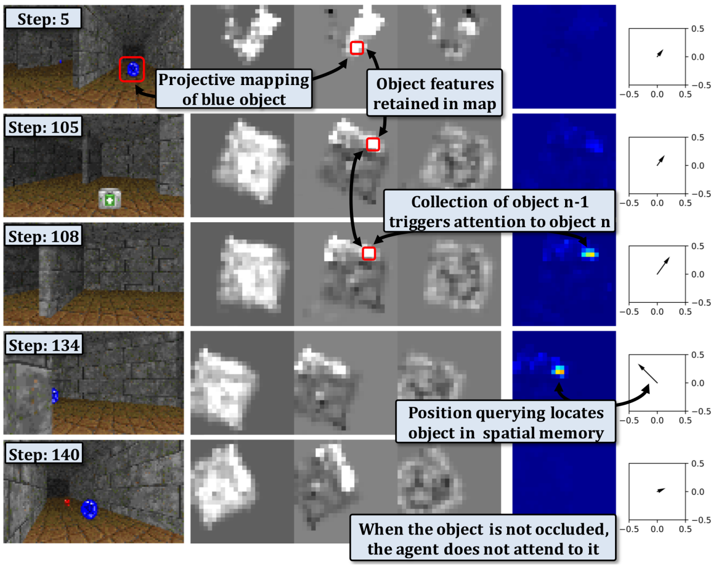


Figure 19. Analysis of the EgoMap for key steps (different rows) during an episode. Left column - RGB observations, central column - the three largest PCA components of features mapped in the spatially structured memory, right - attention heat map (result of the query) and x,y query position vector.

A critical part for solving these kind of problems involving autonomous agents is handling memory and planning. An example can be derived from biology, where an animal that is able to store and recall pertinent information about their environment is likely to exceed the performance of an animal whose behavior is purely reactive. Many control problems in partially observed 3D environments involve long term dependencies and planning. Solving these problems requires agents to learn several key capacities: *spatial reasoning* — to explore the environment in an efficient manner and to learn spatio-temporal regularities and affordances. The agent needs to discover relevant objects, store their positions for later use, their possible interactions and the eventual relationships between the objects and the task at hand. Semantic mapping is a key feature in these tasks. A second feature is *discovering semantics from interactions* — while solutions exist for semantic mapping and semantic SLAM [64], [94], a more interesting problem arises when the semantics of objects and their affordances are not supervised, but defined through the task and thus learned from reward.

We started this work in the end of 2017, following the arrival of C. Wolf and his 2 year delegation in the team between Sept 2017. to Sept. 2019, through combinations of reinforcement learning and deep learning. The underlying scientific challenge here is to automatically learn representations which allow the agent to solve multiple sub problems required 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, which started on October 2018.

First work proposed a new 3D benchmark for Reinforcement learning, which requires high-level reasoning through the automatic discovery of object affordances [58]. Follow-up work proposed EgoMap, a spatially structured metric neural memory architecture integrating projective geometry in deep reinforcement learning, which we show to outperform classical recurrent baselines. In particular, we show that through visualizations that the agents learns to map relevant objects in its spatial memory without any supervision purely from reward (see Fig. 19). Ongoing work aims to propose a fully differentiable topological memory for Deep-RL.

Creating agents capable of high-level reasoning based on structured memory is main topic of the AI Chair "REMEMBER" obtained by C.Wolf in late 2019 and which involves O. Simonin and J. Dibangoye (Inria Chroma) as well as Laetitia Matignon (LIRIS/Univ Lyon 1). The chair is co-financed by ANR, Naver Labs Europe and INSA-Lyon.

7.5.2. Social robot : NAMO extension and RoboCup@home competition

Participants: Jacques Saraydaryan, Fabrice Jumel, Olivier Simonin, Benoit Renault, Laetitia Matignon, Christian Wolf.

Since 3 years, we investigate robot/humanoid navigation and complex tasks in populated environments such as homes :

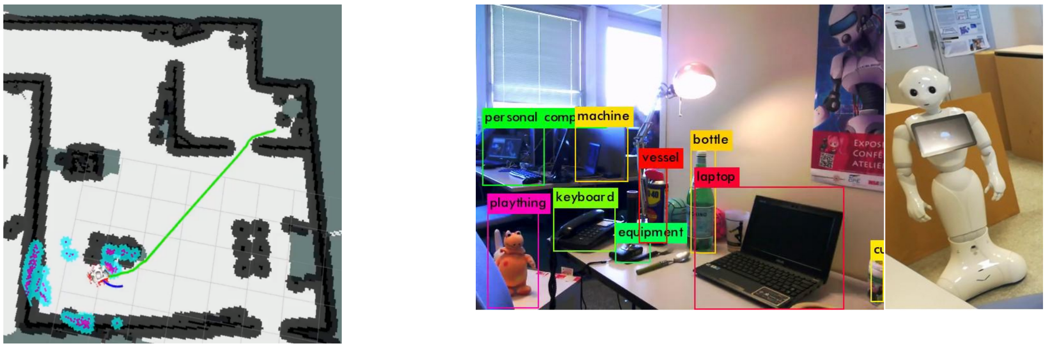


Figure 20. (a) Pepper's navigation and mapping (b) Object detection with Pepper based on vision/deep learning techniques.

- In 2018 we started to study NAMO problems (Navigation Among Movable Obstacles). In his PhD work, Benoit Renault is extending NAMO to Social-NAMO by modeling obstacle hindrance in regards to space access. Defining new spatial cost functions, we extend NAMO algorithms with the ability to maintain area accesses (connectivity) for humans and robots [41]. We also developed a simulator of NAMO problems and algorithms, called S-NAMO-SIM.
- In the context of the **RoboCup** international competition, we created in 2017 the'LyonTech' team, gathering members from Chroma (INSA/CPE/UCBL). We investigated several issues to make humanoid robots able to evolve in a populated indoor environment : decision making and navigation (Fig. 20.a), human and object recognition based on deep learning techniques (Fig. 20.b) and human-robot interaction. In July 2018, we participated for the first time to the RoboCup and we reached the 5th place of the SSL league (Robocup@home with Pepper). In July 2019, we participated to the RoboCup organized in Sydney and we obtained the 3rd place of the SSL league. We also awarded the scientific Best Paper of the RoboCup conference [43].

7.6. Sequential decision-making

This research is the follow up of a subgroup led by Jilles S. Dibangoye carried out during the last four 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.

7.6.1. Optimally solving zero-sum games using centralized planning for decentralized control theory

Participants: Jilles S. Dibangoye, Olivier Buffet [Inria Nancy], Vincent Thomas [Inria Nancy], Abdallah Saffidine [Univ. New South Whales], Christopher Amato [Univ. New Hampshire], François Charpillet [Inria Nancy, Larsen team].

During the last two years, we investigated deep and standard reinforcement learning for solving systems with multiple agents and different information structures. Our preliminary results include:

1. (Theoretical) – As an extension of [68] in the competitive cases, we characterize the optimal solution of two-player fully and partially observable stochastic games.
2. (Theoretical) – We further exhibit new underlying structures of the optimal solution for both non-cooperative two-player settings with information asymmetry, one agent sees what the other does and sees.
3. (Algorithmic) – We extend a non-trivial procedure for computing such optimal solutions.

This work aims at reinforcing a recent 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 [59] for zs-POSGs, and coined them maximin (resp. minimax) equations. Even more importantly, we demonstrated Von Neumann & Morgenstern's minimax theorem [102] [103] holds in zs-POSGs. We further proved that value functions—solutions of maximin (resp. minimax) equations—yield special structures. More specifically, the optimal value functions are Lipchitz-continuous. Together these findings allow us to extend planning 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.

7.6.2. Learning 3D Navigation Protocols on Touch Interfaces with Cooperative Multi-Agent Reinforcement Learning

Participants: Jilles S. Dibangoye, Christian Wolf [INSA Lyon], Quentin Debard [INSA Lyon], Stephane Canu [INSA Rouen].

During the last year, we investigated a number of real-life applications of deep multi-agent reinforcement learning techniques [34]. In particular, we propose to automatically learn a new interaction protocol allowing to map a 2D user input to 3D actions in virtual environments using reinforcement learning (RL). A fundamental problem of RL methods is the vast amount of inter- actions often required, which are difficult to come by when humans are involved. To overcome this limitation, we make use of two collaborative agents. The first agent models the human by learning to perform the 2D finger trajectories. The second agent acts as the interaction protocol, interpreting and translating to 3D operations the 2D finger trajectories from the first agent. We restrict the learned 2D trajectories to be similar to a training set of collected human gestures by first performing state representation learning, prior to reinforcement learning. This state representation learning is addressed by projecting the gestures into a latent space learned by a variational auto encoder (VAE).

7.7. Multi-Robot Routing

7.7.1. Global-local optimization in autonomous multi-vehicle 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 [62]. We were surprised 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.

7.7.2. *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 [66] and iterative local search [101]. 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.

7.7.3. *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 [86], 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-Sat : Multi-robot routing 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 under review.
- 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.

- 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. 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. In the context of the PHC 'DRONEM' project [0] we also developed a flow-based approach to the synchronization problem with the team of Prof. Gabriela Czibula from Babes-Bolyai University in Cluj-Napoca, Romania, see [37].

7.8. Multi-UAV exploration and communication

7.8.1. Multi-UAV Exploration and Visual Coverage of 3D Environments

Participants: Alessandro Renzaglia, Olivier Simonin, Jilles Dibangoye, Vincent Le Doze.

Figure 21. (a) UAVs Chroma simulator (b) Intel Aero quadrotors platform (c) Crazyflie micro-UAV platform extended with UWB decawave chip.

Multi-robot teams, especially when involving aerial vehicles (UAVs [0]), are extremely efficient systems to help humans in acquiring information on large and complex environments. In these scenarios, two fundamental tasks are static coverage and exploration. In both cases, the robots' goal is to navigate through the environment and cooperate to maximize the observed area, either by finding the optimal static configuration which provides the best global view in the case of the coverage or by maximizing the new observed areas at every step until the environment becomes completely known in the case of the exploration.

Although these tasks are usually considered separately in the literature, we proposed a common framework where both problems are formulated as the maximization of online acquired information via the definition of single-robot optimization functions, which differs only slightly in the two cases to take into account the static and dynamic nature of coverage and exploration respectively [0]. A common derivative-free approach based on a stochastic approximation of these functions and their successive optimization is proposed, resulting in a fast and decentralized solution. The locality of this methodology limits however this solution to have local optimality guarantees and specific additional layers are proposed for the two problems to improve the final performance.

[0]Hubert Curien Partnership

[0]Unmanned Aerial Vehicles

[0]A. Renzaglia, J. Dibangoye, V. Le Doze and O. Simonin, "A Common Optimization Framework for Multi-Robot Exploration and Coverage in 3D Environments," *sumbitted to* Journal of Intelligent & Robotic Systems, *under review*.

For the exploration problem, this resulted in a novel decentralized approach which alternates gradient-free stochastic optimization and a frontier-based approach [42] (IROS'19), [47]. Our method allows each robot to generate its own trajectory based on the collected data and the local map built integrating the information shared by its teammates. Whenever a local optimum is reached, which corresponds to a location surrounded by already explored areas, the algorithm identifies the closest frontier to get over it and restarts the local optimization. Its low computational cost, the capability to deal with constraints and the decentralized decision-making make it particularly suitable for multi-robot applications in complex 3D environments.

In the case of visual coverage, we studied how suitable initializations for the UAVs' positions can be computed offline based on a partial knowledge on the environment and how they can affect the final performance of the online measurements-based optimization. The main contribution of this work was thus to add another layer, based on the concept of Centroidal Voronoi Tessellation, to the optimization scheme in order to exploit an a priori sparse information on the environment to cover. The resulting method, taking advantages of the complementary properties of geometric and stochastic optimization, significantly improves the result of the uninitialized solution 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 [88].

Both previous approaches have been tested in realistic simulations based on our extension of Gazebo, called SimuDronesGR (see Fig. 21.a). The development of this UAVs simulator, which includes realistic models of both the environment and the aerial vehicle's dynamics and sensors, is an important current activity in Chroma. Such a simulator has the fundamental role of allowing for realistic tests to validate the developed algorithms and to better prepare the implementation of these solutions on the robotic platform of the team (Intel Aero quadrotors, Fig. 21.b) for real experiments.

7.8.2. Communication-based control of swarm of UAVs

Participants: Remy Grunblatt, Olivier Simonin, Isabelle Guerin-Lassous [Inria/Lyon 1 Dante team], Alexandre Bonnefond.

Intel WiFi controllers are used in many common devices, such as laptops, but also in the Intel Aero Ready-to-Fly UAVs (Unmanned Aerial Vehicle). The mobility capabilities of these devices lead to greater dynamics in radio conditions, and therefore introduce a need for a suitable and efficient rate adaptation algorithm. In the context of the PhD of Remy Grunblatt, we have reverse-engineered the Intel rate adaptation mechanism from the source code of the IwlWifi Linux driver, and we have given, in a comprehensive form, the underlying rate adaptation algorithm named Iwl-Mvm-Rs. We have also implemented the Iwl-Mvm-Rs algorithm in the NS-3 simulator. Thanks to this implementation, we can evaluate the performance of Iwl-Mvm-Rs in different scenarios (static and with mobility, with and without fast fading). We also compared the performances of Iwl-Mvm-Rs with the ones of Minstrel-HT and IdealWifi, also implemented in the NS-3 simulator. This work has been published in ACM MSWiM conference (A) [36].

In the end of 2019, we obtained a DGA/Inria AI project, called "Dynaflock", aiming to extend the flocking approach to control swarm of communicating UAVs. Alexandre Bonnefond started a PhD to elaborate dynamic flocking models based on the link quality, which can be measured online.

7.8.3. Ultra-WideBand based localization & control of micro-UAVs fleets

Participants: Stephane d'Alu, Olivier Simonin, Oana Iova [Inria/INSA Agora team], Hervé Rivano [Inria/INSA Agora team].

The literature on autonomous flight of swarm of UAVs in indoor environments shows it requires the use of an external camera-based localization, i.e. a motion capture system. Indoor flying without such an expensive equipment installed in the infrastructure remains a challenge. To tackle this challenge, we investigate the Ultra-WideBand technology which can be embedded on micro UAVs as a way to estimate inter-drone distances (see Fig. 21.c Crazyflie micro-UAV). In our approach, the distance information is a fundamental building block to perform a self-maintaining formation flight. We defined and experimented a time-of-flight distance computation, using UWB decawave chips. We showed a Crazyflie flying and computing its position in function of three fixed anchors. We also tested a two-UAV flight where inter-distance is measured to avoid collisions. See first results in [33].

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.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Inria ADT 'CORDES' (2017-19) & 'COLOC' (2019-20)

Participants: Olivier Simonin, Vincent Le Doze, Jilles Dibangoye, Alessandro Renzaglia.

The COLOC ADT, which follows the CORDES ADT, aims to coordinate a team of UAVs using both SLAM techniques and communication-based localization, considering outdoor urban environments. These ADT are coordinated by Olivier Simonin. They fund an Inria expert engineer position in Chroma (Vincent Le Doze, 10/17-11/20) focusing on UAVs control and localization. The project provides both a 3D simulator of UAV fleets (SimuDronesGR) and a new experimental platform exploiting IntelAero UAVs.

9.1.2. COMODYS project, FIL (Federation d'Informatique de Lyon), 2017-19

Participants: Laetitia Matignon, Olivier Simonin.

Project between two teams of two laboratories from Lyon : 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.

9.1.3. WIFI-Drones project, FIL (Federation d'Informatique de Lyon), 2019-21

Participants: Remy Grunblatt, Isabelle Guerin-Lassous [Inria/Lyon1 Dante team], Olivier Simonin.

Project between two teams of two laboratories from Lyon : DANTE (LIP) and CHROMA (CITI), entitled "*Performances des communications Wi-Fi dans les réseaux de drones : une approche expérimentale*", 2019-2021. Leader : I. Guerin-Lassous & O. Simonin.

The project aims to experimentally evaluate the Wireless communication in UAVs fleet scenarios. We consider the recent version of Wi-Fi based on 802.11n and 802.11 ac. Experimental measures will be used to build propagation models in order to be integrated in UAVs fleet simulations (in particular with Gazebo and NS3 simulators).

9.2. National Initiatives

9.2.1. ANR

9.2.1.1. ANR JCJC "Plasma" (2019-2023)

The ANR JCJC Plasma, led by Jilles S. Dibangoye, aims at developing a general theory and algorithms with provable guarantees to treat planning and (deep) RL problems arising from the study of multi-agent sequential decision-making, which may be described as Partially Observable Stochastic Games (POSG), see Figure 1. We shall contribute to the development of theoretical foundations of the fields of intelligent agents and MASs by characterizing the underlying structure of the multi-agent decision-making problems and designing scalable and error-bounded algorithms. The research group is made of four senior researchers, O. Simonin, C. Wolf (INSA Lyon), F. Charpillet (Inria Nancy) and O. Buffet (Inria Nancy), and two junior researchers Jilles S. Dibangoye and A. Saffidine (Univeristy of New South Whales). We plan to hire one PhD and one post-doc for two years as well as internships. We received a support for 42-months starting in March 2020 with a financial support of about 254 269,80 euros.

9.2.1.2. ANR "Delicio" (2019-2023)

The ANR Delicio, led by C. Wolf (INSA Lyon, LIRIS), proposes fundamental and applied research in the areas of Machine Learning and Control with applications to drone (UAV) fleet control. The consortium is made of 3 academic partners: INSA-Lyon/LIRIS (C. Wolf and L. Matignon), INSA-Lyon/CICI (J. Dibangoye, O. Simonin, and I. Redko), University Lyon 1/LAGEPP (M. Nadri, V. Andrieu, D. Astolfi, L. bako, and G. Casadei), and ONERA (S. Bertrand, J. Marzat, H. Piet-Lahanier). We plan to hire two Ph.D and two post-doc for one year as well as interships. We received a support for 48-months starting in October 2019 with a financial support of about 540 000 euros.

9.2.1.3. ANR "Valet" (2016-19)

The ANR VALET, led by A. Spalanzani, proposes a novel approach for solving the 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 Irccyn Ecole Centrale de Nantes and the AKKA company. The PhD student (Pavan Vashista) recruited in this project focuses 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, codirected by D. Vaufreydaz (Inria/PervasiveInteraction), has been defended in June 2019.

9.2.1.4. ANR "HIANIC" (2017-21)

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 (Hawai team) and LS2N laboratory (ARMEN and PACCE teams).

9.2.1.5. 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.

9.2.2. FUI Projects

9.2.2.1. 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, with 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.

9.2.2.2. FUI STAR (2018 – 2021)

Participants: Andres Gomez Hernandez, Olivier Simonin, Christian Laugier.

The Project STAR is coordinated by IVECO. The academic partners of the project are Inria Grenoble-Rhône-Alpes, 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 lane Inria is involved in helping design situation awareness perception, especialy 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.

9.2.3. DGA/Inria AI projects

9.2.3.1. "DYNAFLOCK" (2019-2023)

The DYNAFLOCK project, led by O. Simonin, aims to extend flocking-based decentralized control of swarm of UAVs by considering the link quality between communicating entities. The consortium is made of 2 Inria teams from Lyon : Chroma and Dante (involving Prof. I. Guerin-Lassous). The PhD student (Alexandre Bonnefond) recruited in this project aims at defining dynamic flocking models based on the link quality. In 2020, an engineer will be recruited to conduct experiments with a quadrotors platform.
Funding of Dynaflock : $\sim$ 250 K€.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

9.3.1.1. ICT Robotics project "BugWright2" (2020-23)

Success for European H2020 ICT Robotics project application 'BugWright2' (9M€), led by C. Pradalier (CNRS, GeorgiaTech Metz). Chroma is partner and responsible of WP6.

Title : Autonomous Robotic Inspection and Maintenance on Ship Hulls and Storage Tanks

1/01/2020 - 31/12/2023

O. Simonin leads the Multi-Robot Systems work-package (WP6). Chroma will work on multi-robot planning and experiment under environmental constraints. The Agora team is also involved (H. Rivano, O. Iova) to work on robot localization based on the Ultra-WideBand technology.

Funding for Chroma & Agora teams : 600K€

http://dream.georgiatech-metz.fr/research-projects/bugwright2/

9.3.2. Collaborations with Major European Organizations

- ETHZ, Zurich, Autonomous System laboratory, (Switzerland)
- University of Zurich, Robotics and Perception Group (Switzerland) Vision and IMU data Fusion for 3D navigation in GPS denied environment.
- Karlsruhe Institut fur Technologie (KIT, Germany) Autonomous Driving.
- University of Babes-Bolyai, Cluj-Napoca (Romania). Multi-robot patrolling and Machine Learning (PHC "DRONEM" 2017-18).
- Vislab Parma (Italy) Embedded Perception & Autonomous Driving (visits, projects submissions, and book chapter in the new edition of the Handbook of Robotics).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- C. Laugier was General Co-Chair of IEEE/RSJ IROS [0] 2019 (Macau, Nov 2019).
- C. Laugier co-organized in the scope of IEEE/RSJ IROS 2019 and of the IEEE RAS Technical Committee IGV-ITS, the 11th Workshop of the series PPNIV (Planning, Perception and Navigation for Intelligent Vehicles) having attracted about 300 people (see http://project.inria.fr/ppniv19 for more details).

[0] IEEE/RSJ International Conference on Intelligent Robots and Systems

- C. Laugier co-organized at IEEE/RSJ IROS 2019 a Cutting Edge Forum on "Robotics, AI and ITS for Autonomous Driving" having attracted about 150 people (see https://project.inria.fr/ad19/ for more details).
- O. Simonin was General Co-Chair and Co-Organizer with F. Charpillet (Inria Nancy) of the JNRR 2019 conference, the bi-annual conference gathering the French Robotic community, in Vittel (∼ 200 pers.) October 15-17 2019. https://jnrr2019.loria.fr/
- O. Simonin was General and Scientific Chair of JFSMA 2019 (*27emes Journées Francophones sur les Systèmes Multi-Agents*), in Toulouse, as part of PFIA 2019 (Plate-forme IA), July 1-5 2019. https://www.irit.fr/pfia2019/jfsma/

10.1.1.2. Member of the Organizing Committees

- Agostino Martinelli was in the organizing committee of the workshop on "visual-inertial navigation: challenges and applications", held during IROS 2019 in Macau, China.
- C. Laugier was Member of the Organizing Committee of the conference IEEE CIS-RAM 2019 (Bangkok, Nov 2019).
- J. S. Dibangoye and O. Simonin were Members of the Organizing Committee of IA^2 – Institut d'Autonomne en Intelligence Artificielle, Lyon 2019.

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- O. Simonin was Chair of JFSMA Program Committee (27emes journées francophones sur les systèmes multi-agents), Toulouse, France, July 3-5, 2019. He co-edited with S. Combettes (IRIT/Toulouse) the JFSMA 2019 Proceedings, Cépaduès 2019, ISBN 9782364937192.

10.1.2.2. Member of the Conference Program Committees

- C. Laugier served as Associate Editor for the conference IEEE ICRA 2019 (Montreal, May 2019) and also for IEEE ICRA 2020 (Paris, May 2020).
- C. Laugier served as member of the Senior Program Committee of IEEE/RSJ IROS 2019 (Macau, Nov 2019).
- A. Martinelli was Associate Editor for IEEE ICRA 2020.
- A. Spalanzani was PC member of the ECAI 2019 conference.
- J. S. Dibangoye was Senior PC member of the ECAI 2019 conference.
- J. S. Dibangoye was PC member of the IJCAI 2019 conference.
- J. S. Dibangoye was PC member of the AAAI 2020 conference.
- J. S. Dibangoye was PC member of the NeurIPS 2019 conference (selected as one of the best reviewers).
- J. S. Dibangoye was PC member of the PKDD-ECML 2019 conference.
- J. S. Dibangoye was PC member of the JFPDA 2019 conference.
- O. Simonin was Senior PC member of the ECAI 2019 conference.
- O. Simonin was PC member of the IJCAI 2019 conference.
- O. Simonin was PC member of the AAAI 2020 conference.
- O. Simonin was PC member of the ICAPS 2019 conference - Robotics Track (29th International Conference on Automated Planning and Scheduling).
- O. Simonin was PC member of the ECMR 2019 conference (European Conference on Mobile Robots).

10.1.2.3. Reviewer

- A. Martinelli served, in quality of reviewer, for the following conferences: ICRA.

- O. Simonin served, in quality of reviewer, for the following conferences: IROS, ICRA.
- J. S. Dibangoye served, in quality of reviewer, for the following conferences: CDC, ACC, AAMAS, JFPDA.

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 Senior Editor of the journal IEEE Transaction on Intelligent Vehicles
- C. Laugier is member of the Editorial Board of the journal IEEE ROBOMECH.
- O. Simonin was member of the Editorial Board of RIA revue (*Revue d'Intelligence Artificielle*), from 1/1/2018 to 1/06/2019. He is now member of the Editorial Board of the new revue ROIA (*Revue Ouverte d'Intelligence Artificielle*).

10.1.3.2. Reviewer - Reviewing Activities

- Agostino Martinelli served, in quality of reviewer, for the following journals: Transaction on Automatic Control, Automatica, Journal of Robust Control, Transaction on Robotics.
- O. Simonin was reviewer for SMCA journal (Systems, Man and Cybernetics: System).
- O. Simonin was reviewer for RIA revue (Revue d'Intelligence Artificielle).
- J. S. Dibangoye was reviewer for AIJ (Artificial Intelligence Journal).
- J. S. Dibangoye was reviewer for JAIR (Journal of Artificial Intelligence Research).
- J. S. Dibangoye was reviewer for JMLR (Journal of Machine Learning Research).
- J. S. Dibangoye was reviewer for TAC (IEEE Transaction on Automatic Control).

10.1.4. Invited Talks

- J. S. Dibangoye gave an invited talk at the Czech technical University in Prague, Czech Republic, in July 2019.
- J. S. Dibangoye gave an invited talk at the MAFTEC working group of GdR IA, in Caen, France, in February 2019.
- C. Laugier gave an invited talk at the conference IS Auto Europe 2019, Berlin, April 2019. Title: Embedded Sensor Fusion and Perception for Autonomous Vehicles.
- C. Laugier gave an invited talk at the conference IS Auto Europe 2019, Berlin, April 2019. Title: Embedded Sensor Fusion and Perception for Autonomous Vehicles.
- C. Laugier gave an invited talk at the Technological Conference Minalogic B2B 2019, Grenoble, May 2019. Title: Mixing Bayesian and AI approaches for Autonomous Driving.
- C. Laugier gave an invited keynote talk at the conference IEEE World Robotics Conference WRC 2019, Beijing, August 2019. Title: Impact of AI on Autonomous Driving.
- C. Laugier gave an invited Pioneer's talk at the conference IEEE/RSJ IROS 2019, Macau, November 2019. Title: A Journey in the history of Automated Driving.
- C. Laugier gave an invited keynote talk at the Cutting Edge Forum "Robotics, AI and ITS for Autonomous Driving", Conference IEEE/RSJ IROS 2019, Macau, November 2019. Title: Situation Awareness and Decision-making for Autonomous Driving.
- A. Martinelli gave the invited talk "Nonlinear Unknown Input Observability: the General Analytic Solution" on behalf of the Automatic Control Laboratory and the Institute for Dynamic Systems and Control at the ETHZ of Zurich, May, 6, 2019
- A. Martinelli gave invited talk "Nonlinear Unknown Input Observability: the General Analytic Solution" at La Sapienza University, Control lab, Rome, November, 29, 2019

- A. Spalanzani gave a talk that the "Journées Nationales de la Recherche en Robotique" (JNRR) in October 2019.
- O. Simonin gave an invited talk at the GdR Robotics Winter School: Robotica Principia "Decision making in robotics", Centre de recherche Inria Sophia Antipolis – Méditérranée, France. January 21-25, 2019, [55].
- O. Simonin gave an invited talk at the GdR IA2 Autumn School: "Introduction à la coordination multi-agent", CITI Lab/INSA Lyon.

10.1.5. Leadership within the Scientific Community

- C. Laugier is co-chair with Philippe Martinet, Marcelo Ang and Denis Wolf of the IEEE RAS Technical Committee on "Autonomous Ground Vehicles and Intelligent Transportation Systems (AGV-ITS)". He founded this IEEE RAS TC about 15 years ago, as a bridge between the RAS and ITS societies of IEEE.
- C. Laugier is member of the Steering Committee of IEEE/RSJ IROS.
- C. Laugier is member of the Scientific Committee of the French GDR Robotique.
- C. Laugier is member of several International Award Committees for IEEE RAS (Chapter Award Committee) and for IEEE/RSJ IROS.
- O. Simonin is member of the Board of AFIA "*Association Française pour l'Intelligence Artificielle*" (https://afia.asso.fr/).

10.1.6. Scientific Expertise

- O. Simonin was member of the HCERES committee of LIG lab.
- 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 China.

10.1.7. Research Administration

- C. Laugier is a member of several Ministerial and Regional French Committees on Robotics and Autonomous Cars.
- A. Spalanzani was a member of the committee of the ANR project selection in Mobilité et systèmes urbains durables.
- A. Spalanzani was a member of the GDR best Robotics PhD thesis committee.
- O. Simonin is member of the Auvergne-Rhone-Alpes Robotics cluster (Coboteam), for Inria and INSA Lyon entities.
- O. Simonin is member of the Scientific Council of the Digital League (Auvergne-Rhone-Alpes).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

INSA Lyon 5th year : O. Simonin, Resp. of the Robotics option (25 students): AI for Robotics, Multi-Robot Systems, 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 3rd year : Jilles S. Dibangoye, Algorithmics, 24h, L3, Dept. Telecom INSA Lyon, France.

INSA Lyon 3rd year : Jilles S. Dibangoye, WEB, 42h, L3, Dept. Telecom INSA Lyon, France.

INSA Lyon 3rd year : Jilles S. Dibangoye, Operating Systems, 56h, L3, Dept. Telecom INSA Lyon, France.

INSA Lyon 4rd year : Jilles S. Dibangoye, Operating Systems, 16h, Master, Dept. Telecom INSA Lyon, France.

INSA Lyon 5th year : Jilles S. Dibangoye, the Robotics option : AI for Robotics, Robotics projects, 8h, M2, Dept. Telecom INSA Lyon, France.

M2R MoSIG: A. Martinelli, Autonomous Robotics, 12h, ENSIMAG Grenoble.

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.

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)

10.2.2. Supervision

PhD: David Sierra Gonzalez, Towards Human-Like Prediction and Decision-Making for Automated Vehicles in Highway Scenarios, Université Grenoble Alpes, Defended April 1st 2019, C. Laugier, J. Dibangoye, E. Mazer (Inria Pervasive Interaction).

PhD: Pavan Vasishta, Building and Leveraging Prior Knowledge for Predicting Pedestrian Behaviour Around Autonomous Vehicles in Urban Environments, Université Grenoble Alpes, Defended September 30th 2019, A. Spalanzani and D. Vaufreydaz (Inria Pervasive Interaction).

PhD: Mathieu Barbier, Decision making for Intelligent Vehicles, Defended December 11st 2019, C. Laugier, O. Simonin and E. Mazer (Inria Pervasive Interaction).

PhD in progress: Mihai Popescu, Robot fleet mobility under communication constraints, O. Simonin, A. Spalanzani, F. Valois (CITI/Inria Agora).

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.

PhD in progress: Maria Kabtoul, Proactive Navigation in dense crowds, A. Spalanzani and P. Martinet (Inria Chorale).

PhD in progress: Benoit Renault, Navigation coopérative et sociale de robots mobiles en environnement modifiable, O. Simonin and J. Saraydaryan.

PhD in progress: Edward Beeching, Large-scale automatic learning of autonomous agent behavior with structured deep reinforcement learning, C. Wolf, O. Simonin and J. Dibangoye.

PhD in progress: Manon Prédhumeau, Crowd simulation and autonomous vehicle, A. Spalanzani and J. Dugdale (LIG).

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: Alexandre Bonnefond, Large-scale automatic learning of autonomous agent behavior with structured deep reinforcement learning, O. Simonin and I. Guerrin-Lassous (Inria Dante).

Starting PhD: Estéban Carvalho, Safe and aggressive piloting of UAVs, 2019, Ahmad Hably (Gipsa-Lab), Nicolas Marchand (Gipsa-Lab), Jilles S. Dibangoye.

10.2.3. Juries

HDR :

O. Simonin was reviewer and member of the defense committee of the HDR of Charles Lesire (ONERA-ISAE), INP Toulouse, "Architectures délibératives pour la robotique autonome, des algorithmes au logiciel embarqué", March 5th, 2019.

O. Simonin was reviewer and member of the defense committee of the HDR of Abbas-Turki Abdeljalil (UTBM), Université de Haute-Alsace, "Méthodes, modèles et outils pour la régulation coopérative : Application aux véhicules autonomes et connectés", December 10th, 2019.

PhD thesis :

A. Spalanzani was reviewer and member of the defense committee of the PhD thesis of José Mendes Filho, école nationale supérieure de techniques avancées, December 19th, 2019.

O. Simonin was reviewer and member of the defense committee of the PhD thesis of Nicolas Cambier, Université de Technologie de Compiègne (UTC), October 23th, 2019.

O. Simonin was reviewer and member of the defense committee of the PhD thesis of Christophe Reymann, INSA Toulouse, July 8th, 2019.

O. Simonin was member of the defense committee of the PhD thesis of Rustem Abdrakhmanov, Université Clermont Auvergne, June 27th, 2019.

O. Simonin was member of the defense committee of the PhD thesis of Bilel Chenchana, Université de Limoges, March 22th, 2019.

Jilles S. Dibangoye was member of the defense committee of the PhD thesis of Jonathan Cohen, Univ. Caen, June 13th 2019.

C. Laugier and J. Dibangoye were members of the defense committee and co-supervisor of the PhD thesis of David Sierra Gonzalez, UGA Grenoble, April 1st 2019.

C. Laugier and O. Simonin were members of the defense committee and co-supervisor of the PhD thesis of Mathieu Barbier, UGA Grenoble, December 11th 2019.

10.3. Popularization

A. Spalanzani made a cross interview on women in robotics (https://www.inria.fr/centre/nancy/actualites/une-carriere-en-robotique-regards-croises-sur-des-parcours-d-exception)

10.3.1. Articles and contents

O. Simonin gave an interview to Le Monde Informatique - IT Tour à Lyon "*La voiture autonome : mythe ou réalité ?*" (https://www.lemondeinformatique.fr/actualites/lire-interview-video-olivier-simonin-professeur-inria-insa-lyonetgrenoble-77605.html) (26/09/19)

10.3.2. Education

A. Spalanzani made a seminar on "social autonomous vehicle navigation" in the scope of "cycle de conférences ISN".

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] A. RENZAGLIA, J. S. DIBANGOYE, V. LE DOZE, O. SIMONIN.*Combining Stochastic Optimization and Frontiers for Aerial Multi-Robot Exploration of 3D Terrains*, in "IROS 2019 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Macau, China, November 2019, https://hal.inria.fr/hal-02164806

[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] M. BARBIER.*Crossing of Road Intersections: Decision-Making Under Uncertainty for Autonomous Vehicles*, Comue Université Grenoble Alpes, December 2019, https://hal.inria.fr/tel-02424655

[12] D. SIERRA GONZÁLEZ.*Towards Human-Like Prediction and Decision-Making for Automated Vehicles in Highway Scenarios*, Université Grenoble Alpes, April 2019, https://tel.archives-ouvertes.fr/tel-02184362

[13] P. VASISHTA.*Building and Leveraging Prior Knowledge for Predicting Pedestrian Behaviour Around Autonomous Vehicles in UrbanEnvironments*, Inria Grenoble Rhône-Alpes, Université de Grenoble, September 2019, https://hal.archives-ouvertes.fr/tel-02401123

Articles in International Peer-Reviewed Journal

[14] L. BRIÑÓN-ARRANZ, A. RENZAGLIA, L. SCHENATO.*Multi-Robot Symmetric Formations for Gradient and Hessian Estimation with Application to Source Seeking*, in "IEEE Transactions on Robotics", June 2019, vol. 35, n^{o} 3, p. 782-789 [*DOI :* 10.1109/TRO.2019.2895509], https://hal.inria.fr/hal-01991153

[15] Ö. ERKENT, C. WOLF, C. LAUGIER.*End-to-End Learning of Semantic Grid Estimation Deep Neural Network with Occupancy Grids*, in "Unmanned systems", July 2019, vol. 7, n^{o} 3, p. 171-181 [*DOI :* 10.1142/S2301385019410036], https://hal.archives-ouvertes.fr/hal-02302533

[16] F. JUMEL.*Advancing Research at the RoboCup@Home Competition [Competitions]*, in "IEEE Robotics and Automation Magazine", June 2019, vol. 26, n^{o} 2, p. 7-9 [*DOI :* 10.1109/MRA.2019.2908571], https://hal.archives-ouvertes.fr/hal-02305530

[17] A. MARTINELLI.*Nonlinear Unknown Input Observability: Extension of the Observability Rank Condition*, in "IEEE Transactions on Automatic Control", January 2019, vol. 64, n^{o} 1, p. 222 - 237 [*DOI :* 10.1109/TAC.2018.2798806], https://hal.archives-ouvertes.fr/hal-01966303

[18] A. MARTINELLI.*Cooperative Visual-Inertial Odometry: Analysis of Singularities, Degeneracies and Minimal Cases*, in "IEEE Robotics and Automation Letters", 2020, p. 1-11, forthcoming, https://hal.inria.fr/hal-02427991

[19] A. MARTINELLI, A. OLIVA, B. MOURRAIN.*Cooperative Visual-Inertial Sensor Fusion: the Analytic Solution*, in "IEEE Robotics and Automation Letters", 2019, vol. 4, n^{o} 2, p. 453-460 [*DOI :* 10.1109/LRA.2019.2891025], https://hal.archives-ouvertes.fr/hal-01966542

[20] A. MARTINELLI, A. RENZAGLIA, A. OLIVA.*Cooperative Visual-Inertial Sensor Fusion: Fundamental Equations and State Determination in Closed-Form*, in "Autonomous Robots", 2019, p. 1-19, forthcoming [*DOI :* 10.1007/S10514-019-09841-8], https://hal.inria.fr/hal-02013869

[21] P. R. PALAFOX, M. GARZÓN, J. VALENTE, J. J. ROLDÁN, A. BARRIENTOS.*Robust Visual-Aided Autonomous Takeoff, Tracking, and Landing of a Small UAV on a Moving Landing Platform for Life-Long Operation*, in "Applied Sciences", July 2019, vol. 9, n^{o} 13, 2661 [*DOI :* 10.3390/APP9132661], https://hal.inria.fr/hal-02293315

[22] A. RENZAGLIA, L. BRIÑÓN-ARRANZ.*Search and Localization of a Weak Source with a Multi-Robot Formation*, in "Journal of Intelligent and Robotic Systems", 2019, p. 1-12 [*DOI :* 10.1007/S10846-019-01014-0], https://hal.inria.fr/hal-02068180

[23] J. J. ROLDÁN, V. DÍAZ-MAROTO, J. REAL, P. R. PALAFOX, M. GARZÓN, J. VALENTE, A. BARRIENTOS.*Press Start to Play: Classifying Multi-Robot Operators and Predicting Their Strategies through a Videogame*, in "Robotics", September 2019, vol. 8, n^{o} 3, p. 1-14 [*DOI :* 10.3390/ROBOTICS8030053], https://hal.inria.fr/hal-02293309

[24] S. SIERRA M., M. GARZÓN, M. MUNERA, C. CIFUENTES.*Human–Robot–Environment Interaction Interface for Smart Walker Assisted Gait: AGoRA Walker*, in "Sensors", July 2019, vol. 19, n^{o} 13, p. 1-29 [*DOI :* 10.3390/S19132897], https://hal.inria.fr/hal-02182511

[25] J. VALENTE, J. J. ROLDÁN, M. GARZÓN, A. BARRIENTOS.*Towards Airborne Thermography via Low-Cost Thermopile Infrared Sensors*, in "Drones", March 2019, vol. 3, n^{o} 1, p. 1-11 [*DOI :* 10.3390/DRONES3010030], https://hal.inria.fr/hal-02086238

Invited Conferences

[26] C. LAUGIER.*A journey in the history of Automated Driving*, in "IROS 2019 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Macau, China, IEEE, November 2019, p. 1-27, Invited Pionneer's talk at IEEE/RSJ IROS 2019, https://hal.inria.fr/hal-02428196

[27] C. LAUGIER.*Embedded Sensor Fusion and Perception for Autonomous Vehicle*, in "IS Auto Europe 2019 - Image Sensors Automotive Conference", Berlin, Germany, April 2019, p. 1-30, https://hal.inria.fr/hal-02434279

[28] C. LAUGIER.*Impact of AI on Autonomous Driving*, in "WRC 2019 - WRC 2019- IEEE World Robot Conference", Beijing, China, IEEE, August 2019, p. 1-27, https://hal.inria.fr/hal-02429111

[29] C. LAUGIER.*Mixing Bayesian and Artificial Intelligence approaches for Autonomous Driving*, in "Tech M&A 2019 - Minalogic Technical Conference", Grenoble, France, May 2019, https://hal.inria.fr/hal-02434275

[30] C. LAUGIER.*Situation Awareness & Decision-making for Autonomous Driving*, in "IROS 2019 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Macau, China, IEEE, November 2019, p. 1-25, https://hal.inria.fr/hal-02429023

International Conferences with Proceedings

[31] M. BARBIER, A. RENZAGLIA, J. QUILBEUF, L. RUMMELHARD, A. PAIGWAR, C. LAUGIER, A. LEGAY, J. IBAÑEZ-GUZMÁN, O. SIMONIN.*Validation of Perception and Decision-Making Systems for Autonomous Driving via Statistical Model Checking*, in "IV 2019 - 30th IEEE Intelligent Vehicles Symposium", Paris, France, IEEE, June 2019, p. 252-259 [*DOI :* 10.1109/IVS.2019.8813793], https://hal.inria.fr/hal-02127889

[32] R. BRESSON, J. SARAYDARYAN, J. DUGDALE, A. SPALANZANI.*Socially Compliant Navigation in Dense Crowds*, in "IV 2019 - 30th IEEE Intelligent Vehicles Symposium", Paris, France, IEEE, June 2019, p. 64-69 [*DOI :* 10.1109/IVS.2019.8814288], https://hal.inria.fr/hal-02096006

[33] S. D'ALU, O. IOVA, O. SIMONIN, H. RIVANO.*Demo: In-flight Localisation of Micro-UAVs using Ultra-Wide Band*, in "EWSN 2020 - International Conference on Embedded Wireless Systems and Networks", Lyon, France, ACM, February 2020, https://hal.archives-ouvertes.fr/hal-02421783

[34] Q. DEBARD, J. STEEVE DIBANGOYE, S. CANU, C. WOLF.*Learning 3D Navigation Protocols on Touch Interfaces with Cooperative Multi-Agent Reinforcement Learning*, in "ECML PKDD 2019 - European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases", Würzburg, Germany, ECML PKDD 2019 - European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 2019, p. 1-16, https://hal.archives-ouvertes.fr/hal-02302554

[35] N. DESHPANDE, A. SPALANZANI.*Deep Reinforcement Learning based Vehicle Navigation amongst pedestrians using a Grid-based state representation*, in "ITSC 2019 - IEEE Intelligent Transportation Systems Conference", Auckland, New Zealand, IEEE, October 2019, p. 2081-2086 [*DOI :* 10.1109/ITSC.2019.8917299], https://hal.inria.fr/hal-02409042

[36] R. GRÜNBLATT, I. GUÉRIN-LASSOUS, O. SIMONIN.*Simulation and Performance Evaluation of the Intel Rate Adaptation Algorithm*, in "MSWiM 2019 - 22nd ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems", Miami Beach, United States, ACM, November 2019, p. 27-34 [*DOI :* 10.1145/3345768.3355921], https://hal.inria.fr/hal-02282508

[37] V.-S. IONESCU, Z. ONET-MARIAN, M.-G. BADITA, G. CZIBULA, M.-I. POPESCU, J. S. DIBANGOYE, O. SIMONIN.*DynFloR: A Flow Approach for Data Delivery Optimization in Multi-Robot Network Patrolling*, in "KES 2019 - 23rd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems", Budapest, Hungary, September 2019, https://hal.archives-ouvertes.fr/hal-02386692

[38] P. LEDENT, A. PAIGWAR, A. RENZAGLIA, R. MATEESCU, C. LAUGIER.*Formal Validation of Probabilistic Collision Risk Estimation for Autonomous Driving*, in "CIS-RAM 2019 - 9th IEEE International Conference on Cybernetics and Intelligent Systems (CIS) Robotics, Automation and Mechatronics (RAM)", Bangkok, Thailand, IEEE, November 2019, p. 1-6, https://hal.inria.fr/hal-02355551

[39] A. PAIGWAR, Ö. ERKENT, C. WOLF, C. LAUGIER.*Attentional PointNet for 3D-Object Detection in Point Clouds*, in "CVPR 2019 - Workshop on Autonomous driving", Long Beach, California, United States, June 2019, p. 1-10, https://hal.inria.fr/hal-02156555

[40] M. PRÉDHUMEAU, J. DUGDALE, A. SPALANZANI.*Adapting the Social Force Model for Low Density Crowds in Open Environments*, in "15th Social Simulation Conference", Mainz, Germany, September 2019, https://hal.archives-ouvertes.fr/hal-02194735

[41] B. RENAULT, J. SARAYDARYAN, O. SIMONIN.*Towards S-NAMO: Socially-aware Navigation Among Movable Obstacles*, in "RoboCup 2019 - 23rd Annual RoboCup International Symposium", Sydney, Australia, S. CHALUP, T. NIEMUELLER, J. SUTHAKORN, M.-A. WILLIAMS (editors), RoboCup 2019: Robot World Cup XXIII, Springer, July 2019, p. 241-254, https://arxiv.org/abs/1909.10809 [*DOI :* 10.1007/978-3-030-35699-6_19], https://hal.archives-ouvertes.fr/hal-02293242

[42] A. RENZAGLIA, J. S. DIBANGOYE, V. LE DOZE, O. SIMONIN.*Combining Stochastic Optimization and Frontiers for Aerial Multi-Robot Exploration of 3D Terrains*, in "IROS 2019 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Macau, China, November 2019, https://hal.inria.fr/hal-02164806

[43] *Best Paper*
J. SARAYDARYAN, R. LEBER, F. JUMEL.*People management framework using a 2D camera for human-robot social interactions*, in "RoboCup 2019 - 23rd Annual RoboCup International Symposium", Sydney, Australia, Robocup 2019: Robot World Cup XXIII, July 2019, p. 1-13, https://hal.archives-ouvertes.fr/hal-02318916.

[44] D. SIERRA GONZÁLEZ, M. GARZÓN, J. S. DIBANGOYE, C. LAUGIER.*Human-Like Decision-Making for Automated Driving in Highways*, in "ITSC 2019 - 22nd IEEE International Conference on Intelligent Transportation Systems", Auckland, New Zealand, October 2019, p. 1-8, https://hal.inria.fr/hal-02188235

[45] C. YANG, A. RENZAGLIA, A. PAIGWAR, C. LAUGIER, D. WANG.*Driving Behavior Assessment and Anomaly Detection for Intelligent Vehicles*, in "CIS-RAM 2019 - 9th IEEE International Conference on Cybernetics and Intelligent Systems (CIS) Robotics, Automation and Mechatronics (RAM)", Bangkok, Thailand, November 2019, p. 1-6, https://hal.inria.fr/hal-02355809

National Conferences with Proceeding

[46] R. GRÜNBLATT, I. GUÉRIN-LASSOUS, O. SIMONIN.*Study of the Intel WiFi Rate Adaptation Algorithm*, in "CoRes 2019 - Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication", Saint-Laurent-de-la-Cabrerisse, France, June 2019, p. 1-4, https://hal.inria.fr/hal-02126333

[47] A. RENZAGLIA, J. S. DIBANGOYE, V. LE DOZE, O. SIMONIN.*Combiner Optimisation Stochastique et Frontières pour l'Exploration 3D avec un Flotte de Drones*, in "JFSMA 2019 - 27èmes Journées Francophones sur les Systèmes Multi-Agents (JFSMA)", Toulouse, France, July 2019, p. 1-9, https://hal.inria.fr/hal-02160346

Conferences without Proceedings

[48] M. GARZÓN, A. SPALANZANI.*Game theoretic decision making for autonomous vehicles' merge manoeuvre in high traffic scenarios*, in "ITSC 2019 - IEEE Intelligent Transportation Systems Conference", Auckland, New Zealand, IEEE, October 2019, p. 3448-3453 [*DOI :* 10.1109/ITSC.2019.8917314], https://hal.archives-ouvertes.fr/hal-02388757

Scientific Books (or Scientific Book chapters)

[49] M. BARBIER, J. IBAÑEZ-GUZMÁN, C. LAUGIER, O. SIMONIN.*Validation Framework Applied to the Decision-Making Process for an Autonomous Vehicle Crossing Road Intersections*, in "Validation and Verification of Automated Systems", Springer International Publishing, November 2020, p. 179-205 [*DOI :* 10.1007/978-3-030-14628-3_15], https://hal.inria.fr/hal-02424651

Research Reports

[50] R. ASGHAR, M. GARZÓN, J. LUSSEREAU, C. LAUGIER.*Map Relative Localization Based on Visual and Topological Map Matching*, Inria Chroma, September 2019, https://hal.inria.fr/hal-02412837

[51] L. BRIÑÓN-ARRANZ, A. RENZAGLIA, L. SCHENATO.*Technical report on: Multi-Robot Symmetric Formations for Gradient and Hessian Estimation with Application to Source Seeking*, Inria Grenoble Rhône-Alpes, Université de Grenoble, January 2019, p. 1-8, https://hal.inria.fr/hal-01989573

[52] M. A. DIAZ ZAPATA.*Contribution to Panoptic Segmentation*, Inria ; Universidad Autónoma de Occidente, September 2019, n° RT-0506, https://hal.inria.fr/hal-02300774

Other Publications

[53] Q. DEBARD, J. STEEVE DIBANGOYE, S. CANU, C. WOLF.*Learning 3D Navigation Protocols on Touch Interfaces with Cooperative Multi-Agent Reinforcement Learning*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02302526

[54] A. PAIGWAR, Ö. ERKENT, C. WOLF, C. LAUGIER.*Attentional PointNet for 3D-Object Detection in Point Clouds*, June 2019, 1, CVPR 2019 - Workshop on Autonomous Driving, Poster, https://hal.inria.fr/hal-02160868

[55] O. SIMONIN.*Decision Making in (multi) Robot Systems*, January 2019, p. 1-50, Lecture, https://hal.inria.fr/cel-02130231

References in notes

[56] *Global status report on road safety 2018*, in "World Health Organization", 2018

[57] 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

[58] E. BEECHING, C. WOLF, J. DIBANGOYE, O. SIMONIN.*Deep Reinforcement Learning on a Budget: 3D Control and Reasoning Without a Supercomputer*, in "Arxiv pre-print 1904.01806", 2019

[59] R. E. BELLMAN.*Dynamic Programming*, Dover Publications, Incorporated, 1957

[60] 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

[61] 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

[62] 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 Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes (JFPDA 2017)", Caen, France, July 2017, https://hal.archives-ouvertes.fr/hal-01576351

[63] 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

[64] J. CIVERA, D. GALVEZ-LOPEZ, L. RIAZUELO, J. D. TARDÓS, J. M. M. MONTIEL.*Towards semantic SLAM using a monocular camera*, in "IROS", 2011

[65] 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° 1, p. 19–30

[66] 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° 1, p. 4–31

[67] 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

[68] 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

[69] G. FERRER, A. SANFELIU.*Bayesian Human Motion Intentionality Prediction in urban environments*, in "Pattern Recognition Letters", 2014, vol. 44, n° 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

[70] C. FORSTER, M. PIZZOLI, D. SCARAMUZZA.*SVO: Fast Semi-Direct Monocular Visual Odometry*, in "ICRA", 2014

[71] 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/

[72] A. GEIGER, P. LENZ, R. URTASUN.*Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite*, in "Conference on Computer Vision and Pattern Recognition (CVPR)", 2012

[73] K. HE, G. GKIOXARI, P. DOLLÁR, R. GIRSHICK.*Mask r-cnn*, in "Proceedings of the IEEE international conference on computer vision", 2017, p. 2961–2969

[74] 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° 1, p. 158–176

[75] A. KIRILLOV, K. HE, R. GIRSHICK, C. ROTHER, P. DOLLÁR.*Panoptic segmentation*, in "Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition", 2019, p. 9404–9413

[76] 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

[77] 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

[78] 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

[79] 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

[80] 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

[81] 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

[82] 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, ACM, July 2018, p. 559-567

[83] D. NEVEN, B. D. BRABANDERE, S. GEORGOULIS, M. PROESMANS, L. V. GOOL.*Towards End-to-End Lane Detection: an Instance Segmentation Approach*, in "CoRR", 2018, vol. abs/1802.05591, http://arxiv.org/abs/1802.05591

[84] 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

[85] 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

[86] 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

[87] C. R. QI, H. SU, K. MO, L. J. GUIBAS.*Pointnet: Deep learning on point sets for 3d classification and segmentation*, in "Proc. Computer Vision and Pattern Recognition (CVPR), IEEE", 2017, vol. 1, n^o 2, 4

[88] A. RENZAGLIA, J. DIBANGOYE, V. LE DOZE, O. SIMONIN.*Multi-UAV Visual Coverage of Partially Known 3D Surfaces: Voronoi-based Initialization to Improve Local Optimizers*, in "arXiv preprint arXiv:1901.10272", 2019

[89] 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

[90] M. SANDLER, A. HOWARD, M. ZHU, A. ZHMOGINOV, L. C. CHEN.*MobileNetV2: Inverted Residuals and Linear Bottlenecks*, in "CVPR", 2018, p. 4510–4520

[91] 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

[92] 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

[93] 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

[94] K. TATENO, F. TOMBARI, I. LAINA, N. NAVAB.*CNN-SLAM: Real-time dense monocular SLAM with learned depth prediction*, in "CVPR", 2017

[95] M. TEICHMANN, M. WEBER, J. M. ZÖLLNER, R. CIPOLLA, R. URTASUN.*MultiNet: Real-time Joint Semantic Reasoning for Autonomous Driving*, in "CoRR", 2016, vol. abs/1612.07695, http://arxiv.org/abs/1612.07695

[96] S. THRUN, W. BURGARD, D. FOX.*Probabilistic Robotics*, The MIT Press, September 2005, 1

[97] 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

[98] 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

[99] 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

[100] 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

[101] 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

[102] J. VON NEUMANN, O. MORGENSTERN.*Theory of games and economic behavior*, Princeton University Press, 2007

[103] 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
A2.1.6. - Concurrent programming
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
B7.2.1. - Smart vehicles
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

Armen Inants [Inria, Engineer]

PhD Students

Pierre Bouvier [UGA, from Oct 2019]
Lina Marsso [Inria]
Ajay Muroor Nadumane [Inria]
Umar Ozeer [Orange Labs, until Nov 2019]

Visiting Scientist
Luca Di Stefano [Gran Sasso Science Institute, from Mar 2019 until Jul 2019]

Administrative Assistant
Myriam Etienne [Inria, Administrative Assistant]

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 [6] 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*) [47], 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 [35], 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 § 7.5) 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, train supervision 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. Highlights of the Year

5.1. Highlights of the Year

Frédéric Lang, together with Franco Mazzanti from CNR-ISTI/FMT (Pisa, Italy), won all the gold medals for the "Parallel CTL" and "Parallel LTL" tracks of the RERS'2019 (*Rigorous Evaluation of Reactive Systems*) challenge [0]. The goal of these two tracks was to verify 180 properties expressed in the branching-time temporal logic CTL and 180 properties expressed in the linear-time temporal logic LTL. These properties had to be evaluated on various complex systems, having up to 70 concurrent processes and 234 synchronization actions. To attack such difficult problems, Lang and Mazzanti decided to join forces, and managed to evaluate all the 360 properties correctly, by designing new verification algorithms and exploiting the compositional verification techniques of CADP.

6. New Software and Platforms

6.1. CADP

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*) [5] is a toolbox for protocols and distributed systems engineering.

[0] http://rers-challenge.org/2019

In this toolbox, we develop and maintain the following tools:

- CAESAR.ADT [34] 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 [40], [39] 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 [35] 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 [43], CTL [32], ACTL [33], etc.) to be defined in XTL.

- PBG (*Partitioned BCG Graph*) is a file format implementing the theoretical concept of *Partitioned LTS* [38] 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 [45] 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 [36] 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/

6.2. 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 [37], 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/

7. New Results

7.1. New Formal Languages and their Implementations

7.1.1. LOTOS and LNT Specification Languages

Participants: Hubert Garavel, Frédéric Lang, Wendelin Serwe.

LNT [6] [31] 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 § 7.5). 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 2019, a new option `-depend` has been added to the LNT_DEPEND, LNT2LOTOS, and LNT.OPEN tools. LNT_DEPEND now supports the case where the user replaces the predefined LNT modules (e.g., BOOLEAN, NATURAL, etc.) with custom versions. LNT_DEPEND has been made faster and displays better error messages. The LOTOS code generated by LNT2LOTOS for parallel compositions could be semantically incorrect and has been fixed.

We continued working on the TRAIAN compiler for the LOTOS NT language (a predecessor of LNT), which is used for the construction of most CADP compilers and translators.

The version 2.x of TRAIAN that we have been developing for almost 20 years is increasingly difficult to maintain. It consists of 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 2.x only exists in a 32-bit version, and sometimes hits the 4 GB RAM limit when dealing with large compiler specifications, such as those of LNT2LOTOS or EVALUATOR 5.

For this reason, we undertook in 2018 a complete rewrite of TRAIAN (version 3.0) 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.x, 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.x, but should be ultimately capable of bootstrapping itself.

In 2019, we continued the development of TRAIAN 3.0, whose grammar and syntax analysis phase was already almost complete. We fully implemented several static program analysis phases, among which the following:

- binding analysis, which associates a declaration to every identifier occurring in the program (e.g., type, channel, variable, event, etc.)
- typing analysis (including resolution of function name overloading), which associates a type to every expression in the program
- type-productivity and type-finiteness analysis, which check respectively whether a type has at least one value and whether a type has a finite number of values

We also fully implemented the C function generation phase and started to implement the C type generation phase. To avoid problems when switching from TRAIAN 2.x to TRAIAN 3.0, TRAIAN 3.0 generates almost exactly the same code as TRAIAN 2.x. The principal differences concern the numbers used to uniquely identify symbols (variables and functions) in the generated C code, because these are often derived from the syntax tree.

TRAIAN 3.0 is checked regularly against a non-regression test suite consisting of 845 correct and 1545 incorrect programs.

In total, the functionalities that remain to be implemented in TRAIAN 3.0 represent less than 32% of the code of TRAIAN 2.x.

7.1.2. Nested-Unit Petri Nets

Participants: Pierre Bouvier, Hubert Garavel.

Nested-Unit Petri Nets (NUPNs) is a model of computation that can be seen as an upward-compatible extension of P/T nets, which are enriched with structural information on their concurrent and hierarchical structure. Such structural information can easily be produced when NUPNs are generated from higher-level specifications (e.g., process calculi) and allows logarithmic reductions in the number of bits required to represent reachable states, thus enabling verification tools to perform better. For this reason, NUPNs have been so far implemented in thirteen verification tools developed in four countries, and adopted by two international competitions (the Model Checking Contest and the Rigorous Examination of Reactive Systems challenge).

In 2019, a journal article [13] has been published, which formalizes the complete theory of NUPNs.

The development of software tools for NUPNs has steadily progressed. The file format for NUPNs has been enhanced and made more precise; the NUPN_INFO tool has been extended with two new options; the CAESAR.BDD tool as been extended with six new options and its capabilities and efficiency improved in many respects.

We also revisited the problem of decomposing a Petri net into a network of automata, a problem that has been around since the early 70s. We reformulated this problem as the transformation of an ordinary, one-safe Petri net into a unit-safe NUPN. We developed various transformation methods, all of which we implemented in a tool chain that combines NUPN tools with third-party software, such as SAT solvers, SMT solvers, and tools for graph colouring and finding maximal cliques. We performed an extensive evaluation of these methods on a collection of more than 12,000 nets from diverse sources, including nets whose marking graph is too large for being explored exhaustively.

7.1.3. 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 (*Business Process Model and Notation*) standard, and several industrial platforms have been developed for supporting their design, modeling, and simulation.

In collaboration with Francisco Durán (University of Málaga, Spain) and Camilo Rocha (University of Cali, Colombia), we proposed an approach for the modeling and analysis of resource allocation for business processes. Our approach enables the automatic computation of measures for precisely identifying and optimizing the allocation of resources in business processes, including resource usage over time. The proposed analysis, especially suited to support decision-making strategies, is illustrated with a case study of a parcel ordering and delivery by a fleet of drones. This work comprises an encoding of a significant and expressive subset of BPMN in rewriting logic, an executable logic of concurrent change that can naturally deal with states and concurrent computations. The encoding is by itself a formal semantics and interpreter of the BPMN subset that captures all concurrent behavior and thus is used to simulate the concurrent evolution of any business process with a given number of resources and replicas. This work led to two publications, in an international conference [19] and an international journal [12].

7.2. Parallel and Distributed Verification

7.2.1. Debugging of Concurrent Systems using Counterexample Analysis

Participant: 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 a novel approach to improve the usability of model checking by simplifying the comprehension of counterexamples. Our approach takes as input an LTS model and an (unsatisfied) temporal logic property, and operates in four steps. First, all counterexamples for the property are extracted from the model. Second, the model is analyzed to identify the actions that skip from correct to incorrect behaviours (intuitively, these are the most relevant actions from a debugging perspective). Third, using a panel of abstraction techniques, these actions are extracted from the counterexamples. Fourth, 3D visualization techniques are used for highlighting specific regions in the model, where a choice is possible between executing a correct behaviour or falling into an erroneous part of the model, according to the property under analysis. We developed a tool named CLEAR to fully automate our approach, and we applied it on real-world case studies from various application areas for evaluation purposes. This allowed us to identify several patterns corresponding to typical cases of errors (e.g., interleaving, iteration, causality, etc.).

These results led to two publications in international conferences [15] [16] and a publication to appear in an international journal [11].

7.2.2. Eliminating Data Races in Parallel Programs using Model Checking

Participants: Radu Mateescu, Wendelin Serwe.

Parallelization of existing sequential programs to increase their performance and exploit recent multi- and many-core architectures is a challenging but inevitable effort. One increasingly popular parallelization approach is based on OpenMP, which enables the designer to annotate a sequential program with constructs specifying the parallel execution of code blocks. These constructs are then interpreted by the OpenMP compiler and runtime, which assigns blocks to threads running on a parallel architecture. Although this scheme is very flexible and not (very) intrusive, it does not prevent the occurrence of synchronization errors (e.g., deadlocks) or data races on shared variables.

In the framework of the CAPHCA project (see § 9.2.1.1), in collaboration with Eric Jenn and Viet Anh Nguyen (IRT Saint-Exupéry, Toulouse), we proposed an iterative method to assist the OpenMP parallelization by using formal methods and verification. In each iteration, potential data races are identified by applying to the OpenMP program a lockset analysis, which computes the set of shared variables that potentially need to be protected by locks. To avoid the insertion of superfluous locks, an abstract, action-based formal model of the OpenMP program in LNT is extracted and analyzed using the EVALUATOR model checker of CADP. Spurious locks are detected by checking ACTL formulas expressing the absence of concurrent execution of shared variables accesses. This work led to an international publication [28].

7.3. Timed, Probabilistic, and Stochastic Extensions

7.3.1. On-the-fly Model Checking for Extended Regular Probabilistic Operators

Participants: Armen Inants, Radu Mateescu.

Specifying and verifying quantitative properties of concurrent systems requires expressive and user-friendly property languages combining temporal, data-handling, and quantitative aspects. To this aim, we undertook the quantitative analysis of concurrent systems modeled as PTSs (*Probabilistic Transition Systems*), whose actions contain channel names, data values, and probabilities. We proposed a new regular probabilistic operator that extends naturally the Until operators of PCTL (*Probabilistic Computation Tree Logic*) [41], by specifying the probability measure of a path characterized by a generalized regular formula involving arbitrary computations on data values. We devised an on-the-fly model checking method for this new operator, based on a combined local resolution of linear and Boolean equation systems.

In 2019, we continued this work as follows:

- The MCL v4 language was conservatively extended with the new probabilistic operator, leading to a new version MCL v5.
- A new version 5 of the EVALUATOR model checker that handles the MCL v5 language, was added to CADP. EVALUATOR 5 is backward compatible with EVALUATOR 4, to which it adds a new option "`-epsilon`" specifying the precision of floating-point computations. A new version 5 of the MCL_EXPAND tool, the front-end common to the EVALUATOR 3, 4, and 5 model checkers, was added to CADP. This version is upward compatible with the previous one (except for slight changes in some error messages), it corrects a bug and brings some optimizations in the C code generated. Two new manual pages "evaluator5" and "mcl5" have been added.
- For certain probabilistic formulas (e.g., expressing the step-bounded reachability of events), the on-the-fly model checking procedure can be optimized by taking advantage of the possible *query containments*, i.e., implications between instances of the formula with different data parameters. We studied query containment in DHMLR (*Data-based Hennessy-Milner Logic with Recursion*), a parameterized equational formalism used as intermediate language for model checking MCL formulas. Our method consists in detecting, by static analysis, the containment orders present in the DHMLR representation of an MCL formula, and using the information about parameterized Boolean variable implications to improve the convergence of the BES resolution algorithms. We implemented the method in a prototype extension of EVALUATOR 5 and of the CAESAR_SOLVE library for BES resolution, and applied it for verifying probabilistic and also functional properties (e.g., bounded inevitability). The experiments we carried out on self-stabilizing protocols and communication protocols over unreliable channels showed reductions of up to 50% in memory and up to 33% in execution time. This work led to a paper submitted to an international conference.

7.4. Component-Based Architectures for On-the-Fly Verification

7.4.1. Compositional Verification

Participants: Frédéric Lang, Radu Mateescu.

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 2019, in addition to small bug corrections, we updated SVL to support version 5 of EVALUATOR, and we corrected a semantic bug in the expansion of meta-operators of SVL.

In collaboration with Franco Mazzanti (ISTI-CNR, Pisa, Italy), we also used the compositional verification tools of CADP in the framework of the RERS'2019 challenge [0], which consisted in verifying 180 LTL properties and 180 CTL properties on large models of concurrent systems having up to 70 concurrent processes and 234 synchronization actions.

We applied to these examples the *maximal hiding* technique [48], which consists in hiding in the model all actions that are not necessary to verify the property. We combined this technique with compositional minimization (using the smart reduction heuristic implemented in SVL) as follows:

- In a first attempt, we used the technique consisting in applying minimization modulo either strong bisimulation or divbranching (divergence-preserving branching) bisimulation, depending on the fragment of the modal μ-calculus to which the formula belongs, as proposed in [48]. This was more efficient than non-compositional verification on large models, but not sufficient to verify all RERS problems successfully.
- We then proposed a refinement of this approach, which consists in (1) partitioning the actions of the system to be verified into so-called strong and weak actions, depending on the formula, and (2) minimizing modulo divbranching bisimulation all processes and process compositions containing weak actions only. This is an improvement over the previous technique, since divbranching bisimulation can be used to minimize some processes of the system even though the formula does not belong to the fragment of the μ-calculus adequate with divbranching bisimulation (which corresponds to formulas with an empty set of strong actions). This new technique allowed us to verify a lot more problems successfully, but still letting a few of the largest RERS problems unresolved. We published a paper describing the approach in an international conference [23].
- At last, we designed a new bisimulation relation, named *sharp bisimulation*, parameterized by the strong actions of the system, and we implemented a prototype tool that reduces a behavior graph modulo this relation. Sharp bisimulation parameterized by a set S of strong actions is weaker than strong bisimulation, stronger than divbranching bisimulation, and adequate with formulas whose strong actions are included in S. Such a fine-tuning of the bisimulation relation by strong actions allowed us to verify all RERS problems successfully and to win the 2019 challenge. A paper describing the approach was accepted for publication in an international conference.

7.4.2. Other Component Developments

Participants: Hubert Garavel, Frédéric Lang, Philippe Ledent, Radu Mateescu, Wendelin Serwe.

In 2019, several components of CADP have been improved as follows:

- We enhanced the TESTOR tool by adding the possibility to interact with an SUT (*System Under Test*) using its standard input and output.
- We enhanced the XTL compiler with a function converting a transition label into a string (useful for handling the entire content of the label), and we also corrected three bugs.
- We enhanced MCL_EXPAND 5 with a better detection of nondeterminism in probabilistic formulas and a vacuity check for infinite looping operators, and we also corrected a semantic bug.

[0] http://rers-challenge.org/2019

- We enhanced EVALUATOR 5 with more explanative messages about the assignment of probabilities to transitions, and we corrected two bugs in each of the tools EVALUATOR 4 and 5.
- The C code generated by CAESAR has been modified to suppress GCC 6.5 warnings.
- Several changes have been brought to CADP to enable its use on new platforms, including macOS 10.15 "Catalina" and the forthcoming Debian 10.0 Linux distribution. Various bugs specific to Linux and SunOS systems (Solaris or Illumos/OpenIndiana) have been fixed.

7.5. Real-Life Applications and Case Studies

7.5.1. Autonomous Resilience of Distributed IoT Applications in a Fog Environment

Participants: Umar Ozeer, Gwen Salaün.

Recent computing trends have been advocating for more distributed paradigms, namely Fog computing, which extends the capacities of the Cloud at the edge of the network, that is close to end devices and end users in the physical world. The Fog is a key enabler of the Internet of Things (IoT) applications as it resolves some of the needs that the Cloud fails to provide such as low network latencies, privacy, QoS, and geographical requirements. For this reason, the Fog has become increasingly popular and finds application in many fields such as smart homes and cities, agriculture, healthcare, transportation, etc.

The Fog, however, is unstable because it is constituted of billions of heterogeneous devices in a dynamic ecosystem. IoT devices may regularly fail because of bulk production and cheap design. Moreover, the Fog-IoT ecosystem is cyber-physical and thus devices are subjected to external physical world conditions, which increase the occurrence of failures. When failures occur in such an ecosystem, the resulting inconsistencies in the application affect the physical world by inducing hazardous and costly situations.

In the framework of the collaboration with Orange Labs (see § 8.1.1), we proposed an end-to-end autonomic failure management approach for IoT applications deployed in the Fog. The proposed approach recovers from failures in a cyber-physical consistent way. Cyber-physical consistency aims at maintaining a consistent behavior of the application with respect to the physical world, as well as avoiding dangerous and costly circumstances. The approach was validated using model checking techniques to verify important correctness properties. It was then implemented as a framework called F3ARIoT. This framework was evaluated on a smart home application. The results showed the feasibility of deploying F3ARIoT on real Fog-IoT applications as well as its good performances in regards to end user experience.

These results have been published in U. Ozeer's PhD thesis [10] and at an international conference [26]. Another paper was submitted to an international journal.

7.5.2. Verified Composition and Deployment of IoT Applications

Participants: Alejandro Martinez Rivero, Radu Mateescu, Ajay Muroor Nadumane, Gwen Salaün.

The Internet of Things (IoT) applications are built by interconnecting everyday objects over internet. As IoT is becoming popular among consumers, the challenge of making IoT applications easy to design and deploy is more relevant than ever. In 2019, we considered this challenge along two perspectives.

- In the framework of the collaboration with Nokia Bell Labs (see § 8.1.2), we focused on helping consumers to easily design IoT applications that are correct, and also support the deployment of these applications. The correctness of the applications is ensured through formal methods and verification techniques.

  Using W3C Web of Things (WoT) specification as the basis of our work, we extended the specification of objects in WoT with a behavioural model. This allows us to describe formally the composition of objects and thus, to verify their behavioural correctness. Typically, an IoT application is defined using Event-Condition-Action (ECA) rules of the type "IF event THEN action". Our work supports users to specify not only the ECA rules, but also the composition of rules using a simple, yet expressive language. This makes possible the construction of advanced compositions, which would have been hard or sometimes impossible to build using simple ECA rules. Finally, users are provided with an easy-to-deploy solution for these advanced compositions. All these steps were implemented and packaged in a tool named MozART, built on top of Mozilla WebThings platform. LNT is used as the formal specification language, and various tools of CADP are used for verifying the composition. Also, an execution engine based on Mozilla WebThings API was built to support the deployment of advanced compositions. The work has led to the preparation of two conference articles.

- Building IoT applications of added-value from a set of available devices with minimal human intervention is one of the main challenges facing the IoT. This is a difficult task that requires models for specifying objects, in addition to user-friendly and reliable composition techniques which in turn prevent the design of erroneous applications.

  In collaboration with Francisco Durán (University of Málaga, Spain), we tackled this problem by first describing IoT applications using abstract models obtained from existing models of concrete devices. Then, we proposed automated techniques for building compositions of devices using a repository of available devices, and an abstract goal of what the user expects from such compositions. Since the number of possible solutions can be quite high, we used both filtering and ranking techniques to provide the most relevant solutions to users. The provided solutions satisfy the given goal and may be analysed with respect to properties such as deadlock-freeness or unmatched send messages. Finally, the application can be deployed using existing execution engines. This work led to a publication in an international conference [20].

7.5.3. Autonomous Car

Participants: Philippe Ledent, Lina Marsso, Radu Mateescu, Wendelin Serwe.

Autonomous vehicles are complex cyber-physical systems that must satisfy critical correctness requirements to increase the safety of road traffic. The validation of autonomous driving is a challenging field because of the complexity of its key components (perception of the environment, scene interpretation, decision making and undertaking of actions) and the intertwinning of physical and software components. In 2019, we considered this challenge along two lines of work.

- From the embedded software perspective, autonomous cars can be considered as GALS systems, which integrate reactive synchronous components that interact asynchronously. The complexity induced by combining synchronous and asynchronous aspects makes GALS systems difficult to develop and debug.

  In the framework of the ARC6 collaboration (see § 9.1.1), we proposed a testing methodology for GALS systems that leverages conformance test generation for asynchronous systems to automatically derive realistic scenarios (inputs constraints and oracles), which are necessary ingredients for the unit testing of individual synchronous components, and are difficult and error-prone to design manually. The methodology consists of several steps (derivation of asynchronous test cases from a GALS model and a test purpose, projection of the complete test graph on a synchronous component, extraction and execution of test scenarios) and was illustrated on a simple, but relevant example inspired by autonomous cars. These results were published in L. Marsso's PhD thesis [9] and at an international conference [25].

- In collaboration with Christian Laugier, Anshul Paigwar, and Alessandro Renzaglia (CHROMA project-team), we proposed a new approach where formal verification is employed to validate systems with probabilistic predictions. We focused on the risk assessment generated by CMCDOT (*Conditional Monte Carlo Dense Occupancy Tracker*), a probabilistic perception system for autonomous cars. CMCDOT provides an environment representation through Bayesian probabilistic occupancy grids and estimates Time-to-Collision probabilities for every static and dynamic part of the grid in the near future. To validate the probabilistic collision risk estimation, we used the CARLA simulator to generate a large number of realistic intersection crossing scenarios with two vehicles. The set of scenarios is then validated using the XTL model checker, by defining appropriate KPIs (*Key Performance Indicators*) as temporal logic formulas and also performing a quantitative analysis. This work led to a publication at an international conference [24].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

8.1.1. 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 § 7.5.1), under the supervision of Loïc Letondeur (Orange Labs), Gwen Salaün (CONVECS), François Gaël Ottogalli (Orange Labs), and Jean-Marc Vincent (POLARIS project-team).

8.1.2. 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 (see § 7.5.2) supported by formal methods, under the supervision of Gwen Salaün (CONVECS), Radu Mateescu (CONVECS), Ludovic Noirie, and Michel Le Pallec (Nokia Bell Labs).

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. 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), and Ioannis Parissis (LCIS, Valence).

9.2. National Initiatives

9.2.1. PIA (Programme d'Investissements d'Avenir)

9.2.1.1. 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.

9.2.2. Competitivity Clusters

9.2.2.1. SECURIOT-2

Participants: Hubert Garavel [correspondent], Armen Inants, 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.

9.2.3. 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),
- Michel Le Pallec (Nokia Bell Labs, Nozay),
- Chu-Min Li (University of Picardie Jules Verne),
- Ioannis Parissis and Oum-El-Kheir Aktouf (LCIS, Valence),
- Pascal Poizat (LIP6, Paris).

9.3. European Initiatives

9.3.1. 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.

9.4. 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.

9.4.1. Inria International Partners

9.4.1.1. 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.

[0] http://fmics.inria.fr

9.4.2. Other International Collaborations

In 2019, 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),
- FBK, Trento, Italy (Enrico Magnano),
- Aalto University, Finland and Northeastern University, Boston, Massachusetts (Stavros Tripakis),
- Saarland University, Germany (Holger Hermanns),
- Eindhoven University of Technology, The Netherlands (Anton Wijs and Sander de Putter),
- University of Zielona Gora, Poland (Remigiusz Wisniewski).

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- H. Garavel is an invited professor at Saarland University (Germany) as a holder of the Gay-Lussac Humboldt Prize.
- Hernan Ponce de Leon (Fortiss, Munich, Germany) visited us on June 25–26, 2019. He gave a lecture entitled “*BMC with Weak Memory Models*”.
- Hugues Evrard (Google, London, UK) visited us on October 21, 2019. He gave a lecture entitled “*GPU Schedulers: How Fair is Fair Enough?*”.
- Karoliina Lehtinen (University of Liverpool, UK) visited us on October 23, 2019. She gave a lecture entitled “*Quasi-Polynomial Techniques for Parity Games and Other Problems*”.
- Peter Csaba Ölveczky (University of Oslo, Norway) visited us on November 25, 2019. He gave a lecture entitled “*Formal Specification and Analysis of Real-Time Systems in Real-Time Maude*”.

The annual CONVECS seminar was held in Villard-de-Lans (France) on July 1-3, 2019. The following invited scientists attended the seminar:

- Loïc Letondeur (Orange Labs) gave on July 2, 2019 a talk entitled “*Artificial Intelligence and Edge Computing*”.
- Eric Jenn (IRT Saint-Exupéry / Thales Avionics) gave on July 3, 2019 a talk entitled “*Recent Achievements of the CAPHCA Project*”.
- Viet Anh Nguyen (IRT Saint-Exupéry) gave on July 3, 2019 a talk entitled “*Using Model Checking to Identify Timing Interferences on Multicore Processors*”.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- H. Garavel is a member of the model board [0] of MCC (*Model Checking Contest*). In 2019, 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'2019 have been published online [44].

[0] http://mcc.lip6.fr/models.php

- 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, among which 5 were deposited 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.

10.1.1.2. Member of the Organizing Committees

- L. Marsso and A. Muroor Nadumane were publicity chairs (also in charge of the conference Web site and social media feeds) for SEFM'2019 (*17th International Conference on Software Engineering and Formal Methods*), Oslo, Norway, September 16–20, 2019.

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- G. Salaün was co-chair of SEFM'2019.

10.1.2.2. Member of the Conference Program Committees

- H. Garavel was program committee member of FMICS'2019 (*24rd International Conference on Formal Methods for Industrial Critical Systems*), Amsterdam, The Netherlands, August 30–31, 2019.
- R. Mateescu was program committee member of IFIP-ICTSS'2019 (*31st IFIP International Conference on Testing Software and Systems*), Paris, France, October 15–17, 2019.
- G. Salaün was program committee member of SAC-SVT'2019 (*34th ACM Symposium on Applied Computing - Software Verification and Testing Track*), Limassol, Cyprus, April 8–12, 2019.
- G. Salaün was program committee member of FSEN'2019 (*8th IPM International Conference on Fundamentals of Software Engineering*), Tehran, Iran, May 1–3, 2019.
- G. Salaün was program committee member of COORDINATION'2019 (*21st International Conference on Coordination Models and Languages*), Lyngby, Denmark, June 18–21, 2019.
- G. Salaün was program committee member of Microservices'2019, Dortmund, Germany, February 19–21, 2019.
- G. Salaün was program committee member of COMPSAC-ITIP'2019 and COMPSAC-SETA'2019 (*IEEE International Conference on Computers, Software, and Applications – IT in Practice, and Software Engineering Technologies and Applications*), Wisconsin, USA, July 15–19, 2019.
- G. Salaün was program committee member of HPCS-4PAD'2019 and HPCS-SERCO'2019 (*6th International Symposium on Formal Approaches to Parallel and Distributed Systems, and 3rd Special Session on High Performance Services Computing and Internet Technologies*), Dublin, Ireland, July 15–19, 2019.
- G. Salaün was program committee member of FACS'2019 (*16th International Conference on Formal Aspects of Component Software*), Amsterdam, The Netherlands, October 23–25, 2019.
- G. Salaün was program committee member of FOCLASA'2019 (*17th International Workshop on Coordination and Self-Adaptativeness of Software Applications*), Oslo, Norway, September 17, 2019.
- G. Salaün was program committee member of DATAMOD'2019 (*8th International Symposium "From Data to Models and Back"*), Porto, Portugal, October 7–8, 2019.

[0] http://www.mars-workshop.org/

10.1.2.3. Reviewer

- R. Mateescu was a reviewer for the Festschrift in honor of Stefania Gnesi.
- A. Muroor Nadumane was a reviewer for SAC-SVT'2020, FACS'2019, SEFM'2019, COMPSAC-SETA'2019.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- H. Garavel is an editorial board member of STTT (*Springer International Journal on Software Tools for Technology Transfer*).

10.1.3.2. Reviewer - Reviewing Activities

- F. Lang and A. Muroor Nadumane were reviewers for STTT.
- R. Mateescu was a reviewer for ISSE (*Innovations in Systems and Software Engineering*), IST (*Information and Software Technology*), LMCS (*Logical Methods in Computer Science*), and SCP (*Science of Computer Programming*).
- G. Salaün was a reviewer for JISA (*Journal of Internet Services and Applications*), JLAMP (*Journal of Logical and Algebraic Methods in Programming*), JSME (*Journal of Software: Evolution and Process*), LMCS, and SCP.

10.1.4. Software Dissemination and Internet Visibility

The CONVECS project-team distributes several software tools, among which the CADP toolbox.

In 2019, the main facts are the following:

- We prepared and distributed twelve successive versions (2019-a to 2019-l) of CADP.
- We were requested to grant CADP licenses for 340 different computers in the world.

The CONVECS Web site [0] was updated with scientific contents, announcements, publications, etc.

By the end of December 2019, the CADP forum [0], opened in 2007 for discussions regarding the CADP toolbox, had over 435 registered users and over 1940 messages had been exchanged.

Also, for the 2019 edition of the Model Checking Contest, one family of models generated using CADP (totalling 16 Nested-Unit Petri Nets) was provided.

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 OCARINA Tool and its Extension AADL2LNT for Analysing AADL Descriptions [49]
- The LFD-MPI Tool for Lightweight Formal Development of MPI Applications [52], [53]
- The FTRES Tool for Rare Event Simulation in Dynamic Fault Trees [54]
- Formal Analysis of Security Guidelines for Program Certification [55]

Other teams also used the CADP toolbox for various case studies:

- Applying Automata Learning to Embedded Control Software [50]
- Model Checking Based Approach for Compliance Checking [46]
- Finding Conservative Schema Evolutions by Analysing API Changes [51]
- Selection of Model Checking Strategies using Machine Learning [42]
- Verifying Complex Software Control Systems from Test Objectives [29]
- Designing Safe Synchronous Reactive Systems [30]

[0] http://convecs.inria.fr
[0] http://cadp.inria.fr/forum.html

10.1.5. Invited Talks

- H. Garavel participated in the workshop "*Advancing Verification Competitions as a Scientific Method*" organized by the Lorentz Center (Leiden, The Netherlands) on February 18–22, 2019. He gave a lecture entitled "*Managing Large Collections of Benchmarks: An Experiment Report*".
- H. Garavel presented a retrospective of the Rewrite Engines Competitions at the TOOLympics event at ETAPS 2019, Prague, Czech Republic, April 6–7, 2019.
- H. Garavel was one of the six invited speakers at the panel discussion "*Moore's Law, and More?*" organized at the occasion of the 25th Anniversary of the TACAS conference, Prague, Czech Republic, April 7, 2019.
- H. Garavel gave a highlight talk entitled "*From Safe Petri Nets to NUPNs*" at the 8th IFIP WG 1.8 Workshop on Trends in Concurrency Theory, Amsterdam, The Netherlands, August 31, 2019.
- H. Garavel gave a seminar entitled "*Converting Safe Petri Nets to NUPNs*" at Saarland University, Germany, on September 26, 2019.
- R. Mateescu gave a talk entitled "*The MCL Language for Temporal and Probabilistic Analysis of Concurrent Systems*" at the Languages, Semantics, and Compilation seminar of the LIP laboratory, Lyon, September 19, 2019.
- R. Mateescu visited the Gran Sasso Science Institute (L'Aquila, Italy) on November 4–7, 2019. He gave a talk entitled "*An Action-based Model Checking Language for Concurrent Systems*".
- A. Muroor Nadumane gave a talk entitled "*Advanced Rule-based Composition and Deployment for Web of Things*" at the Inria-Nokia Bell Labs meeting held at Inria Rennes on June 11, 2019.

10.1.6. 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, and giving an advice on the recruitment of temporary engineers.
- 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 2019 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.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

CONVECS is a host team for the computer science master MOSIG (*Master of Science in Informatics at Grenoble*), common to Grenoble INP and Université Grenoble Alpes (UGA).

In 2019, we carried out the following teaching activities:

P. Bouvier supervised a group of seven two-person teams of 2nd year students in the context of the "*projet informatique*" (31 hours "*équivalent TD*", consisting of lab exercises and supervision) at PHELMA.

H. Garavel was a jury member for the MOSIG master defenses in June and September 2019.

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.

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.

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 250 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 (third year of university) at IUT1/UGA.

W. Serwe gave the part on "*Behavioural Testing*" (9 hours "*équivalent TD*") of the course "*Verification and Test Theories*" to second year students of the MOSIG.

W. Serwe supervised 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 2019.

10.2.2. Supervision

PhD: Lina Marsso, "*Formal Methods for Testing Networks of Controllers*", Université Grenoble Alpes, December 10, 2019, R. Mateescu, W. Serwe, and I. Parissis

PhD: Umar Ozeer, "*Autonomic Resilience of Applications in a Largely Distributed Cloud Environment*", Université Grenoble Alpes, December 11, 2019, L. Letondeur, G. Salaün, F.-G. Ottogalli, and J.-M. Vincent

PhD in progress: Pierre Bouvier, "*Implémentation et vérification des langages concurrents de nouvelle génération*", Université Grenoble Alpes, since October 2019, H. Garavel and R. Mateescu

PhD in progress: Ajay Muroor Nadumane, "*Softwarization of Everything: IoT Service Composition*", Université Grenoble Alpes, since October 2017, G. Salaün, R. Mateescu, and M. Le Pallec

10.2.3. Juries

- F. Lang was PhD committee member of Sander de Putter's PhD thesis, entitled "*Verification of Concurrent Systems in a Model-Driven Engineering Workflow*", defended at Technische Universiteit Eindhoven (The Netherlands) on January 28, 2019.
- R. Mateescu was reviewer of The Anh Pham's PhD thesis, entitled "*Efficient State-Space Exploration for Asynchronous Distributed Programs*", defended at ENS Rennes on December 6, 2019.
- G. Salaün was committee president of Xavier Etchevers's Habilitation thesis, entitled "*Evolution du déploiement automatisé d'applications : des infrastructures cloud aux enjeux de l'agilité et de l'intelligence ambiante*", defended at Université Grenoble Alpes on November 29, 2019.
- G. Salaün was PhD committee president of Raphaël Jackse's PhD thesis, entitled "*Interactive Runtime Verification*", defended at Université Grenoble Alpes on December 18, 2019.

10.3. Popularization

10.3.1. Interventions

- R. Mateescu has participated to the "*Journée Transfo : Plongez dans le monde du numérique*" organized at Inria Montbonnot on January 24, 2019.
- R. Mateescu has participated to the "*Fête de la Science*" organized at Inria Montbonnot on October 12, 2019.

11. 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.*Nested-unit Petri nets*, in "Journal of Logical and Algebraic Methods in Programming", April 2019, vol. 104, p. 60-85 [*DOI :* 10.1016/J.JLAMP.2018.11.005], https://hal.inria.fr/hal-02072190

[4] 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

[5] 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

[6] 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

[7] 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

[8] 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

[9] L. MARSSO.*On Model-based Testing of GALS Systems*, Université Grenoble Alpes, December 2019

[10] U. OZEER.*Autonomic Resilience of Applications in a Largely Distributed Cloud Environment*, Université Grenoble Alpes, December 2019

Articles in International Peer-Reviewed Journal

[11] G. BARBON, V. LEROY, G. SALAÜN.*Debugging of Behavioural Models using Counterexample Analysis*, in "IEEE Transactions on Software Engineering", 2019, p. 1-14, forthcoming [*DOI :* 10.1109/TSE.2019.2915303], https://hal.inria.fr/hal-02145610

[12] F. DURÁN, C. ROCHA, G. SALAÜN.*A Rewriting Logic Approach to Resource Allocation Analysis in Business Process Models*, in "Science of Computer Programming", November 2019, vol. 183, p. 1-32 [*DOI :* 10.1016/J.SCICO.2019.102303], https://hal.inria.fr/hal-02345895

[13] H. GARAVEL.*Nested-unit Petri nets*, in "Journal of Logical and Algebraic Methods in Programming", April 2019, vol. 104, p. 60-85 [*DOI :* 10.1016/J.JLAMP.2018.11.005], https://hal.inria.fr/hal-02072190

[14] 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

International Conferences with Proceedings

[15] G. BARBON, V. LEROY, G. SALAÜN.*Debugging of Behavioural Models with CLEAR*, in "TACAS 2019 - 25th International Conference on Tools and Algorithms for the Construction and Analysis of Systems", Prague, Czech Republic, Lecture Notes in Computer Science, Springer, April 2019, vol. 11427, p. 386-392 [*DOI :* 10.1007/978-3-030-17462-0_ 26], https://hal.archives-ouvertes.fr/hal-02121180

[16] G. BARBON, V. LEROY, G. SALAÜN, E. YAH.*Visual Debugging of Behavioural Models*, in "ICSE 2019 - IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings", Montreal, Canada, IEEE, May 2019, p. 107-110 [*DOI :* 10.1109/ICSE-COMPANION.2019.00050], https://hal.inria.fr/hal-02145535

[17] E. BARTOCCI, D. BEYER, P. BLACK, G. FEDYUKOVICH, H. GARAVEL, A. HARTMANNS, M. HUISMAN, F. KORDON, J. NAGELE, M. SIGHIREANU, B. STEFFEN, M. SUDA, G. SUTCLIFFE, T. WEBER, A. YAMADA.*TOOLympics 2019: An Overview of Competitions in Formal Methods*, in "25 Years of TACAS: TOOLympics, Held as Part of ETAPS 2019", Prague, Czech Republic, April 2019, p. 3-24 [*DOI :* 10.1007/978-3-030-17502-3_1], https://hal.sorbonne-universite.fr/hal-02094030

[18] F. DURÁN, H. GARAVEL.*The Rewrite Engines Competitions: A RECtrospective*, in "Proceedings of the 25th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS'19), Part III: TOOLympics", Prague, Czech Republic, Lecture Notes in Computer Science, Springer, April 2019, vol. 11429, p. 1-9 [*DOI :* 10.1007/978-3-030-17502-3], https://hal.inria.fr/hal-02133649

[19] F. DURÁN, C. ROCHA, G. SALAÜN.*Analysis of Resource Allocation of BPMN Processes*, in "ICSOC 2019 - 17th International Conference on Service-Oriented Computing", Toulouse, France, Springer, October 2019, p. 452-457 [*DOI :* 10.1007/978-3-030-33702-5_35], https://hal.inria.fr/hal-02345879

[20] F. DURÁN, G. SALAÜN, A. KRISHNA.*Automated Composition, Analysis and Deployment of IoT Applications*, in "TOOLS 2019 - 51st International Conference on Software Technology: Methods and Tools", Innopolis, Russia, Springer, October 2019, p. 252-268 [*DOI :* 10.1007/978-3-030-29852-4_21], https://hal.inria.fr/hal-02345865

[21] A. KRISHNA, M. LE PALLEC, R. MATEESCU, L. NOIRIE, G. SALAÜN.*IoT Composer: Composition and Deployment of IoT Applications*, in "ICSE 2019 - IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings", Montreal, Canada, IEEE, May 2019, p. 19-22 [*DOI :* 10.1109/ICSE-COMPANION.2019.00028], https://hal.inria.fr/hal-02146569

[22] A. KRISHNA, M. LE PALLEC, R. MATEESCU, L. NOIRIE, G. SALAÜN.*Rigorous Design and Deployment of IoT Applications*, in "FormaliSE 2019 - 7th International Conference on Formal Methods in Software Engineering", Montreal, Canada, ACM, May 2019, p. 21-30 [*DOI :* 10.1109/FORMALISE.2019.00011], https://hal.inria.fr/hal-02146553

[23] F. LANG, R. MATEESCU, F. MAZZANTI.*Compositional Verification of Concurrent Systems by Combining Bisimulations*, in "FM 2019 - 23rd International Conference on Formal Methods", Porto, Portugal, Lecture Notes in Computer Science, Springer Verlag, October 2019, vol. 11800, p. 196-213 [*DOI :* 10.1007/978-3-030-30942-8_13], https://hal.inria.fr/hal-02295459

[24] P. LEDENT, A. PAIGWAR, A. RENZAGLIA, R. MATEESCU, C. LAUGIER.*Formal Validation of Probabilistic Collision Risk Estimation for Autonomous Driving*, in "CIS-RAM 2019 - 9th IEEE International Conference on Cybernetics and Intelligent Systems (CIS) Robotics, Automation and Mechatronics (RAM)", Bangkok, Thailand, IEEE, November 2019, p. 1-6, https://hal.inria.fr/hal-02355551

[25] L. MARSSO, R. MATEESCU, I. PARISSIS, W. SERWE.*Asynchronous Testing of Synchronous Components in GALS Systems*, in "IFM'2019 - 15th International Conference on Integrated Formal Methods", Bergen, Norway, Lecture Notes in Computer Science, Springer Verlag, November 2019, vol. 11918, p. 360-378 [*DOI :* 10.1007/978-3-030-34968-4_20], https://hal.archives-ouvertes.fr/hal-02394989

[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 - 22nd Conference on Innovation in Clouds, Internet and Networks", Paris, France, IEEE, February 2019, p. 230-232 [*DOI :* 10.1109/ICIN.2019.8685909], https://hal.archives-ouvertes.fr/hal-01979686

Scientific Books (or Scientific Book chapters)

[27] H. GARAVEL, R. MATEESCU.*Reflections on Bernhard Steffen's Physics of Software Tools*, in "Models, Mindsets, Meta: The What, the How, and the Why Not?", Springer Verlag, June 2019, p. 186-207 [*DOI :* 10.1007/978-3-030-22348-9_12], https://hal.inria.fr/hal-02394588

[28] V.-A. NGUYEN, W. SERWE, R. MATEESCU, E. JENN.*Hunting Superfluous Locks with Model Checking*, in "From Software Engineering to Formal Methods and Tools, and Back", Lecture Notes in Computer Science, Springer Verlag, October 2019, vol. 11865, p. 416-432 [*DOI :* 10.1007/978-3-030-30985-5_24], https://hal.inria.fr/hal-02314088

References in notes

[29] R. AMEUR-BOULIFA, A. CAVALLI, S. MAAG.*Verifying Complex Software Control Systems from Test Objectives: Application to the ETCS System*, in "Proceedings of the 14th International Conference on Software Technologies (ICSOFT'2019), Prague, Czech Republic", SciTePress, July 2019, p. 397–406

[30] S. CHABANE, R. AMEUR-BOULIFA, M. MOHAMED.*Vers une conception de systèmes réactifs synchrones sûrs*, in "Actes du 12ème Colloque sur la Modélisation des Systèmes Réactifs (MSR'2019), Angers, France", November 2019

[31] 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

[32] 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

[33] 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

[34] 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

[35] 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

[36] 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

[37] 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

[38] 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)

[39] 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

[40] 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

[41] 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

[42] M. HENDRIKS.*Can Machine Learning Automatically Choose your Best Model Checking Strategy?*, EEMCS: Electrical Engineering, Mathematics and Computer Science, University of Twente, 2019

[43] M. HENNESSY, R. MILNER.*Algebraic Laws for Nondeterminism and Concurrency*, in "Journal of the ACM", 1985, vol. 32, p. 137–161

[44] 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 2019 Edition of the Model Checking Contest*, April 2019, http://mcc.lip6.fr/2019/results.php

[45] J. MAGEE, J. KRAMER.*Concurrency: State Models and Java Programs*, 2006, Wiley, April 2006

[46] F. MARTINELLI, F. MERCALDO, V. NARDONE, A. ORLANDO, A. SANTONE, G. VAGLINI.*Model Checking Based Approach for Compliance Checking*, in "Journal of Information Technology and Control", May 2019, vol. 48, n^{o} 2, p. 278–298

[47] 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

[48] R. MATEESCU, A. WIJS.*Property-dependent Reductions Adequate with Divergence-sensitive Branching Bisimilarity*, in "Science of Computer Programming", 2014, vol. 96, p. 354–376

[49] H. MKAOUAR.*A Formal Approach for Real-time Systems Engineering*, University of Sfax, Tunisia, February 2019

[50] J. MOERMAN.*Nominal Techniques and Black Box Testing for Automata Learning*, Radboud University, Nijmegen, The Netherlands, July 2019

[51] L. A. OUBELLI, Y. A. AMEUR, J. BEDOUET, B. CHAUSSERIE-LAPREE, B. LARZUL.*Finding Conservative Schema Evolutions by Analysing API Changes*, in "Proceedings of the 31st International Conference on Software Engineering and Knowledge Engineering (SEKE'2019), Lisbon, Portugal", KSI Research Inc., July 2019, p. 748–777

[52] N. ROSA, H. KAMAL, A. WAGNER.*A LOTOS-based Lightweight Approach to Formally Verify MPI Applications*, Universidade Federal de Pernambuco, Centro de Informática, 2015

[53] N. ROSA, A. WAGNER, H. KAMAL.*Towards Lightweight Formal Development of MPI Applications*, in "Proceedings of 36th WoTUG Conference on Concurrent and Parallel Programming – Communicating Process Architectures (CPA'2015), Kent, UK", Concurrent Systems Engineering Series, Open Channel Publishing, August 2015, vol. 69, p. 67–85

[54] E. RUIJTERS, D. REIJSBERGEN, P.-T. DE BOER, M. STOELINGA.*Rare Event Simulation for Dynamic Fault Trees*, in "Reliability Engineering & System Safety", June 2019, vol. 186, p. 220–231

[55] Z. ZHIOUA, R. AMEUR-BOULIFA, Y. ROUDIER.*Framework for the Formal Specification and Verification of Security Guidelines*, in "Advances in Science, Technology and Engineering Systems Journal", January 2018, vol. 3, n^{o} 1, p. 38–48

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
A2.2.3. - Memory management
A2.2.4. - Parallel architectures
A2.2.5. - Run-time systems
A2.2.6. - GPGPU, FPGA...
A2.3.2. - Cyber-physical systems
A4.4. - Security of equipment and software
A7.1. - Algorithms

Other Research Topics and Application Domains:
B4.5. - Energy consumption
B5.3. - Nanotechnology
B6.1.2. - Software evolution, maintenance
B6.6. - Embedded systems
B6.7. - Computer Industry (harware, equipments...)
B9.1. - Education

1. Team, Visitors, External Collaborators

Research Scientists

Fabrice Rastello [Team leader, Inria, Senior Researcher, HDR]
Frederic Desprez [Inria, Senior Researcher, HDR]

Faculty Members

Florent Bouchez - Tichadou [Univ Grenoble Alpes, Associate Professor]
Francois Broquedis [Institut polytechnique de Grenoble, Associate Professor]
Ylies Falcone [Univ Grenoble Alpes, Associate Professor]
Manuel Selva [Institut polytechnique de Grenoble, Associate Professor, from Sep 2019]

Technical Staff

Ali Kassem [Inria, Engineer]
Chukri Soueidi [Inria, Engineer, from Oct 2019]

PhD Students

Théo Barollet [Inria, PhD Student, from Oct 2019]
Georgios Christodoulis [Univ Grenoble Alpes, PhD Student, until Dec 2019]
Nicolas Derumigny [Univ Grenoble Alpes, PhD Student, from Sep 2019]
Fabian Gruber [Univ Grenoble Alpes, PhD Student, until Sep 2019]
Raphael Jakse [Univ Grenoble Alpes, PhD Student, until Sep 2019]
Auguste Olivry [Univ Grenoble Alpes, PhD Student, from Sep 2019]
Mathieu Stoffel [Bull, PhD Student]
Nicolas Tollenaere [Inria, PhD Student, from Apr 2019]

Post-Doctoral Fellow

Manuel Selva [Univ Grenoble Alpes, Post-Doctoral Fellow, until Aug 2019]

Administrative Assistant

Maria Immaculada Presseguer [Inria, Administrative Assistant]

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. More specifically, an important activity concerns the optimization of machine learning applications for some high-performance accelerators. Members of CORSE already maintain fruitful and strong collaborations with several companies such as STMICROELECTRONICS, Atos/Bull, Kalray.

5. New Software and Platforms

5.1. 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

5.2. 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: Fabrice Rastello
- Publications: hal-02060796v1 - hal-01967828v2

5.3. GUS

KEYWORDS: CPU - Microarchitecture simulation - Performance analysis - Dynamic Analysis

FUNCTIONAL DESCRIPTION: GUS' goal is to detect performance bottlenecks at the very low level on monothread applications by the use of sensitivity analysis. It is coded as a QEMU plug-in in order to collect runtime information that are later treated by the generic CPU model.

- Contact: Nicolas Derumigny
- URL: https://gitlab.inria.fr/nderumig/gus

5.4. Pipedream

KEYWORDS: Performance analysis - CPU - Reverse engineering

SCIENTIFIC DESCRIPTION: Pipedream reverse engineers the following performance characteristics: - Instruction latency – The number of cycles an instruction requires to execute. - Peak micro-op retirement rate – How many fused micro-ops the CPU can retire per cycle. - Micro-fusion – The number of fused micro-ops an instruction decomposes into. - Micro-op decomposition and micro-op port usage – The list of unfused micro-ops every instruction decomposes into and the list of execution ports every one of these micro-ops can execute on.

The first step of the reverse engineering process consists of generating a number of microbenchmarks. Pipedream then runs these benchmark, measuring their performance using hardware counters. The latency, throughput, and micro-fusion of different instructions can then be read directly from these measurements.

The process of finding port mappings, i.e. micro-op decompositions and micro-op port usage, however, is more involved. For this purpose, we have defined a variation of the maximum flow problem which we call the "instruction flow problem". We have developed a linear program (LP) formulation of the instruction flow problem which can be used to calculate the peak IPC and micro-operations per cycle (MPC) a benchmark kernel can theoretically achieve with a given port mapping. The actual port mapping of the underlying hardware is then determined by finding the mapping for which the throughput predicted by instruction flow best matches the actual measured IPC and MPC.

FUNCTIONAL DESCRIPTION: Pipedream is a tool for measuring specific performance characteristics of CPUs It is used to build the performance model of another tool called Gus (https://gitlab.inria.fr/nderumig/gus). Pipedream finds measured performance characteristics such as the throughput and latency of instructions by running a large set of automatically generated microbenchmarks. The tool can also find port mappings, a model of part of the CPU instruction scheduler, by analysing performance measurements of specially crafted microkernels using a LP solver. We have used it to produce a port mapping for the Intel Skylake CPU architecture. Pipedream is able to find the port mappings for some instructions for which existing approaches fall back to manual analysis.

- Contact: Nicolas Derumigny
- URL: https://gitlab.inria.fr/fgruber/pipedream

5.5. Platforms

5.5.1. Grid'5000

Grid'5000 [0] 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: 14828 cores, 829 compute-nodes grouped in homogeneous clusters located in 8 sites in France connected through a dedicated network (Renater), 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 GRID5000 GIS.

5.5.2. SILECS/SLICES

Frédéric Desprez is co-PI with Serge Fdida (Université Sorbonne) of the SILECS [0] infrastructure (IR ministère) which goal is to provide an experimental platform for experimental computer Science (Internet of things, clouds, HPC, big data, IA, wireless technologies, ...). This new infrastructure is based on two existing infrastructures, Grid'5000 and FIT. A European infrastructure (SLICES) is also currently designed with other european partners (Spain, Cyprus, Greece, Netherland, Switzerland, Poland, ...).

6. New Results

6.1. Compiler Optimizations and Analysis

Participants: Fabrice Rastello, Manuel Selva, Fabian Grüber, Diogo Sampaio [CORSE, Inria], Christophe Guillon [STMicroelectronics], P. Sadayappan [OSU, USA], Louis-Noël Pouchet [CSU, USA], Atanas Rountev [OSU, USA], Richard Veras [LSU, USA], Rui Li [UoU, USA], Aravind Sukumaran-Rajam [OSU, USA], Tse Meng Low [CMU, USA].

Our current efforts with regard to code optimization follows two directions. 1. The first consists in improving compiler optimization techniques by considering pattern specific applications such as those related to machine learning. Our first result presented at SC 2019 [10] focuses on tensor contractions. 2. The second consists in developing dynamic analysis based performance debugging tools. Our first results published at PPoPP 2019 [9] and TACO 2020 [7] shows a scalable approach that compresses an execution trace obtained from binary instrumentation and analyses it using a polyhedral compiler.

6.1.1. Analytical Cache Modeling and Tilesize Optimization for Tensor Contractions

Data movement between processor and memory hierarchy is a fundamental bottleneck that limits the performance of many applications on modern computer architectures. Tiling and loop permutation are key techniques for improving data locality. However, selecting effective tile-sizes and loop permutations is particularly challenging for tensor contractions due to the large number of loops. Even state-of-the-art compilers usually produce sub-optimal tile-sizes and loop permutations, as they rely on naïve cost models. In this work we provide an analytical model based approach to multilevel tile size optimization and permutation selection for tensor contractions. Our experimental results show that this approach achieves comparable or better performance than state-of-the-art frameworks and libraries for tensor contractions.

[0] https://www.grid5000.fr/
[0] https://www.silecs.net/

This work is the fruit of the collaboration 8.3.1.1 with OSU. It has been presented at ACM/IEEE International Conference for High Performance Computing, Networking, Storage, and Analysis, SC 2019 [10].

6.1.2. *Profiling-based Polyhedral Optimization Feedback*

This work addresses the problem of reconstructing a compact (static) representation of a binary execution, automatically detecting hot regions and enabling precise feedback about optimization opportunities potentially missed by the compiler. Our framework handles codes with irregular accesses, pointers with indirections, inter-procedural or recursive loop regions. By enabling binary execution analysis we are able to discover run-time properties (i.e., the ability to form a compact representation) as well as inter-procedural optimization opportunities that cannot be uncovered by standard static analyses. Our design choices were driven towards achieving portability, both in terms of targeted architecture, but also in terms of programming environment (e.g., being robust to arbitrary programming language, compiler, use of third-party binaries, etc.).

A compact and yet precise inter-procedural dynamic dependence graph (DDG) is first computed via: 1. a new instrumentation framework based on QEMU; 2. the use of a new concept of inter-procedural loop-nesting tree; 3. followed by new techniques we introduce for folding, clamping, and widening of the DDG to agglomerate dynamic dependence instances into polyhedra of integer points whenever possible. State-of-the-art polyhedral analysis and transformation systems we specifically modified to provide useful feedback to the user is then used. We extensively evaluate our tool on numerous benchmarks, demonstrating the pratical usefulness of our tool-chain.

This work is the fruit of the collaboration 8.3.1.1 with OSU and and the past collaboration Nano2017 with STMicroelectronics. The main contributions has been presented at the ACM conference on Principles and Practice of Parallel Programming, PPoPP 2019 [9]. The new techniques that allow to build the polyhedral representation from the instrumented execution in a scalable way lead to a separate publication in the ACM Transactions on Architecture and Code Optimization, TACO 2020 [7].

6.2. Extraction of Periodic Patterns of Scientific Applications to Identify DVFS Opportunities

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 developed a tool suite resorting to fine-grain profiling and periodicity analysis to identify optimization opportunities for both performance and power-efficiency. It leverages the fact that a large share of HPC parallel applications are constituted of a restrained set of compute kernels executed a huge number of times to extract periodic patterns representative of the aforementioned kernels. By doing so, our tool offers a simple and condensed proxy to analyze and predict the behavior of complex parallel applications. For instance, we were able to identify and extract periodic patterns for a panel of reference HPC applications such as NAMD and NEMO. Then, as an example of the many ways to exploit the aforementioned extracted periodic patterns, we evaluated the impact of the CPU frequency on the latter. As a result, we were able to identify DVFS opportunities we plan to exploit in a future work.

6.3. Runtime Monitoring, Verification, and Enforcement

Participants: Antoine El-Hokayem [Univ. Grenoble Alpes, Verimag], Yliès Falcone, Thierry Jéron [Inria Rennes], Ali Kassem, Hervé Marchand [Inria Rennes], Srinivas Pinisetty [IIT Bhubaneswar], Matthieu

Renard [Foxi], Antoine Rollet [Université de Bordeaux], César Sànchez [IMDEA Madrid], Gerardo Schneider [University of Gothenborg].

Our contributions in the domain of runtime monitoring, verification, and enforcement are threefold. First, we contributed to the publication of general papers aimed to structure the community by publishing a tutorial on runtime enforcement of timed properties [16], a review of the first five years of the competition on runtime verification [15] and a survey of future challenges of runtime verification [6]. We also concluded some other previous work by realizing journal publications on the topics of decentralized runtime verification [3] and on runtime enforcement of timed properties [5]. We started a new activity on monitoring for security properties, and more particularly on the detection of fault-injection attacks [12].

6.3.1. On the Runtime Enforcement of Timed Properties

This work [16] is concerned with runtime enforcement which refers to the theories, techniques, and tools for enforcing correct behavior of systems at runtime. We are interested in such behaviors described by specifications that feature timing constraints formalized in what is generally referred to as timed properties. This tutorial presents a gentle introduction to runtime enforcement (of timed properties). First, we present a taxonomy of the main principles and concepts involved in runtime enforcement. Then, we give a brief overview of a line of research on theoretical runtime enforcement where timed properties are described by timed automata and feature uncontrollable events. Then, we mention some tools capable of runtime enforcement, and we present the TiPEX tool dedicated to timed properties. Finally, we present some open challenges and avenues for future work.

6.3.2. Detecting Fault Injection Attacks with Runtime Verification

This work [12] is concerned with fault injections which are increasingly used to attack/test secure applications. In this paper, we define formal models of runtime monitors that can detect fault injections that result in test inversion attacks and arbitrary jumps in the control flow. Runtime verification monitors offer several advantages. The code implementing a monitor is small compared to the entire application code. Monitors have a formal semantics; and we prove that they effectively detect attacks. Each monitor is a module dedicated to detecting an attack and can be deployed as needed to secure the application. A monitor can run separately from the application or it can be weaved inside the application. Our monitors have been validated by detecting simulated attacks on a program that verifies a user PIN.

6.3.3. International Competition on Runtime Verification (CRV)

In this work [15], we review the first five years of the international Competition on Runtime Verification (CRV), which began in 2014. Runtime verification focuses on verifying system executions directly and is a useful lightweight technique to complement static verification techniques. The competition has gone through a number of changes since its introduction, which we highlight in this paper.

6.3.4. A Survey of Challenges for Runtime Verification from Advanced Application Domains (beyond software)

In this work [6], we survey the future challenges for runtime verification. Typically, the two main activities in runtime verification efforts are the process of creating monitors from specifications, and the algorithms for the evaluation of traces against the generated monitors. Other activities involve the instrumentation of the system to generate the trace and the communication between the system under analysis and the monitor. Most of the applications in runtime verification have been focused on the dynamic analysis of software, even though there are many more potential applications to other computational devices and target systems. In this paper we present a collection of challenges for runtime verification extracted from concrete application domains, focusing on the difficulties that must be overcome to tackle these specific challenges. The computational models that characterize these domains require to devise new techniques beyond the current state of the art in runtime verification.

6.3.5. On the Monitoring of Decentralized Specifications Semantics, Properties, Analysis, and Simulation

In this work [3], we define two complementary approaches to monitor decentralized systems. The first relies on those with a centralized specification, i.e, when the specification is written for the behavior of the entire system. To do so, our approach introduces a data-structure that i) keeps track of the execution of an automaton, ii) has predictable parameters and size, and iii) guarantees strong eventual consistency. The second approach defines decentralized specifications wherein multiple specifications are provided for separate parts of the system. We study two properties of decentralized specifications pertaining to monitorability and compatibility between specification and architecture. We also present a general algorithm for monitoring decentralized specifications. We map three existing algorithms to our approaches and provide a framework for analyzing their behavior. Furthermore, we introduce THEMIS, a framework for designing such decentralized algorithms and simulating their behavior. We show the usage of THEMIS to compare multiple algorithms and verify the trends predicted by the analysis by studying two scenarios: a synthetic benchmark and a real example.

6.3.6. Optimal Enforcement of (timed) Properties with Uncontrollable Events

This work deals with runtime enforcement of untimed and timed properties with uncontrollable events [5]. Runtime enforcement consists in defining and using mechanisms that modify the executions of a running system to ensure their correctness with respect to a desired property. We introduce a framework that takes as input any regular (timed) property described by a deterministic automaton over an alphabet of events, with some of these events being uncontrollable. An uncontrollable event cannot be delayed nor intercepted by an enforcement mechanism. Enforcement mechanisms should satisfy important properties, namely soundness, compliance and optimality – meaning that enforcement mechanisms should output as soon as possible correct executions that are as close as possible to the input execution. We define the conditions for a property to be enforceable with uncontrollable events. Moreover, we synthesise sound, compliant and optimal descriptions of runtime enforcement mechanisms at two levels of abstraction to facilitate their design and implementation.

6.4. Teaching of Algorithms, Programming, Debugging, and Automata

Participants: Florent Bouchez Tichadou, YlIès Falcone, Théo Barollet, Antoine Clavel, Thomas Hervé, Anthony Martinez, Beryl Piasentin, Steven Sengchanh.

This domain is a new axis of the Corse team. Our goal here is to combine our expertise in compilation and teaching to help teachers and learners in computer science fields such as programming, algorithms, data strucures, automata, or more generally computing litteracy. The most important project in this regard is the automated generation and recommendation of exercises using artificial intelligence, a thesis that started this year. Other projects focus on tools to help learning through visualization (data structures, debugger, automata) or gamification (AppoLab), and are the source of many internships that give younger students experience in a research team.

6.4.1. AI4HI: Artificial Intelligence for Human Intelligence

In an ideal educative world, each learner would have access to individual pedagogical help, tailored to its needs. For instance, a tutor who could rapidly react to the questions, and propose pedagogical contents that match the learner's kills, and who could identify and work on his or her weaknesses. However, the real world imposes constraints that make this individual pedagogical help hard to achieve.

The goal of the AI4HI project is to combine the new advances in artificial intelligence with the team's skills in compilation and teaching to aid teaching through the automated generation and recommendation of exercises to learners. In particular, we target the teaching of programming and debugging to novices. This system will propose exercises that match the learners' needs and hence improve the learning, progression, and self-confidence of learners.

This projet has received an "Action Exploratoire" funding from Inria and Théo Barollet started his PhD this September so is still in its early stages.

6.4.2. AppoLab

Classical teaching of algorithms and low-level data structures is often tedious and unappealing to students. AppoLab is an online platform to engage students in their learning by including gamification in Problem-Based Learning. In its core, it is a server with scripted “exercises”. Students can communicate with the server manually, but ultimately they 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.

6.4.3. Data Structures and Program Visualization at Runtime

Debuggers are powerful tools to observe a program behaviour and find bugs but they have a hard learning curve. 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 developed a prototype, Moly, a GDB extension that explores a program runtime memory and analyze its data structures. It provides an interface with an external visualizer, Lotos, through a formatted output. Work has also started to include a tutorial on how to use GDB and these extensions.

6.4.4. Aude

Aude is a pedagogical software for manipulating, learning, and teaching finite state automata and the automata theory. It is used by the students in the second year of the bachelor in computer science at Univ. Grenoble Alpes. It allows students to get acquainted and autonomously work on the concepts involved in the theory of regular languages and automata.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. Atos/Bull

- Title: Static and dynamic approaches for the optimization of the energy consumption associated with applications of the High Performance Computing (HPC) field
- CORSE participants: François Broquedis, Frédéric Desprez, Mathieu Stoffel
- Partner: Atos/Bull
- Duration: February 2018 - February 2021
- Abstract: The purpose of this project is to dynamically improve the energy consumption of HPC applications on large-scale platforms. It relies on an adaptation of the CPU frequency at runtime, based on the analysis of hardware-related metrics to determine an *application profile*. This profile is then split into different *phases*, each of which being associated to a best CPU frequency, depending on its nature (CPU bound, memory bound, ...). This project is funding the PhD of Mathieu Stoffel, and the corresponding development is to be integrated into *Bull Dynamic Power Optimizer*, a software suite developed by Atos/Bull.

7.2. Bilateral Grants with Industry

7.2.1. ES3CAP

- Title: Embedded Smart Safe Secure Computing Autonomous Platform
- CORSE participants: Fabrice Rastello, Nicolas Tolenaere
- Duration: July 2018 - August 2021
- INRIA Partners: AOSTE, PARKAS, CHROMA

- Other Partners: Renault-Nissan, EasyMile, Safran E&D, MBDA, ANSYS/ESterel Technologies, Kronno-Safe, Prove & Run, Kalray, Prophesee, CEA
- Abstract: The objective of ES3CAP is to develop a tool-chain that targets multi- and many-core architectures for critical systems. In particular it should address the different challenges related to making existing critical systems solutions (heterogeneous, decentralized, single-core, single-task) match the industrial constraints targeted by Kalray's MPPA (MPPA, high-performance, real-time, safety, security). Considered applications are autonmous driving, drones, avionics, and defense. CORSE is involved in the optimization of machine learning algorithms for many-core architectures.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. HEAVEN Persyval Project

- Title: HEterogenous Architectures: Versatile Exploitation and programmiNg
- HEAVEN leaders: François Broquedis, Olivier Muller [TIMA lab]
- CORSE participants: François Broquedis, Frédéric Desprez, Georgios Christodoulis, Manuel Selva
- Duration: September 2015 - December 2019
- Abstract: The main objective of this project was to improve the accessibility of heterogeneous architectures comprising FPGA accelerators with portability and real experimentation in mind. The portability criterion allows application programmers to benefit from FPGA devices with only small modifications to their applications. It was achieved by extending a standard parallel programming environment already targeting heterogeneous architectures comprising CPUs and GPUs. During the project, we developed an operational prototype targeting Xilinx FPGAs. Experiments have been conducted using both matrix multiplication and Cholesky decomposition kernels. These experiments have shown the usability of the framework and its very low overhead. This framework opens the path for challenging questions regarding the scheduling of heterogeneous applications targeting FPGAs.

8.2. National Initiatives

8.2.1. 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.

8.3. International Initiatives

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

8.3.1.1. IOComplexity

Title: Automatic characterization of data movement complexity

International Partner (Institution - Laboratory - Researcher):

Ohio State University (United States). P. Sadayappan

Colorado State University (United States). Louis-Noël Pouchet

Start year: 2018

See also: https://team.inria.fr/corse/iocomplexity/

The goal of this project is to extend techniques for automatic characterization 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 characterization; 3. Performance model.

8.4. International Research Visitors

8.4.1. Visits to International Teams

8.4.1.1. Research Stays Abroad

- Fabrice Rastello visited the University of Utah to work with P. Sadayappan during the month of November. He worked on abstract simulation, and optimization of pattern specific programs.
- Nicolas Derumigny visited the University of Utah to work with P. Sadayappan during the month of November. He worked on abstract simulation.
- Nicolas Tollenaere visited the University of Utah to work with P. Sadayappan during the month of November. He worked on abstract simulation, and optimization of convolutions
- Theo Barollet visited the Colorado State University to work with Steve Kommrusch during the month of October. He worked on graph neural networks.
- Nicolas Tollenaere visited the university of Utah to work with P. Sadayappan during the month of August. He worked on optimizing packing and transposition of tensors.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. Member of the Organizing Committees

- Fabrice Rastello: Steering Committee ACM/IEEE CGO; Steering Committee “Journées française de la compilation”

9.1.2. Scientific Events: Selection

9.1.2.1. Member of the Conference Program Committees

- Frédéric Desprez: IPDPS, HPC, CCGRID, ParCo, ICA3PP, CloudCom
- François Broquedis: IPDPS, COMPAS
- Fabrice Rastello: ACM SIGPLAN/SIGBED LCTES 2019; IEEE/ACM SIGPLAN CGO 2020; IEEE/ACM SIGARCH PACT 2019; CADO 2019

- Yliès Falcone: AFADL'19, SHPCS'19, IFIP-ICTSS'19, RV'19, FDL'19, 4PAD'19, TASE'19, SAC-SVT'19
- Manuel Selva: ReCoSoC'19

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Frédéric Desprez: IEEE Transactions on Cloud Computing (associate editor)

9.1.3.2. Reviewer - Reviewing Activities

- Frédéric Desprez: ACM Transactions on Internet of Things, IEEE Access
- Fabrice Rastello: ACM TACO, ACM TOPLAS

9.1.4. 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

9.1.5. Scientific Expertise

- Frédéric Desprez: Genci: attribution heures de calcul CT6
- Frédéric Desprez: Groupe de travail "Cloud pour l'IA" d'Allistène
- Frédéric Desprez: Comité des sages IRIT
- Yliès Falcone: Review for ANR call for projects

9.1.6. Research Administration

- Frédéric Desprez: Deputy Scientific Director at INRIA
- Frédéric Desprez: Director of the GIS GRID5000
- Frédéric Desprez: Conseil Scientifique ESIEE Paris
- Frédéric Desprez: Groupe de travail "Infrastructures" Inria

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

License 3: François Broquedis, Imperative programming using python, 40 hours, Grenoble Institute of Technology (Ensimag)

License 3: François Broquedis, Introduction to UNIX, 20 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, 30 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, Object-Oriented Programming, 20 hours, Grenoble Institute of Technology (Ensimag)

François Broquedis is in responsible of the first year of Ensimag

Master: Florent Bouchez Tichadou, Algorithmic Problem Solving, 41 hours, M1 MoSIG

Licence: Florent Bouchez Tichadou, Algorithms languages and programming, 160 hours, L2 UGA

Licence: Florent Bouchez Tichadou was responsible of the second year (L2) of INF (informatique) and MIN (mathématiques et informatique) students at UGA from January to June, eq. 42 hours

Licence: Florent Bouchez Tichadou, Software Project, 10 hours, L3 UGA

Data Asperger, Florent Bouchez Tichadou, Formation d'autistes Asperger aux métiers du développement informatique et analyse des données. Bloc Algorithmique. 20 hours, GEM & Le Campus Numérique.

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, was co-responsible of the first year of the International Master of Computer Science (Univ. Grenoble Alpes and INP ENSIMAG) until August 2019.

License 3: Manuel Selva, Imperative programming using python, 60 hours, Grenoble Institute of Technology (Ensimag)

License 3: Manuel Selva, Introduction to UNIX, 15 hours, Grenoble Institute of Technology (Ensimag)

Master 1: Manuel Selva, Operating systems and concurrent programming, 15 hours, Grenoble Institute of Technology (Ensimag)

9.2.2. Supervision

PhD: Georgios Christodoulis, Adaptation of a heterogeneous run-time system to efficiently exploit FPGA, December 5, 2019, 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: Fabian Grüber, Performance Debugging Toolbox for Binaries: Sensitivity Analysis and Dependence Profiling, December 17 2019, advised by Fabrice Rastello

PhD: Raphaël Jakse, Interactive Runtime Verification, December 18 2019, advised by Jean-François Méhaut and Yliès Falcone.

PhD in progress: Auguste Olivry, Data Locality and Parallelism Optimization for Linear and Multilinear Algebra, September 2019, advised by Fabrice Rastello.

PhD in progress: Nicolas Tollenaere, Optimizing ML algorithms for MPPA Asics, April 2019, advised by Fabrice Rastello.

PhD in progress: Nicolas Derumigny, Automatic generation of performance models for heterogeneous architectures, September 2019, advised by Fabrice Rastello.

PhD in progress: Théo Barollet, Problem-based learning: automatic generation and recommendation of programming exercises, September 2019, advised by Florent Bouchez Tichadou and Fabrice Rastello.

9.2.3. Juries

9.2.3.1. Fabrice Rastello

Fabian Grüber, advisor, *Performance Debugging Toolbox for Binaries: Sensitivity Analysis and Dependence Profiling*, PhD, Université Grenoble Alpes, December 17 2019

Maxime Schmitt, reviewer, *Automatic Generation of Adaptative Codes*, PhD, Université de Strasbourg, September 30 2019

9.2.3.2. Frédéric Desprez

Michael Mercier, examiner, *Contribution to High Performance Computing and Big Data Infrastructure Convergence*, PhD, Université Grenoble Alpes, January 07 2019

Alexandre Veith, examiner, *Quality of Service Aware Mechanisms for (Re)Configuring Data Stream Processing Applications on Highly Distributed Infrastructure*, PhD, Lyon, September 23 2019

Nathanaël Cheriere, reviewer, *Towards Malleable Distributed Storage Systems From Models to Practice*, PhD, Ecole normale supérieure de Rennes, November 5 2019

Ayham Kassab, reviewer/president, *Optimisation de l'ordonnancement de calculs parallèles et de l'engagement de sources d'énergie renouvelable pour l'alimentation des centres de calcul*, PhD, Université de Bourgogne Franche Comté, November 14 2019

Georgios Christodoulis, advisor, *Adaptation of a heterogeneous run-time system to efficiently exploit FPGA*, PhD, Université Grenoble Alpes, December 5, 2019

9.2.3.3. François Broquedis

- Georgios Christodoulis, advisor, *Adaptation of a heterogeneous run-time system to efficiently exploit FPGA*, PhD, Université Grenoble Alpes, December 5, 2019

9.2.3.4. Yliès Falcone

- Raphaël Jakse, advisor, *Interactive Runtime Verification*, PhD, Université Grenoble Alpes, December 18 2019.

9.3. Popularization

9.3.1. Education

DIU EIL, Florent Bouchez Tichadou, Formation des enseignants des lycées suite à la réforme du Bac et l'introduction de l'option informatique en 1ère et Terminale (NSI). Bloc algorithmique par Apprentissage Par Problèmes (APP). Académie de Grenoble ainsi que les enseignants dans les lycées français à l'étranger.

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] R. JAKSE.*Interactive runtime verification*, CORSE - Compiler Optimization and Run-time Systems ; Université Grenoble - Alpes ; LIG (Laboratoire informatique de Grenoble) ; Inria Grenoble Rhône-Alpes, December 2019, https://hal.inria.fr/tel-02460734

Articles in International Peer-Reviewed Journal

[2] E. CRUZ, M. DIENER, L. LIMA PILLA, P. NAVAUX.*EagerMap: A Task Mapping Algorithm to Improve Communication and Load Balancing in Clusters of Multicore Systems*, in "ACM Transactions on Parallel Computing", 2019, vol. 5, n^{o} 4, 17 [*DOI :* 10.1145/3309711], https://hal.archives-ouvertes.fr/hal-02062952

[3] A. EL-HOKAYEM, Y. FALCONE.*On the Monitoring of Decentralized Specifications: Semantics, Properties, Analysis, and Simulation*, in "ACM Transactions on Software Engineering and Methodology", September 2019, p. 1-57, https://hal.archives-ouvertes.fr/hal-02283429

[4] P. H. PENNA, A. T. A. GOMES, M. CASTRO, P. PLENTZ, H. C. D. FREITAS, F. BROQUEDIS, J.-F. MÉHAUT.*A Comprehensive Performance Evaluation of the BinLPT Workload-Aware Loop Scheduler*, in "Concurrency and Computation: Practice and Experience", February 2019, vol. 31, n^{o} 18, p. 1-22 [*DOI :* 10.1002/CPE.5170], 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] C. SANCHEZ, G. SCHNEIDER, W. AHRENDT, E. BARTOCCI, D. BIANCULLI, C. COLOMBO, Y. FALCONE, A. FRANCALANZA, S. KRSTIĆ, J. LOURENÇO, D. NICKOVIC, G. PACE, J. RUFINO, J. SIGNOLES, D. TRAYTEL, A. WEISS.*A survey of challenges for runtime verification from advanced application domains (beyond software)*, in "Formal Methods in System Design", November 2019, vol. 54, n^{o} 3, p. 279-335 [*DOI :* 10.1007/S10703-019-00337-W], https://hal.archives-ouvertes.fr/hal-02433558

[7] M. SELVA, F. GRUBER, D. SAMPAIO, C. GUILLON, L.-N. POUCHET, F. RASTELLO.*Building a Polyhedral Representation from an Instrumented Execution: Making Dynamic Analyses of Non-Affine Programs Scalable*, in "ACM Transactions on Architecture and Code Optimization", December 2019, vol. 16, n^{o} 4, p. 1-26 [*DOI :* 10.1145/3363785], https://hal.inria.fr/hal-02418987

International Conferences with Proceedings

[8] F. AIT SALAHT, F. DESPREZ, A. LEBRE, C. PRUD'HOMME, M. ABDERRAHIM.*Service Placement in Fog Computing Using Constraint Programming*, in "SCC 2019 : IEEE International Conference on Services Computing", Milan, Italy, IEEE, July 2019, p. 19-27 [*DOI :* 10.1109/SCC.2019.00017], https://hal.archives-ouvertes.fr/hal-02108806

[9] F. GRUBER, M. SELVA, D. SAMPAIO, C. GUILLON, A. MOYNAULT, L.-N. POUCHET, F. RASTELLO.*Data-Flow/Dependence Profiling for Structured Transformations*, in "PPoPP 2019 - 24th Symposium on Principles and Practice of Parallel Programming", Washington, D.C., United States, ACM, February 2019, p. 173-185 [*DOI :* 10.1145/3293883.3295737], https://hal.inria.fr/hal-02060796

[10] R. LI, A. SUKUMARAN-RAJAM, R. VERAS, T. M. LOW, F. RASTELLO, A. ROUNTEV, P. SADAYAPPAN.*Analytical Cache Modeling and Tilesize Optimization for Tensor Contractions*, in "SC 2019 - International Conference for High Performance Computing, Networking, Storage and Analysis", Denver, United States, ACM Press, November 2019, p. 1-13 [*DOI :* 10.1145/3295500.3356218], https://hal.inria.fr/hal-02418875

[11] F. ZANON BOITO, J.-F. MÉHAUT, T. DEUTSCH, B. VIDEAU, F. DESPREZ.*Instrumental Data Management and Scientific Workflow Execution: the CEA case study*, in "IPDPSW 2019 - International Parallel and Distributed Processing Symposium Workshops", Rio de Janeiro, Brazil, IEEE, May 2019, p. 850-857 [*DOI :* 10.1109/IPDPSW.2019.00139], https://hal.inria.fr/hal-02076963

Conferences without Proceedings

[12] A. KASSEM, Y. FALCONE.*Detecting Fault Injection Attacks with Runtime Verification*, in "SPRO 2019 - 3rd International Workshop on Software PROtection", Londres, United Kingdom, ACM, November 2019, p. 65-76 [*DOI :* 10.1145/3338503.3357724], https://hal.archives-ouvertes.fr/hal-02283434

[13] P. H. PENNA, J. SOUTO, D. F. LIMA, M. CASTRO, F. BROQUEDIS, H. H. FREITAS, J.-F. MÉHAUT.*On the Performance and Isolation of Asymmetric Microkernel Design for Lightweight Manycores*, in "SBESC 2019 - IX Brazilian Symposium on Computing Systems Engineering", Natal, Brazil, November 2019, p. 1-31, https://hal.archives-ouvertes.fr/hal-02297637

[14] P. H. PENNA, M. SOUZA, E. P. JUNIOR, J. SOUTO, M. CASTRO, F. BROQUEDIS, H. COTA DE FREITAS, J.-F. MÉHAUT.*RMem: An OS Service for Transparent Remote Memory Access in Lightweight Manycores*, in "MultiProg 2019 - 25th International Workshop on Programmability and Architectures for Heterogeneous Multicores", Valencia, Spain, High-Performance and Embedded Architectures and Compilers Workshops (HiPEAC Workshops), January 2019, p. 1-16, https://hal.archives-ouvertes.fr/hal-01986366

Scientific Books (or Scientific Book chapters)

[15] E. BARTOCCI, Y. FALCONE, G. REGER.*International Competition on Runtime Verification (CRV)*, in "Tools and Algorithms for the Construction and Analysis of Systems. TACAS 2019", Springer, April 2019, p. 41-49 [*DOI :* 10.1007/978-3-030-17502-3_3], https://hal.archives-ouvertes.fr/hal-02433552

[16] Y. FALCONE, S. PINISETTY.*On the Runtime Enforcement of Timed Properties*, in "Proceedings of the Runtime Verification 2019 conference", Springer, October 2019, p. 48-69 [*DOI :* 10.1007/978-3-030-32079-9_4], https://hal.archives-ouvertes.fr/hal-02433521

Research Reports

[17] F. AIT SALAHT, F. DESPREZ, A. LEBRE.*An overview of service placement problem in Fog and Edge Computing*, Univ Lyon, EnsL, UCBL, CNRS, Inria, LIP, LYON, France, October 2019, n° RR-9295, p. 1-43, https://hal.inria.fr/hal-02313711

[18] 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, p. 1-24, https://hal.inria.fr/hal-01967828

[19] R. JAKSE, Y. FALCONE, J.-F. MÉHAUT.*Interactive Runtime Verification: Formal Models, Algorithms, and Implementation*, UGA (Université Grenoble Alpes) ; LIG (Laboratoire informatique de Grenoble) ; Inria Grenoble Rhône-Alpes, Université de Grenoble, July 2019, https://hal.inria.fr/hal-02190656

Other Publications

[20] A. OLIVRY, J. LANGOU, L.-N. POUCHET, P. SADAYAPPAN, F. RASTELLO.*Automated Derivation of Parametric Data Movement Lower Bounds for Affine Programs*, December 2019, https://arxiv.org/abs/1911.06664 - working paper or preprint, https://hal.inria.fr/hal-02421026

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
A1.1.5. - Exascale
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
A2.6.4. - Ressource management
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

Raphaël Bleuse [Univ Grenoble Alpes, Associate Professor, from Sep 2019]
Gwenaël Delaval [Univ Grenoble Alpes, Associate Professor]
Stéphane Mocanu [Institut polytechnique de Grenoble, Associate Professor]

External Collaborators
Soguy Mak Kare Gueye [Atos, until Jul 2019]
Bogdan Robu [Univ Grenoble Alpes]

Technical Staff
Jan Pristas [Institut polytechnique de Grenoble, Engineer, from Feb 2019 until Jul 2019]

PhD Student
Neil Ayeb [Orange Labs, PhD Student]

Post-Doctoral Fellows
Quang Pham Tran Anh [Inria/Nokia, co-advised EPI Dyonisos, Rennes, until nov. 2019]
Hamza Sahli [Inria, co-advised EPI Stack, Nantes, from jan. 2019]

Administrative Assistant
Maria Immaculada Presseguer [Inria, Administrative Assistant]

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) [7], 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 [5], and of domain specific languages, following e.g., a component-based approach [4], [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], 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 [3], multi-core HPC architectures [10], Dynamic Partial Reconfiguration in FPGA-based hardware [2] and the IoT and smart environments [8].

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. Highlights of the Year

5.1. Highlights of the Year

The Ctrl-A team co-organised, in cooperation with the Gipsa-lab laboratory, the 40th International Summer School in Grenoble, with a special topic on Control of Computing Systems, on 9-13th of September 2019.

Invited speakers were international specialists of the field, from USA, Europe and France. Full information and programme are available : http://www.gipsa-lab.fr/summerschool/auto2019

6. New Software and Platforms

6.1. 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

7. New Results

7.1. Programming support for Autonomic Computing

7.1.1. Reactive languages

Participants: Gwenaël Delaval, Lucie Muller, 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 [5].

An ongoing topic is 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).

Recent work concerns compilation and diagnosis for discrete controller synthesis. The compilation involving a phase of controller synthesis can fail to find a solution, if the problem is overconstrained. The compiler does notify so to the programmer, but the latter would need a diagnosis in order to understand where and how to debug the program. Such diagnosis is made especially difficult by the declarative nature of the synthesis.

This was the object of the M1 TER internship of Lucie Muller [19].

7.1.2. 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 7.1.1 [4]. 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 also proposed a DSL called Ctrl-DPR [6], allowing designers to easily generate Autonomic Managers for DPR FPGA systems (see Section 7.2.3).

Ongoing work involves a generalization from our past experiences in software components, DPR FPGA, as well as IoT [8], and Cyberphysical Systems. As we observed a similarity in objects and structures (e.g., tasks, implementation 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 7.1.1. This direction will also lead us to study the definition of software architecture patterns for multiple loops Autonomic Managers, particularly hierarchical, with lower layers autonomy alleviating management burden from the upper layers as in Section 7.2.

7.2. 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 [7], [9].

7.2.1. Self-adaptative distributed systems

Participants: Quang Pham Tran Anh, Eric Rutten, Hamza Sahli.

Complex Autonomic Computing Systems, as found typically in distributed systems, must involve multiple management loops, addressing different subproblems of the general management, and using different modeling, decision and control approaches (discrete [3], continuous, stochastic, machine-learning based, ...) They are generally addressing deployment and allocation of computations on resources w.r.t. QoS, load, faults, ... but following different, complementary approaches. The similarities and recurring patterns are considered as in Section 7.1.2 . Their execution needs to be distributed w.r.t. different characteristics such as latency (as in Fog and Edge Computing) or load. We are studying Software Architectures to address the design of such complex systems.

7.2.1.1. Self-adaptation of micro-services in Fog/Edge and Cloud computing

Fog systems are a recent trend of distributed computing having vastly ubiquitous architectures and distinct requirements making their design difficult and complex. Fog computing is based on leveraging both resource-scarce computing nodes around the Edge to perform latency and delay sensitive tasks and Cloud servers for the more intensive computation.

In this work, we present a formal model defining spatial and structural aspects of Fog-based systems using Bigraphical Reactive Systems, a fully graphical process algebraic formalism. The model is extended with reaction rules to represent the dynamic behavior of Fog systems in terms of self-adaptation. The notion of bigraph patterns is used in conjunction with boolean and temporal operators to encode spatio-temporal properties inherent to Fog systems and applications. The feasibility of the modelling approach is demonstrated via a motivating case study and various self-adaptation scenarios.

This work is done in cooperation with the Inria team Stack in Nantes, and published in the FOCLASA workshop, co-located with the SFEM conference [13].

7.2.1.2. Autonomic management in Software Defined Networks

In the framework of our cooperation with Nokia Bell-labs (See Section 8.1.2), and the Dyonisos team at Inria Rennes, we are considering the management of Software Defined Networks (SDN), involving Data-Centers and accelerators.

The main approach AI / Machine Learning approaches, developed in Rennes. An ongoing topic is to consider that these reinforcement learning based approaches involve questions of trust, and we are beginning to consider their composition with controllers based e.g. on Control Theory, in order to maintain guarantees on the behaviors of the managed system.

7.2.2. High-Performance Grid Computing

Cloud and HPC (High-Performance Computing) systems have increasingly become more varying in their behavior, in particular in aspects such as performance and power consumption, and the fact that they are becoming less predictable demands more runtime management [10].

7.2.2.1. A Control-Theory based approach to minimize cluster underuse

Participants: Abdul Hafeez Ali, Raphaël Bleuse, Bogdan Robu, Eric Rutten.

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. In this work, we consider autonomic administration in HPC systems for scientific workflows management through a control theoretical approach. We propose a model described by parameters related to the key aspects of the infrastructure thus achieving a deterministic dynamical representation that covers the diverse and time-varying behaviors of the real computing system. We propose a model-predictive control loop to achieve two different objectives: maximize cluster utilization by best-effort jobs and control the file server's load in the presence of external disturbances. The accuracy of the prediction relies on a parameter estimation scheme based on the EKF (Extended Kalman Filter) to adjust the predictive-model to the real system, making the approach adaptive to parametric variations in the infrastructure. The closed loop strategy shows performance improvement and consequently a reduction in the total computation time. The problem is addressed in a general way, to allow the implementation on similar HPC platforms, as well as scalability to different infrastructures.

This work is done in cooperation with the Datamove team of Inria/LIG, and Gipsa-lab. Some results were published in the CCTA conference [14]. It was the topic of the Master's thesis of Abdul Hafeez Ali [16].

7.2.2.2. Combining Scheduling and Autonomic Computing for Parallel Computing Resource Management

Participants: Raphaël Bleuse, Eric Rutten.

This research topic aims at studying the relationships between scheduling and autonomic computing techniques to manage resources for parallel computing platforms. The performance of such platforms has greatly improved (149 petaflops as of November 2019 [20]) at the cost of a greater complexity: the platforms now contain several millions of computing units. While these computation units are diverse, one has to consider other constraints such as the amount of free memory, the available bandwidth, or the energetic envelope. The variety of resources to manage builds complexity up on its own. For example, the performance of the platforms depends on the sequencing of the operations, the structure (or lack thereof) of the processed data, or the combination of application running simultaneously.

Scheduling techniques offer great tools to study/guaranty performances of the platforms, but they often rely on complex modeling of the platforms. They furthermore face scaling difficulties to match the complexity of new platforms. Autonomic computing manages the platform during runtime (on-line) in order to respond to the variability. This approach is structured around the concept of feedback loops.

The scheduling community has studied techniques relying on autonomic notions, but it has failed to link the notions up. We are starting to address this topic.

7.2.3. High-Performance Embedded Computing

Participants: Soguy Mak Kare Gueye, Stéphane Mocanu, Eric Rutten.

This topics build upon our experience in reconfiguration control in DPR FPGA [2].

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. This work is in the framework of the ANR project HPeC (see Section 9.2.1).

7.2.3.1. DPR FPGA and discrete control for reconfiguration

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.

Ongoing work concerns experimental validation, where upon the availability of hardware implementations of vision, detection and tracking tasks, a demonstrator is being built integrating our controller.

7.2.3.2. 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 of the PhD of Chabha Hireche, defended in nov. 2019 [17].

7.2.4. IoT and Cyberphysical Systems

Participants: Neil Ayeb, Ayan Hore, Fabien Lefevre, Stéphane Mocanu, Jan Pristas, Eric Rutten, Gaetan Sorin, Mohsen Zargarani.

7.2.4.1. 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 (DM) 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.

Device Management is currently industrially deployed for LAN devices, phones and workstation management. Internet of Things (IoT) devices are massive, dynamic, heterogeneous, and inter-operable. Existing solutions are not suitable for IoT management. This work in an industrial environment addresses these limitations with a novel autonomic and distributed approach for the DM.

This work is in the framework of the Inria/Orange labs joint laboratory (see Section 8.1.1), and supported by the CIFRE PhD thesis grant of Neïl Ayeb, starting dec. 2017. It was awarded a best paper distinction at the Doctoral Symposium of ICAC 2019 [12].

7.2.4.2. Security in SCADA industrial systems

We focus mainly on vulnerability search, automatic attack vectors synthesis and intrusion detection [11]. 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 has been 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 9.2.2).

One of important results is the realization of a Hardware-in-the-loop SCADA Cyberange based on a electronic interface card that allows to interface real-world PLC with a software simulation [21]. The entire system is available in open-source including the electronic card fabrication files (http://gics-hil.gforge.inria.fr/). Interfacing system allow connection with various commercial simulation software but also with “home made” simulators [15]. The work is also supported by Grenoble Alpes Cybersecurity Institute (see Section 9.1.1) and Pulse program of IRT NANOELEC.

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. A first approach was developed in the M2R internship by Ayan Hore [18]

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

8.1.1. Orange

We have a cooperation with Orange labs, around a CIFRE PhD grant, on the topic of autonomic device management (see Section 7.2.4.1). This activity is part of the Inria/Orange joint laboratory.

8.1.2. Nokia / Bell labs

We have 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 Autonomic management in Software Defined Networks. This activity is part of the Inria/ Nokia / Bell labs joint laboratory.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. 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 currently funding two “alternance” student positions and a PhD position might be provided in September 2020 and supervised by Stephane Mocanu.

9.2. National Initiatives

9.2.1. 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 has been 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.

9.2.2. 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.

9.2.3. IRT Nanoelec Pulse program

The Pulse program aims the development for SCADA cybersecurity demonstrators. It has funded a Master grant in 2019 and two master grants in 2020. A PhD position was also approved for September 2020 and it will be co-supervised by Stéphane Mocanu

9.2.4. 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).

9.2.5. 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).

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

Program: ECSEL

Project acronym: CPS4EU

Project title: Cyber Physical Systems for Europe

Duration: july 2019 - june 2022

Coordinator: VALEO

Other partners: 38 participants

Abstract: CPS4EU proposes to address technical issues and organizational issues in an integrated way. Hence, CPS4EU promotes a high level of sharing, so that an operational ecosystem, with adequate skills and expertise all along the value chain can enable, at the end of the project, the European industry to lead strategic markets based on CPS technologies.

In this project, the Ctrl-A team is involved in WP4 and WP9 mainly, on topics of Software Architectures for Self-Adaptive systems in CPS, and our main industrial collaboration is with RTE.

9.4. International Initiatives

9.4.1. 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 9th Workshop of the JLESC at Knoxville, TE, USA, in April 2019, and visited ANL in Chicago.

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

9.4.2. Inria International Partners

9.4.2.1. 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).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

Eric Rutten is co-chairing, with Bogdan Robu (Gipsa-lab), the program of the 40th Summer School of Automatic Control, Grenoble, September 2019, on the special topic of Control for Computing Systems.

http://www.gipsa-lab.fr/summerschool/auto2019

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

Eric Rutten is co-chairing, with Françoise Baude, Antonio Fileri, and Nicola Capodieci, the 5th International Workshop on Autonomic High Performance Computing (AHPC 2019), http://hpcs2019.cisedu.info/2-conference/workshops/workshop09-ahpc

as part of The 17th International Conference on High Performance Computing & Simulation (HPCS 2019), July 15 – 19, 2019 Dublin, Ireland, http://hpcs2019.cisedu.info/

10.1.2.2. Member of the Conference Program Committees

Eric Rutten is PC member for :

- international conferences
  - ICAC 2019 (16th IEEE International Conference on Autonomic Computing), Umeøa, Sweden, June 16-20, 2 019 (http://icac2019.cs.umu.se)
  - 17th High Performance Computing & Simulation Conference (HPCS 2019). July 15 – 19, 2019 Dublin, Ireland (http://hpcs2019.cisedu.info/)
  - 15th Workshop on Discrete Event Systems, WODES'20, Rio de Janeiro, May, 13th-15th, 2020. (https://wodes2020.eventos.ufrj.br)
  - CPS&IoT'2020 - the 8th International Conference on Cyber-Physical Systems and Internet-of-Things, Budva, Montenegro, June 8-11, 2020 (http://embeddedcomputing.me/en/cps-iot-2020)
  - (Associate Editor) 4th IEEE Conference on Control Technology and Applications, CCTA 2020, Montreal, August, 2020 (http://ccta2020.ieeecss.org/)
  - 7th International Conference on Control, Decision and Information Technologies, CoDIT'20, June 29 - July 02, 2020, Prague, Czech Republic (https://codit2020.com/)
  - MSR 19, Modélisation des Systèmes Réactifs 2019 LARIS Université d'Angers, France, du 13 au 15 novembre 2019 (http://msr2019.laris.univ-angers.fr/)

10.1.3. Journal

10.1.3.1. Reviewer - Reviewing Activities

Eric Rutten is reviewer for

- Science of Computer Programming
- journal of Discrete Event Dynamic System (jDEDS)

10.1.4. Invited Talks

Eric Rutten was invited to give a talk at :

- Workshop on Integrating HPC and FPGAs, December 11, 2018 in Okinawa, Japan ; co-held with the International Workshop on FPGA for HPC (IWFH), the Joint Laboratory on Extreme Scale Computing (JLESC), and the Field-Programmable Technology (FPT'18).

  http://www.fpt18.sakura.ne.jp/workshop.html

  https://collab.cels.anl.gov/display/HPCFPGA/HPC-FPGA
- 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
- Languages, Compilation, and Semantics LIP Seminar, 10th Edition, Lyon, 2019, March 21

  http://perso.ens-lyon.fr/christian.perez/journeelangages
- 15th Cloud Control Workshop, Sandhamn, Sweden June 11–13, 2019

http://cloudresearch.org/workshops/15th/

10.1.5. 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).

10.1.6. Scientific Expertise

Eric Rutten was reviewer for the ANR : *Appel à projets Générique* 2019, PRCI international

10.1.7. 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.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Licence : R. Bleuse, C language, 12h lab, L2, Univ. Grenoble Alpes

Licence : R. Bleuse, methodology of software development, 30h, L2, Univ. 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

10.2.2. Supervision

- HdR : Stéphane Mocanu ; Cyberdéfense des infrastructures critiques ; COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES ; January 2019.
- PhD : Chabha Hireche ; Etude et implémentation d'une approche probabiliste de contrôle de mission de drone autonome ; Université de Bretagne Occidentale ; nov. 2019 ; 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).

10.2.3. Juries

Stéphane Mocanu was member of the jury of the PhD defense of Chabha Hireche, Université de Bretagne Occidentale, nov. 2019 (examinator).

Eric Rutten was member of the following juries :

- habilitation (HdR) defense, Stéphane Mocanu, UNIVERSITÉ GRENOBLE ALPES, 16 jan. 2019 (examinator)
- PhD defense of Anis Mezni, INSA Lyon, 13 may 2019 (reviewer)
- PhD defense of Mete Özbaltan, U. Liverpool, 10 dec 2019 (reviewer)
- PhD of Romolo Marotta, Uni. Roma Sapienza, dec 2019 (reviewer)

10.3. Popularization

10.3.1. Interventions

- National events:

  During the 2019 edition of Fête de la Science, Raphaël Bleuse co-hosted with Clément Mommessin (DataMove team) six one-hour popularization sessions. These sessions introduce the concept of infinity, and detail key milestones towards its construction in the modern mathematics for high-school students.

11. 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° 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° 7, 18, https://hal.inria.fr/hal-01242853

[4] 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

[5] 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° 4, p. 385-418, http://dx.doi.org/10.1007/s10626-013-0163-5

[6] 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

[7] 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

[8] 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

[9] 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

[10] 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

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] S. MOCANU.*Cyberdéfense des infrastructures critiques*, COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES, January 2019, Habilitation à diriger des recherches, https://hal.archives-ouvertes.fr/tel-02418978

International Conferences with Proceedings

[12] N. AYEB, E. RUTTEN, S. BOLLE, T. COUPAYE, M. DOUET.*Towards an Autonomic and Distributed Device Management for the Internet of Things*, in "FAS*W 2019 - IEEE 4th International Workshops on Foundations and Applications of Self* Systems", Umea, Sweden, 16th IEEE International Conference on Autonomic Computing (ICAC), IEEE, June 2019, p. 246-248, Best paper award at the Doctoral Symposium of ICAC 2019 [*DOI :* 10.1109/FAS-W.2019.00065], https://hal.archives-ouvertes.fr/hal-02295409

[13] H. SAHLI, T. LEDOUX, E. RUTTEN.*Modeling Self-Adaptive Fog Systems Using Bigraphs*, in "FOCLASA 2019 - 17th International Workshop on coordination and Self-Adaptativeness of Sotware applications", Oslo, Norway, September 2019, p. 1-16, https://hal.inria.fr/hal-02271394

[14] A. G. YABO, B. ROBU, O. RICHARD, B. BZEZNIK, E. RUTTEN.*A control-theory approach for cluster autonomic management: maximizing usage while avoiding overload*, in "CCTA 2019 - 3rd IEEE Conference on Control Technology and Applications", Hong-Kong, China, Proceedings of the 2019 IEEE Conference on Control Technology and Applications (CCTA), IEEE, August 2019, p. 189-195, https://hal.archives-ouvertes.fr/hal-02294272

Conferences without Proceedings

[15] S. MOCANU, M. PUYS, P.-H. THEVENON.*An Open-Source Hardware-In-The-Loop Virtualization System for Cybersecurity Studies of SCADA Systems*, in "C&esar 2019 - Virtualization and Cybersecurity", Rennes, France, November 2019, p. 1-16, https://hal.archives-ouvertes.fr/hal-02371133

References in notes

[16] A. H. ALI.*Controlling Computer Cluster Overload: Intelligent Control Using Nonsupervised Learning*, Masters Thesis, M2R MISCIT, Université Grenoble Alpes, France, 2019

[17] C. HIRECHE.*Etude et implémentation sur SoC-FPGA d'une méthode probabiliste pour le contrôle de mission de véhicule autonome*, PhD Thesis, Université de Bretagne Occidentale, November 2019

[18] A. HORE.*Analysis and identification of reaction mechanisms for self-protection in cybersecurity of SCADA systems*, Masters Thesis, M2R MOSIG, Université Grenoble Alpes, France, 2019

[19] L. MULLER.*Recherches de diagnostics pour un outil de synthèse de contrôleurs*, TER Report, M1, Université Grenoble Alpes, France, 2019

[20] E. STROHMAIER, J. DONGARRA, H. SIMON, M. MEUER.*TOP500 list*, https://www.top500.org/lists/

[21] M. ZARGARANI.*Hardware in the loop simulation of physical processes: a Tennessee-Eastman-like chemical process*, Masters Thesis, M2R MISCIT, Université Grenoble Alpes, France, 2019

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]
Rémi Gribonval [Inria, Senior Researcher, from Aug 2019, HDR]
Philippe Nain [Inria, Senior Researcher, HDR]

Faculty Members

Thomas Begin [Univ de Claude Bernard, Associate Professor, HDR]
Anthony Busson [Univ de Claude Bernard, Professor]
Christophe Crespelle [Univ de Claude Bernard, Associate Professor, HDR]
Marion Foare [Ecole supérieure de chimie physique électronique de Lyon, Associate Professor]
Isabelle Guérin Lassous [Univ de Claude Bernard, Professor, HDR]
Márton Karsai [Ecole Normale Supérieure Lyon, Associate Professor, until Aug 2019, HDR]

External Collaborators

Eric Philippe Guichard [Ecole Normale Supérieure Lyon, Maitre de Conférence ENSSIB, HDR]
Márton Karsai [Professor ECU, Budapest, Hungary, since Sep 2019, HDR]

Technical Staff

Hakim Hadj Djilani [Inria, Engineer, from Sep 2019]
Sébastien Lerique [Inria, Engineer, from Nov 2019]

PhD Students

Lafdal Abdelwedoud [Ambassade de France en Mauritanie]
Amélie Barbe [Univ de Lyon]
Esteban Bautista Ruiz [CONACyT fellowship until Dec 2019]
Nour El Houda Bouzouita [Ecole Normale Supérieure Lyon]
Sicheng Dai [Univ de Lyon]
Gaetan Frusque [Ecole Normale Supérieure Lyon]
Rémy Grünblatt [Inria]
Jacob Levy Abitbol [Inria, until Nov 2019]
Marija Stojanova [Ecole Normale Supérieure Lyon]
Samuel Unicomb [Inria]

Post-Doctoral Fellows

Sébastien Lerique [Inria, until Oct 2019]
Mikhail Tsitsvero [Univ de Lyon, until Jun 2019]

Visiting Scientists

Alexandre Brandwajn [University of California, Santa Cruz, from Feb to Mar 2019]
Juan Pablo Astudillo [Universitat Politècnica de Catalunya, PhD, from Apr 2019 until Jul 2019]
Dorsaf Ghozlani [Visiting PhD from Ecole Nationale d'Ingénieurs de Tunis, from Jun 2019 until Aug 2019]
Jaqueline Oliveira [Visiting PhD from Brazil with a CAPES fellowship, from July 2019 to December 2019]
Laetitia Gauvin [Research Leader ISI foundation, Torino, Italy]

Administrative Assistant

Sophie Gérard [Inria]

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, Rémi Gribonval, Marion Foare, Márton Karsai.

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 [74].

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 [57]. 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 [69] and references therein. Other viewpoints can be found as well, including multi-resolution Markov models [77], Bayesian networks or distributed processing over sensor networks [68]. 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 [58].

The arrival of Rémi Gribonval in August 2019 brought a new dimension to the research program of this theme. Specialist of parsimonious representations of large data sets, R. Gribonval will develop at Dante a specific activity related to the sparsification of resources (computing and storage but also regarding the data volumes) in the context of machine and deep learning. This new orientation of Dante will be elaborated and fully integrated to the objectives of the future Inria project that will be proposed after Dante.

3.2. Theory and Structure of dynamic Networks

Participants: Christophe Crespelle, Anthony Busson, Márton Karsai, Éric Guichard.

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 [71]. 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 [71], [62], [78], [67], to describe the observed structural and temporal heterogeneities [55], [62], [56], to detect and measure emerging community structures [59], [75], [76], to see how the functionality of networks determines their evolving structure [66], and to determine what kinds of correlations play a role in their dynamics [63], [65], [70].

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 [73]. 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 [64] and its co-evolution with ongoing processes like spreading phenomena, synchronisation, evolution of consensus, random walk etc. [64], [72]. 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.

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 [60] 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 [61]. 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 9.2.8).

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

Dante is member of the following new projects accepted in 2019, for funding by ANR:

- **DataRedux** Big data reduction for predictive computational modelling. Consortium: Dante (ENS de Lyon), Centre Physique Théorique Marseille (CNRS), Pierre Louis Institute of Epidemiology and Public Health (INSERM).
- **Darling** Distributed adaptation and learning over graphs. Consortium: Observatoire Côte d'Azur (U. Nice), Dante & LP (ENS de Lyon), L2S (Centrale Supelec).
- **CCS 2021** Márton Karsai and IXXI obtained the right to organise Conference on Complex Systems in Lyon in 2021.

5.1.1. Awards

- Article [10] has been highlighted as the cover page article of the journal MDPI.

BEST PAPERS AWARDS :

[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 "PAM 2019 - 20th Passive and Active Measurements Conference", Puerto Varas, Chile, Springer, March 2019, p. 197-209, Best paper award. [*DOI :* 10.1007/978-3-030-15986-3_13], https://hal.inria.fr/hal-01970596

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: Frederic 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: Frederic 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 Vandaele
- 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, Rémi Gribonval, Marion Foare, Thomas Begin, Esteban Bautista Ruiz, Gaetan Frusque, Amélie Barbe, Mikhail Tsitsvero, Marija Stojanova, Márton Karsai, Sébastien Lerique, Jacobo Levy Abitbol.

7.1.1. L^γ -PageRank for Semi-Supervised Learning

Participants: Paulo Gonçalves, Esteban Bautista Ruiz.

PageRank for Semi-Supervised Learning has shown to leverage data structures and limited tagged examples to yield meaningful classification. Despite successes, classification performance can still be improved, particularly in cases of fuzzy graphs or unbalanced labeled data. To address such limitations, a novel approach based on powers of the Laplacian matrix L^γ ($\gamma > 0$), referred to as L^γ-PageRank, is proposed. Its theoretical study shows that it operates on signed graphs, where nodes belonging to one same class are more likely to share positive edges while nodes from different classes are more likely to be connected with negative edges. It is shown that by selecting an optimal γ, classification performance can be significantly enhanced. A procedure for the automated estimation of the optimal γ, from a unique observation of data, is devised and assessed. Experiments on several datasets demonstrate the effectiveness of both L^γ-PageRank classification and the optimal γ estimation. [11]

7.1.2. Designing Convex Combination of Graph Filters

Participant: Paulo Gonçalves.

In this work, we studied the problem of parametric modeling of network-structured signals with graph filters. Unlike the popular polynomial graph filters, which are based on a single graph shift operator, we considered convex combinations of graph shift operators particularly adapted to directed graphs. As the resulting modeling problem is not convex, we reformulated it as a convex optimization problem which can be solved efficiently. Experiments on real-world data structured by undirected and directed graphs were conducted. The results showed the effectiveness of this method compared to other methods reported in the literature. [18]

7.1.3. Optimal transport under regularity constraints for domain adaptation between graphs with attributes

Participants: Paulo Gonçalves, Amélie Barbe.

In this work, we addresses the problem of domain adaptation between two graphs by optimal transport. We aimed at benefiting from the knowledge of a labeled source graph to improve the classification of nodes in an unlabeled target graph. We focused on the setting where a set of features is associated to each node of the graphs. We proposed an original method that optimizes a transportation plan from the source to the target that *(i)* preserves the structures transported between the graphs and *(ii)* prevents the mapping from transporting two source nodes with different labels to the same destination. [30]

7.1.4. Sparse tensor dimensionality reduction with application to the clustering of functional connectivity in the brain

Participants: Paulo Gonçalves, Gaetan Frusque.

Functional connectivity (FC) is a graph-like data structure commonly used by neuroscientists to study the dynamic behaviour of the brain activity. However, these analyses rapidly become complex and time-consuming, as the number of connectivity components to be studied is quadratic with the number of electrodes. In our work, we addressed the problem of clustering FC into relevant ensembles of simultaneously activated components that reveal characteristic patterns of the epileptic seizures of a given patient. While $k-$means is certainly the most popular method for data clustering, it is known to perform badly on large dimensional data sets, and to be highly sensitive to noise. To overcome the co-called curse of dimensionality, we proposed a new tensor decomposition to reduce the size of the data set formed by FC time series recorded for several seizures, before applying k-means. Our contribution is twofold: First, we derived a method that we compared to the state of the art, emphasizing one variant that imposes sparsity constraints. Second, we conducted a real case study, applying the proposed sparse tensor decomposition to epileptic data in order to infer the functional connectivity graph dynamics corresponding to the different stages of an epileptic seizure. [31], [47]

7.1.5. Graph signal processing to model WLANs performances

Participants: Paulo Gonçalves, Thomas Begin, Marija Stojanova.

As WLANs have become part of our everyday life, there is an increasing need for more transmission capacity and wireless coverage. In response to this growing need, network administrators tend to intensify the deployment of Access Points (APs). However, if not correctly done, this AP densification may lead to badly planned and uncoordinated networks with sub-optimal use of the available resources. In this work, we propose a data-driven approach using graph signal processing and a set of input/output signals to capture the behavior of a WLAN and derive a predictive performance model. Given the simplicity and the novelty of the proposed model, we believe that its relative error of around 10-20% in modeling and 25% in prediction may represent a promising start for new approaches in the modeling of WLANs. [33]

7.1.6. Joint embedding of structure and features via graph convolutional networks

Participants: Márton Karsai, Sébastien Lerique.

We propose *AN2VEC*, a node embedding method 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 develop a multitask GCN Variational Autoencoder where different dimensions of the generated embeddings can be dedicated to encoding feature information, network structure, and shared feature-network information. We explore the interaction between these disentangled characters by comparing the embedding reconstruction performance to a baseline case where no shared information is extracted. We use synthetic datasets with different levels of interdependency between feature and network characters and show (i) that shallow embeddings relying on shared information perform better than the corresponding reference with unshared information, (ii) that this performance gap increases with the correlation between network and feature structure, and (iii) that our embedding is able to capture joint information of structure and features. Our method can be relevant for the analysis and prediction of any featured network structure ranging from online social systems to network medicine. [51]

7.2. Computational Human Dynamics and Temporal Networks

Participants: Márton Karsai, Sébastien Lerique, Jacobo Levy Abitbol, Samuel Unicomb, Sicheng Dai.

7.2.1. Optimal Proxy Selection for Socioeconomic Status Inference on Twitter

Participants: Márton Karsai, Jacobo Levy Abitbol.

The socioeconomic status of people depends on a combination of individual characteristics and environmental variables, thus its inference from online behavioral data is a difficult task. Attributes like user semantics in communication, habitat, occupation, or social network are all known to be determinant predictors of this feature. In this paper we propose three different data collection and combination methods to first estimate and, in turn, infer the socioeconomic status of French Twitter users from their online semantics. Our methods are based on open census data, crawled professional profiles, and remotely sensed, expert annotated information on living environment. Our inference models reach similar performance of earlier results with the advantage of relying on broadly available datasets and of providing a generalizable framework to estimate socioeconomic status of large numbers of Twitter users. These results may contribute to the scientific discussion on social stratification and inequalities, and may fuel several applications. [19]

7.2.2. Randomized reference models for temporal networks

Participant: Márton Karsai.

In this paper we propose a unified framework for classifying and understanding microcanonical RRMs (MRRMs). Focusing on temporal networks, we use this framework to build a taxonomy of MRRMs that proposes a canonical naming convention, classifies them, and deduces their effects on a range of important network features. We furthermore show that certain classes of compatible MRRMs may be applied in sequential composition to generate over a hundred new MRRMs from the existing ones surveyed in this article. We provide two tutorials showing applications of the MRRM framework to empirical temporal networks: 1) to analyze how different features of a network affect other features and 2) to analyze how such features affect a dynamic process in the network. We finally survey applications of MRRMs found in literature. [48]

7.2.3. Reentrant phase transitions in threshold driven contagion on multiplex networks

Participants: Márton Karsai, Samuel Unicomb.

Models of threshold driven contagion explain the cascading spread of information, behavior, systemic risk, and epidemics on social, financial and biological networks. At odds with empirical observation, these models predict that single-layer unweighted networks become resistant to global cascades after reaching sufficient connectivity. We investigate threshold driven contagion on weight heterogeneous multiplex networks and show that they can remain susceptible to global cascades at any level of connectivity, and with increasing edge density pass through alternating phases of stability and instability in the form of reentrant phase transitions of contagion. Our results provide a novel theoretical explanation for the observation of large scale contagion in highly connected but heterogeneous networks. [23]

7.2.4. Interactional and informational attention on Twitter

Twitter may be considered as a decentralized social information processing platform whose users constantly receive their followees' information feeds, which they may in turn dispatch to their followers. This decentralization is not devoid of hierarchy and heterogeneity, both in terms of activity and attention. In particular, we appraise the distribution of attention at the collective and individual level, which exhibits the existence of attentional constraints and focus effects. We observe that most users usually concentrate their attention on a limited core of peers and topics, and discuss the relationship between interactional and informational attention processes – all of which, we suggest, may be useful to refine influence models by enabling the consideration of differential attention likelihood depending on users, their activity levels and peers' positions. [10]

7.2.5. Efficient limited time reachability estimation in temporal networks

Participant: Márton Karsai.

In this paper we propose a probabilistic counting algorithm, which gives simultaneous and precise estimates of the in- and out-reachability (with any chosen waiting-time limit) for every starting event in a temporal network. Our method is scalable allowing measurements for temporal networks with hundreds of millions of events. This opens up the possibility to analyse reachability, spreading processes, and other dynamics in large temporal networks in completely new ways; to compute centralities based on global reachability for all events; or to find with high probability the exact node and time, which could lead to the largest epidemic outbreak. [52]

7.2.6. weg2vec: Event embedding for temporal networks

Participant: Márton Karsai.

Network embedding techniques are powerful to capture structural regularities in networks and to identify similarities between their local fabrics. However, conventional network embedding models are developed for static structures, commonly consider nodes only and they are seriously challenged when the network is varying in time. Temporal networks may provide an advantage in the description of real systems, but they code more complex information, which could be effectively represented only by a handful of methods so far. Here, we propose a new method of event embedding of temporal networks, called weg2vec, which builds on temporal and structural similarities of events to learn a low dimensional representation of a temporal network. This projection successfully captures latent structures and similarities between events involving different nodes at different times and provides ways to predict the final outcome of spreading processes unfolding on the temporal structure. [53]

7.3. Communication Networks

Participants: Thomas Begin, Anthony Busson, Isabelle Guérin Lassous, Marion Foare, Philippe Nain, Lafdal Abdelwedoud, Marija Stojanova, Rémy Grünblatt, Juan Pablo Astudillo.

7.3.1. Quantum communications

In [29] we investigate the performance of a quantum switch serving a set of users. The function of the switch is to convert bipartite entanglement generated over individual links connecting each user to the switch, into bipartite or tripartite entangled states among (pairs or groups of) users at the highest possible rates at a fixed ratio. Such entanglement can then be converted to quantum-secure shared secret bits among pairs or triples of users using E91-like Quantum Key Distribution (QKD) protocols. The switch can store a certain number of qubits in a quantum memory for a certain length of time, and can make two-qubit Bell-basis measurements or three-qubit GHZ-basis projective measurements on qubits held in the memory. We model a set of randomized switching policies. Discovering that some are better than others, we present analytical results for the case where the switch stores one qubit per user at a given time step, and find that the best policies outperform a time division multiplexing (TDM) policy for sharing the switch between bipartite and tripartite entanglement generation. This performance improvement decreases as the number of users grows. The model is easily augmented to study the capacity region in the presence of qubit decoherence, obtaining similar results. Moreover, decoherence appears to have little effect on capacity. We also study a smaller class of policies when the switch can store two qubits per user.

7.3.2. Resource Allocation

In [28] we consider assignment policies that allocate resources to users, where both resources and users are located on a one-dimensional line $[0, \infty\infty)$. First, we consider unidirectional assignment policies that allocate resources only to users located to their left. We propose the Move to Right (MTR) policy, which scans from left to right assigning nearest rightmost available resource to a user, and contrast it to the Unidirectional Gale-Shapley (UGS) matching policy. While both policies among all unidirectional policies minimize the expected distance traveled by a request (request distance), MTR is fairer. Moreover, we show that when user and resource locations are modeled by statistical point processes, and resources are allowed to satisfy more than one user, the spatial system under unidirectional policies can be mapped into bulk service queueing systems, thus allowing the application of many queueing theory results that yield closed-form expressions. As we consider a case where different resources can satisfy different numbers of users, we also generate new results for bulk service queues. We also consider bidirectional policies where there are no directional restrictions on resource allocation and develop an algorithm for computing the optimal assignment which is more efficient than known algorithms in the literature when there are more resources than users. Finally, numerical evaluation of performance of unidirectional and bidirectional allocation schemes yields design guidelines beneficial for resource placement.

7.3.3. VoD broadcasting over vehicular networks

Participants: Thomas Begin, Anthony Busson, Isabelle Guérin Lassous.

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. The network infrastructure comprises IEEE 802.11p RSUs (Road Side Units) that are deployed along the highway and deliver video content to traveling vehicles. In this paper, we propose a simple analytical and yet accurate solution to estimate two key performance parameters for a VoD platform: (i) the average download data rate experienced by vehicles over their journey and (ii) the average "interruption time", which corresponds to the fraction of time the video playback of a given vehicle is interrupted because of an empty buffer. Through multiple examples, we investigate the influence of several parameters (e.g., the video bit rate, the number of vehicles, the distance between RSUs, the vehicle velocity) on these two performance parameters whose outcome may help the sizing of an IEEE 802.11p-based VoD platform [12].

7.3.4. Performance Evaluation of Channel Bonding in IEEE 802.11ac

Participants: Thomas Begin, Anthony Busson, Marija Stojanova.

WLANs grow in popularity in home, public, and work environments, resulting in constantly increasing demands for wireless coverage and capacity. There exist two dominant strategies that help solve the problem of WLAN capacity: the deployment of more APs and enhancement of the standards in use. These policies result in

WLANs containing a larger number of more complex devices, making the prediction of the network's behavior an even more elaborate problem. Because of these issues, WLANs are prone to inefficient configurations. In this paper, we propose a Markovian continuous time model that aims at predicting the throughputs achieved by all the WLAN's APs as a function of the network's topology and the AP's throughput demands. By means of simulation, we show that our model achieves mean relative errors of less than 10% for networks of different sizes and with diverse node configurations. The model is adapted to the specificities of the IEEE 802.11ac standard amendment and can be used to solve problems such as channel assignment or channel bonding. We derive guidelines on the best practice in channel bonding given a performance metric and for different MCS indexes, frame aggregation rates, saturation levels, and network topologies. We then put our findings to the test by identifying the optimal channel bonding combination in a WLAN containing a diverse set of nodes.

7.3.5. Distributed Congestion Control mechanism for NANs

Participants: Thomas Begin, Anthony Busson, Juan Pablo Astudillo.

The need for significant improvements in the management and efficient use of electrical energy has led to the evolution from the traditional electrical infrastructures towards modern Smart Grid networks. Taking into account the critical importance of this type of networks, multiple research groups focus their work on issues related to the generation, transport and consumption of electrical energy. One of the key research points is the data communication network associated with the electricity transport infrastructure, and specifically the network that interconnects the devices in consumers' homes, the so-called Neighborhood Area Networks (NANs). In this paper, a new distributed congestion control mechanism is proposed, implemented and evaluated for NANs. Besides, different priorities have been considered for the traffic flows transmitted by different applications. The main goal is to provide with the needed Quality of Service (QoS) to all traffic flows, especially in high traffic load situations. The proposal is evaluated in the context of a wireless ad hoc network made up by a set of smart meter devices, using the Ad hoc On-Demand Distance Vector (AODV) routing protocol and the IEEE 802.11ac physical layer standard. The application of the proposed congestion control mechanism, together with the necessary modifications made to the AODV protocol, lead to performance improvements in terms of packet delivery ratio, network throughput and transit time, fairness between different traffic sources and QoS provision [35].

7.3.6. Simulation and Performance Evaluation of the Intel Rate Adaptation Algorithm

Participants: Rémy Grünblatt, Isabelle Guérin-Lassous.

With the rise of the complexity of the IEEE 802.11 standard, rate adaptation algorithms have to deal with a large set of values for all the different parameters having an impact on the network throughput. Simple trial-and-error algorithms can no longer explore solution space in reasonable time and smart solutions are required. Most of the WiFi controllers rely on proprietary code and the used rate adaptation algorithms in these controllers are unknown. Very few WiFi controllers expose their rate adaptation algorithms if they do not rely on the MINSTREL-HT algorithm which is implemented in the mac80211 component of the Linux kernel. Intel WiFi controllers come with their own rate adaptation algorithms that are implemented in the Intel IWLWIFI Linux Driver which is open-source.

In this work, we have reverse-engineered the Intel rate adaptation mechanism from the source code of the IWLWIFI Linux driver, and we give, in a comprehensive form, the underlying rate adaptation algorithm named IWL-MVM-RS. We describe the different mechanisms used to seek the best throughput adapted to the network conditions. We have also implemented the IWL-MVM-RS algorithm in the NS-3 simulator. Thanks to this implementation, we can evaluate the performance of IWL-MVM-RS in different scenarios (static and with mobility, with and without fast fading). We also compare the performances of IWL-MVM-RS with the ones of MINSTREL-HT and IDEALWIFI, also implemented in the NS-3 simulator [26], [32].

7.3.7. A Passive Method to Infer the Weighted Conflict Graph of a IEEE 802.11 Network

Participants: Lafdal Abdelwedoud, Anthony Busson, Isabelle Guérin-Lassous, Marion Foare.

Wi-Fi networks often consist of several Access Points (APs) to form an Extended Service Set. These APs may interfere with each other as soon as they use the same channel or overlapping channels. A classical model to describe interference is the conflict graph. As the interference level varies in the network and in time, we consider a weighted conflict graph. In this work, we propose a method to infer the weights of the conflict graph of a Wi-Fi network.

Weights represent the proportion of activity from a neighbor detected by the Clear Channel Assessment mechanism. Our method relies on a theoretical model based on Markov networks applied to a decomposition of the original conflict graph. The input of our solution is the activity measured at each AP, measurements available in practice. The proposed method is validated through ns-3 simulations performed for different scenarios. Results show that our solution is able to accurately estimate the weights of the conflict graph. [24], [34].

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. 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, Rémi Gribonval, Marion Foare, Amélie Barbe, Gaetan Frusque.

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.

9.1.3. FIL PerfWiFi

Participants: Guérin-Lassous Isabelle [correspondant], Grünblatt Rémy.

Duration of the project: **January 2019 - December 2020**.

The goal of the project **PerfWiFi** is to set up a Wi-Fi experimental platform that will be, in the future, open to interested researchers. This platform consists in devices (cards, routers) implementing the last versions of Wi-Fi (Wi-Fi 5 and Wi-Fi 6) and with different chipsets from different manufacturers. This platform will also be interconnected to a fleet of UAVs equipped with Wi-Fi interfaces. The Wi-Fi devices are chosen to be as open as possible in order to have a large set of possibilities in terms of parameterization of the Wi-Fi parameters.

In 2019, a first version of the platform has been set up along with a set of software tools to automatically launch Wi-Fi experiments. The first experiments can monitor, during a long period, all the possible Wi-Fi channels and their medium use ratio. We intend to provide these data via an open website.

9.1.4. FIL ALIENOR

Participant: Begin Thomas [correspondant].

Duration of the project: **January 2019 - December 2020**.

The goal of ALIENOR (ArtificiaL IntElligence-assisted NetwORks) is to develop an approach to dynamically select adequate values for the IEEE 802.11 parameters related to the Rate Adaptation (RA) mechanism to the WLAN context. The search for an adequate setting for the RA parameters is made complex due to the vast number of parameters (e.g., the used amendment of 802.11, the channel transmission rate, the number of competing nodes, the Frame Error Rate (FER), the offered load, and the transport protocol to name a few) that may affect a WLAN behavior.

In ALIENOR, we propose to explore a new approach to determine an adequate setting of the RA parameters using a data-driven approach based on techniques of Machine Learning (ML) in Artificial Intelligence (AI). Our approach consists of three stages. First, we will build a large dataset of measurements that will serve as the training set. Second, we will use ML techniques to discover a function that fits the mapping between the dataset output and the inputs. Lastly, WLAN devices will embed and use this learned function to predict (approximately) what will be their attained throughput under various possible settings of their RA, and then select their best option.

9.1.5. ENS Lyon project Vehicular project

Participants: Begin Thomas [correspondant], Guérin Lassous Isabelle, Busson Anthony.

Duration of the project: **January 2017 - December 2020**.

The goal of this project is to design new performance tools to improve the sharing of communication resources in vehicular networks. In particular, we focus on the use case of delivering a Video on Demand service to vehicles traveling along a highway. Through the development of a simple and yet accurate performance modeling approach, we were able to demonstrate the feasibility of using IEEE 802.11p to deliver video content to vehicles. Our work also underlines the benefit of blocking the lowest transmission rates for the sake of a collective gain in terms of attained throughput and interruption time in the video playback. This somehow surprising property derives from the well-established performance anomaly of 802.11-based networks.

9.2. National Initiatives

9.2.1. ANR DataRedux

Participants: Paulo Gonçalves [correspondant], Rémi Gribonval, Marion Foare.

Duration of the project: **February 2020 - January 2024**.
DataRedux puts forward an innovative framework to reduce networked data complexity while preserving its richness, by working at intermediate scales (“mesoscales”). Our objective is to reach a fundamental breakthrough in the theoretical understanding and representation of rich and complex networked datasets for use in predictive data-driven models. Our main novelty is to define network reduction techniques in relation with the dynamical processes occurring on the networks. To this aim, we will develop methods to go from data to information and knowledge at different scales in a human-accessible way by extracting structures from high-resolution, diverse and heterogeneous data. Our methodology will involve the identification of the most relevant subparts of time-resolved datasets while remapping the remaining parts of the system, the simultaneous structural-temporal representations of time-varying networks, the development of parsimonious data representations extracting meaningful structures at mesoscales (“mesostructures”), and the building of models of interactions that include mesostructures of various types. Our aim is to identify data aggregation methods at intermediate scales and new types of data representations in relation with dynamical processes, that carry the richness of information of the original data, while keeping their most relevant patterns for their manageable integration in data-driven numerical models for decision making and actionable insights.

9.2.2. ANR Darling

Participants: Paulo Gonçalves [correspondant], Rémi Gribonval, Marion Foare.

Duration of the project: **February 2020 - January 2024**.
This project meets the compelling demand of developing a unified framework for distributed knowledge extraction and learning from graph data streaming using in-network adaptive processing, and adjoining powerful recent mathematical tools to analyze and improve performances. The project draws on three major parallel directions of research: network diffusion, signal processing on graphs, and random matrix theory which DARLING aims at unifying into a holistic dynamic network processing framework. Signal processing on graphs has recently provided a comprehensive set of basic instruments allowing for signal on graph filtering or sampling, but it is limited to static signal models. Network diffusion on the opposite inherently assumes models of time varying graphs and signals, and has pursued the path of proposing and understanding the performance of distributed dynamic inference on graphs. Both areas are however limited by their assuming either deterministic graph or signal models, thereby entailing often inflexible and difficult-to-grasp theoretical results. Random matrix theory for random graph inference has taken a parallel road in explicitly studying the performance, thereby drawing limitations and providing directions of improvement, of graph-based algorithms (e.g., spectral clustering methods). The ambition of DARLING lies in the development of network diffusion-type algorithms anchored in the graph signal processing lore, rather than heuristics, which shall systematically be analyzed and improved through random matrix analysis on elementary graph models. We believe that this original communion of as yet remote areas has the potential to path the pave to the emergence of the critically needed future field of dynamical network signal processing.

9.2.3. Equipex FIT (Futur Internet of Things)

Participant: Éric Fleury [correspondant].

Duration of the project: **February 2011 - December 2019**.
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. 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.

9.2.4. ANR SoSweet

Participant: Márton Karsai [correspondant].

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.

9.2.5. ANR DylNet

Participant: Márton Karsai [correspondant].

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.

9.2.6. Inria PRE LIAISON

Participant: Márton Karsai [correspondant].

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.

9.2.7. HOTNET - IXXI

Participant: Márton Karsai [correspondant].

Duration of the project: **January 2019 - December 2021**.

The purpose of the HOTNet (Higher-order representation of temporal networks) project is to develop a pipeline for the embedding of temporal networks that captures higher order correlations relevant for dynamical processes. We propose to detach from the straightforward representations of networks — as successions of static networks — by focusing on representations that better reflects the higher-order neighbourhood and temporal paths. To project plans to develop a framework that learns from this representation an embedding sufficient to estimate the outcome of spreading processes that might take place on top of the original network.

This is a small-scale collaborative project funded by the IXXI Complex System Institute to foster collaborations between MK and Laetitia Gauvin (ISI Torino) for the period of 2019-2021.

9.2.8. Inria & HCERES

Participant: Éric Guichard [correspondant].

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,

9.2.9. Inria IPL BetterNet

Participant: Éric Guichard.

An Observatory to Measure and Improve Internet Service Access from User Experience.

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. International Initiatives

9.3.1. Participation in Other International Programs

9.3.1.1. International Initiatives

MOTif

Title: Mobile phone sensing of human dynamics in techno-social environment

International Partners (Institution - Laboratory - Researcher):

Universidad de Buenos Aires (Argentina) - Instituto de Cálculo - Alejo Salles

Universidade Federal de Minas Gerais (Brazil) - Jussara M. Almeida

Duration: 2018 - 2019

Start year: 2018

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.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Jaqueline Faria has been a long term visitor in the DANTE team as a visiting PhD student from the PUC Minas University of Belo Horizonte (Brazil). Her stay between May-December was supported by the CAPES.
- Alexandre Brandwajn from University of California, Santa Cruz (USA) has been a visiting Professor in the DANTE team between Feb and Mar 2019.
- Dorsaf Ghozlani, PhD student at Ecole Nationale d’Ingénieurs de Tunis, has been a visitor in the Dante team from April to July 2019.

9.4.1.1. Internships

- Maxime De Freitas, Télécom Physique Strasbourg, from Jun 2019 until Aug 2019.
- Julien Alamelle, Université Claude Bernard Lyon 1, from Oct 2019 until Dec 2019.
- Juan Pablo Astudillo, Universitat Politècnica de Catalunya, PhD, from Apr 2019 until Jul 2019.
- Simon Fernandez, Master 2 student, ENS Lyon, from February 2019 until June 2019.
- Paul Grangette, Master 2 student, Université Claude from November 2019 to July 2020 (work-study contract).

9.4.2. Visits to International Teams

9.4.2.1. Research Stays Abroad

- Christophe Crespelle is on leave with a Marie Sklodowska-Curie Grant from EU. He is currently at the University of Bergen (Norway) until February 1st, 2020.

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'19 Conference in September 2019 in Singapore
- Márton Karsai was in the organizing team (general chair) of the Machine Learning and Network Science Satellite of the NetSci'19 Conference in June 2019 in Burlington (VT)

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- Thomas Begin was co-Chair for the National Conference Algotel 2019.

10.1.2.2. Member of the Conference Program Committees

- Thomas Begin has been member of the PC for the conference IEEE LCN 2019
- Isabelle Guérin Lassous has been a member of the PC for the following conferences in 2019: ACM MSWiM, Mascots, IEEE ICC, IEEE Globecom, IEEE WCNC.
- Márton Karsai has been the member of the PC of the conferences: NetSci, CompleNet, IC2S2, Algotel, NetSciX, CSSCCS, MLNS, Complex Networks, CCS, IC2S2
- Anthony Busson has been member of the PC for Algotel 2019, FNC 2019.
- Christophe Crespelle

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 is member of the editorial boards of PLoS ONE
- 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).
- Thomas Begin was a reviewer for IEEE Transactions on Network and Service Management, IEEE Transactions on Parallel and Distributed Systems, and Performance Evaluation.
- Anthony Busson was a reviewer for Computer communication, Sensors, IEEE wireless communications letters.

10.1.4. Invited Talks

- Thomas Begin gave the following invited talk:

  Contributions to the Performance Modeling of Computer Networks - CITI lab (Avril 2019, Lyon, France)
- Isabelle Guérin-Lassous gave the following invited talks:

UAVs and cellular networks: which usage, which constraints, which performance? - Entretiens Jacques Cartier (November 2019, Montreal, Canada)

How to improve the performance of a Wi-Fi network without changing the Wi-Fi technologie? - Séminaire ENS Rennes (September 2019, Rennes, France)

- Márton Karsai gave the following invited talks:

  Lecture in the Business Analytics and Data Mining MSc program - Bocconi University (7 March 2019, Milano, Italy)

  Complex System Academy - Université de Côte d'Azure (25 March 2019, Nice, France)

  YEP XV - Information Diffusion on Random Networks - TU/e, EUROSTAT (27 April 2019, Eindhoven, The Netherlands)

  Lectures about Artificial Intelligence - EM Lyon Business School - AIM Institute (4 April 2019, Lyon, France)

  MOTIF Conference on Human Dynamics - Universidade Federal de Minas Gerais (10 April 2019, Belo Horizonte, Brazil)

  UNICEF Innovation seminar - UNICEF Innovation Office - United Nations (4 June 2019, New York, NY, USA)

  MédiaLab seminar - SciencePo Paris (25 June 2019, Paris, France)

  AGRANDA Simposio Argentino de Ciencia de Datos Grandes Datos - 48 JAIIO (17 Septemner 2019, Salta, Argentina)

  Kickoff meeting of the DYNASNET Synergy ERC project - Central European University (23 September 2019, Budapest, Hungary)

  Dondena Seminar Series- Bocconi University (14 October 2019, Milano, Italy)

  Agglomeration and Social Networks Seminar - MTA Center for Economic and Regional Studies (8 November 2019, Budapest, Hungary)

  BigData and Mobility Workshop - University of Havana (25-26 November, Havana, Cuba)

10.1.5. Scientific Expertise

- Isabelle Guérin Lassous

  has been a member of the strategic consulting of the Fondation Blaise Pascal since September 2019.

  has been a member of the scientific council of the Société Informatique de France since November 2019

- 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).

  is a member of the international evaluation board of the doctoral program *Filosofia da Ciencia, Tecnologia, Arte e Sociedade* of the University of Lisbon.

  is the manager of the RAIL (Réseau de l'Atelier Internet Lyonnais), founded in 2017 and supported by IXXI and Enssib.

10.1.6. Research Administration

- Paulo Gonçalves

  is scientific liaison officer for international relations in Inria Research Centre of Rhône-Alpes.

  is a member of the executive committee of the Milyon labex and referent for its valorisation committee.

is correspondant for the theme "Big Data" of the *Fédération d'Informatique de Lyon*.

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.
- Anthony Busson

  is member of the Thesis Commission at LIP.

  is head of the computer science department at IUT (Institut Universitaire de Technologie) - Université Lyon Claude Bernard Lyon 1.
- Thomas Begin

  is an elected member of the Council of the LIP laboratory.

  is an elected member of the department council of the Computer Science department of Université Lyon 1.
- Márton Karsai

  is the co-responsible for the M2 master program in Modelling of Complex Systems at ENS Lyon

  is the elected council member of the Complex System Society (2015-)

  is the elected member of executive committee of the Complex System Society (2018-)

  is the elected member of the steering committee of the IXXI Complex System Institute (2017-)

  is the member of the computational infrastructure board of LIP

  is the member of the communication board of LIP
- Éric Guichard is a member of the steering committee of the IXXI Complex System Institute

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

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.

Responsabilités

Thomas Begin : Head of the Master 2 speciality SRIV (Systèmes, Réseaux et Infrastructures Virtuelles) at UCBL - http://master-info.univ-lyon1.fr/SRIV

Isabelle Guérin Lassous has been appointed as the president of the CAPES NSI (Numérique et Science Informatique) committee since November 2019. This committee will be in charge of the recruitment of the computer science teachers in high schools.

Anthony Busson is head of the computer science department at IUT Lyon 1 Doua.

10.2.2. Supervision

PhD defense: Esteban BAUTISTA RUIZ, "Laplacian Powers for Graph-Based Semi-Supervised Learning", November 27, 2019. P. Gonçalves (Dir.).

PhD defense: Marija STOJANOVA, "Performance Modelling of IEEE 802.11 networks" ,December 16, 2019. T. Begin (Dir.)

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, M. Karsai director

PhD in progress: Jacobo Levy ABITBOL, Information diffusion and language evolution on dynamical social networks, Oct 2016, M. Karsai director and E. Fleury

PhD in progress: Sicheng DAI, Dynamic Multilayer Network Modelling, M. Karsai director. Started October 1st, 2017.

PhD in progress: Gaetan FRUSQUE, *Modal Decompositions of Dynamic Graphs : Application in Neurosciences*, P. Gonçalves (Dir. and 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 (Dir. with M. Sebban and P. Borgnat, co-advisors). 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 reviewer of the Ph.D thesis examination board of Xiaoyi MAI, Centrale Supelec, October 2019.

Paulo Gonçalves was member of the Ph.D thesis examination boards of Karina Ashurbekova (Université Grenoble Alpes, December 2019).

Isabelle Guérin Lassous was a reviewer of the Ph.D thesis examination boards of Fannia Pacheco (UPPA, November 2019) and Pierre Brunisholz (Université Grenoble Alpes, May 2019).

Isabelle Guérin Lassous was a reviewer of the HDR examination board of Géraldine Texier (IMT-Atlantique, December 2019).

Isabelle Guérin Lassous was the president of the Ph.D thesis examination boards of Marc Heinrich (Université de Lyon, July 2019) and Thierry Arrabal (Université Franche-Comté, November 2019).

Isabelle Guérin Lassous was the president of the HDR examination boards of Christophe Alias (ENS de Lyon, May 2019), Marton Karsai (ENS de Lyon, April 2019) and Razvan Stanica (INSA de Lyon, November 2019).

Isabelle Guérin Lassous was a member of the Ph.D thesis examination board of Jalal Rachad (Telecom Paris, December 2019).

Isabelle Guérin Lassous was a member of the HDR examination board of Katia Jaffres-Runser (INP Toulouse, July 2019).

Thomas Begin was a reviewer of the Ph.D. thesis examination board of Amira KAMLI, Télécom-ParisSud, October 2019.

Márton Karsai was the member of the Prix de Thèse Systm`es Complexes - Complex System Institute, July 2019

Márton Karsai was the member of the PhD jury of Jordam Cambe (ENS Lyon), October 2019

Anthony Busson was a reviewer of the Ph.D thesis examination board of Jalal Rachad (Telecom Paris, December 2019).

Anthony Busson was a reviewer of the Ph.D thesis examination board of Henri-Joseph Audeoud (Université Grenoble Alpes, December 2019).

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

Isabelle Guérin Lassous has been the managing director of the Foundation Blaise Pascal until the end of June 2019. 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. Interventions

Isabelle Guérin Lassous participated to the project Science XX Elles organized by the association Femmes & Sciences, CNRS and Ecole Normale Supérieure de Lyon. A portrait has been realized and has been exhibited in from of the Musée des Confluences during La Fête de la Science in October 2019 (http://www.femmesetsciences.fr/lascience-taillexxelles/lyon/isabelle-guerin-lassous/. She also participated to a speed meeting with families visiting the Musée des Confluences during which she has promoting her research activity.

Isabelle Guérin Lassous spent one day in high school (lycée St Just in Lyon) to promote computer science and to present female computer scientists in from of terminal classes.

Isabelle Guérin Lassous participated to the Challenge Innovatech organized by the association Elles Bougent. She was a member of the selection committe who selected a scientific project designed by a group of young girls (in high school).

11. Bibliography

Major publications by the team in recent years

[1] T. BEGIN, B. BAYNAT, I. GUÉRIN-LASSOUS, T. ABREU.*Performance analysis of multi-hop flows in IEEE 802.11 networks: A flexible and accurate modeling framework*, in "Performance Evaluation", February 2016, vol. 96, p. 12–32 [*DOI :* 10.1016/J.PEVA.2015.12.003], https://hal.archives-ouvertes.fr/hal-01246822

[2] A. BRANDWAJN, T. BEGIN.*Multi-server preemptive priority queue with general arrivals and service times*, in "Performance Evaluation", 2017 [*DOI :* 10.1016/J.PEVA.2017.08.003], https://hal.inria.fr/hal-01581118

[3] A. T. GIANG, A. BUSSON, A. LAMBERT, D. GRUYER.*Spatial capacity of IEEE 802.11p based VANET: models, simulations and experimentations*, in "IEEE Transactions on Vehicular Technology", January 2015 [*DOI :* 10.1109/TVT.2015.2474156], https://hal.archives-ouvertes.fr/hal-01217564

[4] B. GIRAULT, P. GONÇALVES, É. FLEURY.*Translation on Graphs: An Isometric Shift Operator*, in "IEEE Signal Processing Letters", December 2015, vol. 22, n^{o} 12, p. 2416–2420 [*DOI :* 10.1109/LSP.2015.2488279], https://hal.inria.fr/hal-01221562

[5] M. KARSAI, G. IÑIGUEZ, K. KASKI, J. KERTÉSZ.*Complex contagion process in spreading of online innovation*, in "Journal of the Royal Society Interface", October 2014, vol. 11, 8, 20140694 [*DOI :* 10.1098/RSIF.2014.0694], https://hal.inria.fr/hal-01100359

[6] T. OBADIA, R. SILHOL, L. OPATOWSKI, L. TEMIME, J. LEGRAND, A. C. THIÉBAUT, J.-L. HERRMANN, É. FLEURY, D. GUILLEMOT, P.-Y. BOELLE.*Detailed Contact Data and the Dissemination of Staphylococcus aureus in Hospitals*, in "PLoS Computational Biology", 2015, vol. 11, n^{o} 3, on behalf of the I-Bird study Group [*DOI :* 10.1371/JOURNAL.PCBI.1004170], http://hal.upmc.fr/hal-01134050

[7] M. TIZZONI, K. SUN, D. BENUSIGLIO, M. KARSAI, N. PERRA.*The Scaling of Human Contacts in Reaction-Diffusion Processes on Heterogeneous Metapopulation Networks*, in "Scientific Reports", October 2015, vol. 5, n^{o} 15111, https://hal.inria.fr/hal-01100351

Publications of the year

Doctoral Dissertations and Habilitation Theses

[8] E. BAUTISTA.*Laplacian Powers for Graph-Based Semi-Supervised Learning*, Ecole Doctorale en Informatique et Mathématiques de Lyon, November 2019, https://hal.inria.fr/tel-02437253

[9] M. KARSAI.*Computational Human Dynamics: People, Networks, and Collective Phenomena*, Ecole normale supérieure de Lyon ; Laboratoire de l'Informatique du Parallélisme, April 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02151333

Articles in International Peer-Reviewed Journal

[10] A. BALTZER, M. KARSAI, C. ROTH.*Interactional and Informational Attention on Twitter*, in "Information", August 2019, vol. 10, n^{o} 8, p. 1-16 [*DOI :* 10.3390/INFO10080250], https://halshs.archives-ouvertes.fr/halshs-02378190

[11] E. BAUTISTA, P. ABRY, P. GONÇALVES.*L^γ -PageRank for Semi-Supervised Learning*, in "Applied Network Science", 2019, p. 1-20, forthcoming, https://hal.inria.fr/hal-02063780

[12] T. BEGIN, A. BUSSON, I. GUÉRIN-LASSOUS, A. BOUKERCHE.*Performance Analysis of Video on Demand in an IEEE 802.11p-based Vehicular Network*, in "Computer Communications", August 2019, vol. 146, p. 174-185 [*DOI :* 10.1016/J.COMCOM.2019.08.006], https://hal.archives-ouvertes.fr/hal-02277855

[13] 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, p. 1-14 [*DOI :* 10.1109/TNET.2018.2890512], https://hal.inria.fr/hal-01672961

[14] A. DUVAL, T. OBADIA, P.-Y. BOËLLE, É. FLEURY, J.-L. HERRMANN, D. GUILLEMOT, L. TEMIME, L. OPATOWSKI.*Close proximity interactions support transmission of ESBL-K. pneumoniae but not ESBL-E. coli in healthcare settings*, in "PLoS Computational Biology", May 2019, vol. 15, n^{o} 5, e1006496 [*DOI :* 10.1371/JOURNAL.PCBI.1006496], https://www.hal.inserm.fr/inserm-02153937

[15] M. FOARE, N. PUSTELNIK, L. CONDAT.*Semi-Linearized Proximal Alternating Minimization for a Discrete Mumford-Shah Model*, in "IEEE Transactions on Image Processing", 2019, p. 1-13, forthcoming [*DOI :* 10.1109/TIP.2019.2944561], https://hal.archives-ouvertes.fr/hal-01782346

[16] É. GUICHARD.*Cartographie et visualisation*, in "Annales des mines - Responsabilité et environnement", 2019, n^{o} 94, p. 38-41, https://halshs.archives-ouvertes.fr/halshs-02097578

[17] É. GUICHARD.*Les nouveaux maîtres de l'écriture du monde*, in "Contemporary French and Francophone Studies", 2019, vol. 23, n^{o} 4, p. 1-9, forthcoming, https://hal.archives-ouvertes.fr/hal-02310616

[18] F. HUA, C. RICHARD, C. JIE, H. WANG, P. BORGNAT, P. GONÇALVES.*Designing Convex Combination of Graph Filters*, in "IEEE Signal Processing Letters", 2019, forthcoming, https://hal.inria.fr/hal-02367868

[19] J. LEVY ABITBOL, É. FLEURY, M. KARSAI.*Optimal Proxy Selection for Socioeconomic Status Inference on Twitter*, in "Complexity", May 2019, vol. 2019, p. 1-15 [*DOI :* 10.1155/2019/6059673], https://hal.inria.fr/hal-02388403

[20] Y. LÉO, C. CRESPELLE, É. FLEURY.*Non-altering time scales for aggregation of dynamic networks into series of graphs*, in "Computer Networks", 2019, vol. 148, p. 108-119 [*DOI :* 10.1016/J.COMNET.2018.11.006], https://hal.inria.fr/hal-01969504

[21] F. PINOTTI, É. FLEURY, D. GUILLEMOT, P.-Y. BÖELLE, C. POLETTO.*Host contact dynamics shapes richness and dominance of pathogen strains*, in "PLoS Computational Biology", May 2019, vol. 15, n^{o} 5, e1006530 [*DOI :* 10.1371/JOURNAL.PCBI.1006530], https://www.hal.inserm.fr/inserm-02153754

[22] B. RICAUD, P. BORGNAT, N. TREMBLAY, P. GONÇALVES, P. VANDERGHEYNST.*Fourier could be a Data Scientist: from Graph Fourier Transform to Signal Processing on Graphs*, in "Comptes Rendus Physique", September 2019, p. 474-488 [*DOI :* 10.1016/J.CRHY.2019.08.003], https://hal.inria.fr/hal-02304584

[23] S. UNICOMB, G. IÑIGUEZ, J. KERTÉSZ, M. KARSAI.*Reentrant phase transitions in threshold driven contagion on multiplex networks*, in "Physical Review E ", October 2019, vol. 100, n^{o} 4, p. 1192-1199 [*DOI :* 10.1103/PHYSREVE.100.040301], https://hal.inria.fr/hal-02388406

International Conferences with Proceedings

[24] L. ABDELWEDOUD, A. BUSSON, I. GUÉRIN-LASSOUS, M. FOARE.*A Passive Method to Infer the Weighted Conflict Graph of an IEEE 802.11 Network*, in "AdHoc-Now 2019 - 18th International Conference on Ad Hoc Networks and Wireless", Luxembourg, Luxembourg, October 2019, p. 304-316 [*DOI :* 10.1007/978-3-030-31831-4_21], https://hal.inria.fr/hal-02404943

[25] *Best Paper*
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 "PAM 2019 - 20th Passive and Active Measurements Conference", Puerto Varas, Chile, Springer, March 2019, p. 197-209, Best paper award. [*DOI :* 10.1007/978-3-030-15986-3_13], https://hal.inria.fr/hal-01970596.

[26] R. GRÜNBLATT, I. GUÉRIN-LASSOUS, O. SIMONIN.*Simulation and Performance Evaluation of the Intel Rate Adaptation Algorithm*, in "MSWiM 2019 - 22nd ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems", Miami Beach, United States, ACM, November 2019, p. 27-34 [*DOI :* 10.1145/3345768.3355921], https://hal.inria.fr/hal-02282508

[27] S. GUPTA, R. GRIBONVAL, L. DAUDET, I. DOKMANIĆ.*Don't take it lightly: Phasing optical random projections with unknown operators*, in "NeurIPS 2019 - Thirty-third Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, p. 1-13, https://arxiv.org/abs/1907.01703 , https://hal.inria.fr/hal-02342280

[28] N. K. PANIGRAHY, P. BASU, P. NAIN, D. TOWSLEY, A. SWAMI, K. S. CHAN, K. K. LEUNG.*Resource Allocation in One-dimensional Distributed Service Networks*, in "MASCOTS 2019 - 27th IEEE International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems", Rennes, France, October 2019, p. 14-26 [*DOI :* 10.1109/MASCOTS.2019.00013], https://hal.inria.fr/hal-02267631

[29] G. VARDOYAN, S. GUHA, P. NAIN, D. TOWSLEY.*On the Capacity Region of Bipartite and Tripartite Entanglement Switching and Key Distribution*, in "QCRYPT 2019 - 9th International Conference on Quantum Cryptography", Montreal, Canada, August 2019, p. 1-3, https://hal.inria.fr/hal-02424441

National Conferences with Proceeding

[30] D. BARBE, P. BORGNAT, P. GONÇALVES, M. SEBBAN.*Transport Optimal sous Contrainte de Régularité pour l'Adaptation de Domaines entre Graphes avec Attributs*, in "GRETSI 2019 - XXVIIème Colloque francophonede traitement du signal et des images", Lille, France, August 2019, p. 1-4, https://hal.inria.fr/hal-02154883

[31] G. FRUSQUE, J. JUNG, P. BORGNAT, P. GONÇALVES.*Réduction de dimension tensorielle parcimonieuse: Application au clustering de connectivité fonctionnelle*, in "GRETSI 2019 - XXVIIème Colloque francophonede traitement du signal et des images", Lille, France, August 2019, p. 1-4, https://hal.inria.fr/hal-02154888

[32] R. GRÜNBLATT, I. GUÉRIN-LASSOUS, O. SIMONIN.*Study of the Intel WiFi Rate Adaptation Algorithm*, in "CoRes 2019 - Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication", Saint-Laurent-de-la-Cabrerisse, France, June 2019, p. 1-4, https://hal.inria.fr/hal-02126333

[33] M. STOJANOVA, T. BEGIN, P. GONÇALVES.*Traitement du signal sur graphe pour modéliser les WLANs*, in "GRETSI 2019 - XXVIIème Colloque francophonede traitement du signal et des images", Lille, France, August 2019, p. 1-4, https://hal.inria.fr/hal-02154891

Conferences without Proceedings

[34] L. ABDELWEDOUD, A. BUSSON, I. GUÉRIN-LASSOUS.*Inférence des poids d'un graphe des conflits modélisant un réseau Wi-Fi*, in "CoRes 2019 - Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication", Saint-Laurent-de-la-Cabrerisse, France, June 2019, p. 1-4, https://hal.archives-ouvertes.fr/hal-02122772

[35] J. P. ASTUDILLO LEO´N, T. BEGIN, A. BUSSON, L. J. DE LA CRUZ LLOPIS.*Towards a Distributed Congestion Control mechanism for Smart Grid Neighborhood Area Networks*, in "PE-WASUN 2019 - Sixteenth ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks", Miami, United States, ACM, November 2019, https://hal.archives-ouvertes.fr/hal-02309054

[36] 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 "CILPR 2019 - XXIXe Congrès international de linguistique et de philologie romanes", Copenhague, Denmark, July 2019, https://hal.archives-ouvertes.fr/hal-01970511

[37] G. FRUSQUE, J. JUNG, P. BORGNAT, P. GONÇALVES.*Sparse tensor dimensionality reduction with application to clustering of functional connectivity*, in "Wavelets and Sparsity XVIII", San Diego, United States, SPIE, December 2019, 22 [*DOI :* 10.1117/12.2529595], https://hal.archives-ouvertes.fr/hal-02399385

[38] P. STOCK, A. JOULIN, R. GRIBONVAL, B. GRAHAM, H. JÉGOU.*And the Bit Goes Down: Revisiting the Quantization of Neural Networks*, in "ICLR 2020 - Eighth International Conference on Learning Representations", Addis-Abeba, Ethiopia, February 2020, p. 1-11, https://hal.archives-ouvertes.fr/hal-02434572

[39] M. STOJANOVA, T. BEGIN, A. BUSSON.*Modélisation des réseaux IEEE 802.11 : Diviser pour régner*, in "CoRes 2019 - Rencontres Francophones sur la Conception de Protocoles, l'Évaluation de Performance et l'Expérimentation des Réseaux de Communication", Saint Laurent de la Cabrerisse, France, June 2019, p. 1-4, https://hal.archives-ouvertes.fr/hal-02121254

Scientific Books (or Scientific Book chapters)

[40] 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, forthcoming, https://hal.archives-ouvertes.fr/hal-01673099

[41] É. GUICHARD.*Digital Humanities do not exist*, in "Amériques / Europe, les Humanités numériques en partage? Enjeux, innovations et perspectives", C. ILLOUZ, A. HUERTA (editors), Presses de l'université de la Rochelle, 2020, forthcoming, https://hal.archives-ouvertes.fr/hal-02403315

[42] É. GUICHARD.*Parler du virtuel aux temps du numérique*, in "La technologie une et multiple", J.- C. BEAUNE, M. SALHAB (editors), La technologie une et multiple, Université Libano-Française (ULF) et Éditions universitaires du Liban, 2020, forthcoming, https://hal.archives-ouvertes.fr/hal-02403360

[43] 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", 2019, https://hal.archives-ouvertes.fr/hal-01672289

[44] 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", M. NAGATSU, A. RUZZENE (editors), Bloomsbury, May 2019, https://hal.archives-ouvertes.fr/hal-01672283

Other Publications

[45] 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

[46] S. FOUCART, R. GRIBONVAL, L. JACQUES, H. RAUHUT.*Jointly Low-Rank and Bisparse Recovery: Questions and Partial Answers*, October 2019, https://arxiv.org/abs/1902.04731 - working paper or preprint, https://hal.inria.fr/hal-02062891

[47] G. FRUSQUE, J. JUNG, P. BORGNAT, P. GONÇALVES.*Sparse tensor dimensionality reduction with application to clustering of functional connectivity*, June 2019, working paper or preprint, https://hal.inria.fr/hal-02154903

[48] L. GAUVIN, M. GÉNOIS, M. KARSAI, M. KIVELÄ, T. TAKAGUCHI, E. VALDANO, C. L. VESTERGAARD.*Randomized reference models for temporal networks*, March 2019, https://arxiv.org/abs/1806.04032 - Simplified presentation of temporal network features, added figures to illustrate important concepts, and added a new example application of MRRMs, https://hal.archives-ouvertes.fr/hal-01817633

[49] R. GRIBONVAL, M. NIKOLOVA.*A characterization of proximity operators*, November 2019, https://arxiv.org/abs/1807.04014 - working paper or preprint, https://hal.inria.fr/hal-01835101

[50] Y. KALOGA, M. FOARE, N. PUSTELNIK, P. JENSEN.*Discrete Mumford-Shah on graph for mixing matrix estimation*, May 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01948570

[51] S. LERIQUE, J. L. ABITBOL, M. KARSAI.*Joint embedding of structure and features via graph convolutional networks*, December 2019, https://arxiv.org/abs/1905.08636 - working paper or preprint, https://hal.inria.fr/hal-02388402

[52] A. B. MODIRI, M. KARSAI, M. KIVELÄ.*Efficient limited time reachability estimation in temporal networks*, December 2019, https://arxiv.org/abs/1908.11831 - Software implementation available at https://doi.org/10.5281/zenodo.3369893, https://hal.inria.fr/hal-02388408

[53] M. TORRICELLI, M. KARSAI, L. GAUVIN.*weg2vec: Event embedding for temporal networks*, December 2019, https://arxiv.org/abs/1911.02425 - working paper or preprint, https://hal.inria.fr/hal-02388409

[54] S. UNICOMB, G. IÑIGUEZ, J. KERTÉSZ, M. KARSAI.*Reentrant phase transitions in threshold driven contagion on multiplex networks*, March 2019, https://arxiv.org/abs/1901.08306 - 13 pages, 7 figures, https://hal.inria.fr/hal-02065731

References in notes

[55] R. ALBERT, A.-L. BARABÁSI.*Statistical mechanics of complex networks*, in "Reviews of Modern Physics", 2002, vol. 74

[56] A.-L. BARABÁSI.*The origin of bursts and heavy tails in human dynamics*, in "Nature", 2005, vol. 435, 207

[57] 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^o 21, p. 7426-7431

[58] W. J. FITZGERALD, R. L. SMITH, A. T. WALDEN.*Nonlinear and Nonstationary Signal Processing*, Cambridge University Press, Cambridge, 2001

[59] S. FORTUNATO.*Community detection in graphs*, in "Physics Reports", 2010, vol. 486, p. 75-174

[60] 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

[61] 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

[62] P. HOLME, J. SARAMÄKI.*Temporal networks*, in "Physics Reports", 2012, vol. 519, p. 97-125

[63] 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

[64] M. KARSAI, N. PERRA, A. VESPIGNANI.*A. Random Walks and Search in Time-Varying Networks*, 2013, arXiv:1303.5966

[65] 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

[66] L. KOVANEN, M. KARSAI, K. KASKI, J. KERTÉSZ, J. SARAMÄKI.*Temporal motifs in time-dependent networks*, in "J. Stat. Mech.", 2011

[67] 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^o 4

[68] Z. Q. LUO, M. GASTPAR, J. LIU, A. SWAMI.*Distributed Signal Processing in Sensor Networks*, in "IEEE Signal Processing Mag", 2006, vol. 23

[69] B. A. MILLER, N. T. BLISS, P. J. WOLFE.*Towards Signal Processing Theory for Graphs and Non-Euclidian Data*, in "ICASSP", Dallas, IEEE, 2010

[70] G. MIRITELLO, E. MORO, R. LARA.*Dynamical strength of social ties in information spreading*, in "Phys. Rev. E", 2011, vol. 83

[71] M. E. J. NEWMAN.*Networks: An Introduction,*, Oxford University Press, 2010

[72] 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

[73] N. PERRA, B. GONÇALVES, R. PASTOR-SATORRAS, A. VESPIGNANI.*Activity driven modeling of time varying networks*, in "Scientific Reports", 2012, vol. 2, n^{o} 469

[74] 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^{o} 3, p. 83–98

[75] 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

[76] 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

[77] A. S. WILLSKY.*Multiresolution Statistical Models for Signal and Image Processing*, in "Proceedings of the IEEE", 2002, vol. 90

[78] K. ZHAO, M. KARSAI, G. BIANCONI.*Entropy of Dynamical Social Networks*, in "PLoS ONE", 2011, vol. 6, n^{o} 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

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

Bruno Raffin [Team leader, Inria, Senior Researcher, HDR]
Fanny Dufosse [Inria, Researcher]
Francieli Zanon-Boito [Inria, Researcher, until Aug 2019]

Faculty Members

Yves Denneulin [Institut polytechnique de Grenoble, Professor]
Pierre-François 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]

Technical Staff

Michael Mercier [Inria, Engineer, until Feb 2019]
Pierre Neyron [CNRS, Engineer]
Samir Noir [Inria, Engineer, from Nov 2019]
Albin Petit [Inria, Engineer, from Apr 2019]
Theophile Terraz [Inria, Engineer]

PhD Students

Danilo Carastan Dos Santos [Universidade Federal do ABC, PhD Student]
Vincent Fagnon [Univ Grenoble Alpes, PhD Student, from Oct 2019]
Adrien Faure [Bull ATOS, PhD Student, granted by CIFRE]
Sebastian Friedemann [Inria, PhD Student, from Mar 2019]
Mohammed Khatiri [Univ Grenoble Alpes, PhD Student]
Angan Mitra [Qarnot Computing, PhD Student, from Mar 2019, granted by CIFRE]
Clement Mommessin [Institut polytechnique de Grenoble, PhD Student]
Ioannis Panagiotas [Inria, PhD Student]
Paul Youssef [Univ Grenoble Alpes, PhD Student, from Nov 2019]
Salah Zrigui [Univ Grenoble Alpes, PhD Student]

Post-Doctoral Fellows

Essam Morsi [Inria, Post-Doctoral Fellow, from Nov 2019]
Millian Poquet [Inria, Post-Doctoral Fellow, from Sep 2019]
Malin Rau [Institut polytechnique de Grenoble, Post-Doctoral Fellow, from Jul 2019]

Visiting Scientist

Alfredo Goldman [Universidade de São Paulo, from Jun 2019 until Jul 2019]

Administrative Assistant

Annie Simon [Inria, Administrative Assistant]

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

[0] Top500 Ranking, http://www.top500.org
[0] 10^{15} floating point operations per second

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 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.

[0] SciDAC Review, 2010, http://scidacreview.org/1001/pdf/hardware.pdf

[0] https://www.whitehouse.gov/the-press-office/2015/07/29/executive-order-creating-national-strategic-computing-initiative

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.

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

[0] Standard Workload Format: http://www.cs.huji.ac.il/labs/parallel/workload/swf.html

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.

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. 6.4) 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.

5. Highlights of the Year

5.1. Highlights of the Year

- Pierre Neyron received the Médaille de Cristal CNRS 2019 (http://www.cnrs.fr/fr/personne/pierre-neyron)
- Denis Trystram leading the Edge Intelligence chair of the new Institute of Artificial Intelligence of Univ. Grenoble Alpes (MIA@Grenoble-Alpes).
- Best Paper Awards at CCGrid 2019
- Outstanding Paper Award at HPCS 2019

BEST PAPERS AWARDS :

[12]
D. CARASTAN-SANTOS, R. Y. DE CAMARGO, D. TRYSTRAM, S. ZRIGUI.*One can only gain by replacing EASY Backfilling: A simple scheduling policies case study*, in "CCGrid 2019 - International Symposium in Cluster, Cloud, and Grid Computing", Larnaca, Cyprus, IEEE, May 2019, p. 1-10 [*DOI :* 10.1109/CCGRID.2019.00010], https://hal.archives-ouvertes.fr/hal-02237895

[15]
F. ZANON BOITO, R. NOU, L. LIMA PILLA, J. LUCA BEZ, J.-F. MÉHAUT, T. CORTES, P. O. NAVAUX.*On server-side file access pattern matching*, in "HPCS 2019 - 17th International Conference on High Performance Computing & Simulation", Dublin, Ireland, IEEE, 2019, p. 1-8, outstanding paper award, https://hal.inria.fr/hal-02079899

6. New Software and Platforms

6.1. 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

6.2. 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

6.3. 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: Theophile 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

6.4. Platforms

6.4.1. Grid'5000 (https://www.grid5000.fr/) and Meso Center Ciment (http://ciment.univ-grenoble-alpes.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).

7. New Results

7.1. Integration of High Performance Computing and Data Analytics

7.1.1. In Situ Processing Model

The work in [2] focuses on proposing a model for in situ analysis taking into account memory constraints. This model is used to provide different scheduling policies to determine both the number of resources that should be dedicated to analysis functions, and that schedule efficiently these functions. We evaluate them and show the importance of considering memory constraints when choosing in between in situ and in transit resource allocation.

7.1.2. I/O Characterization

I/O operations are the bottleneck of several HPC applications due to the difference between process- ing and data access speeds. Hence, it is important to understand and characterize the typical I/O behavior of these applications, so we can identify problems in HPC architectures and propose solutions. In [3], we conducted an extensive analysis to collect and analyze information about applications that run in the Santos Dumont supercomputer, deployed in the National Laboratory for Scientific Computing (LNCC), in Brazil. In [9], we propose an I/O characterization approach that uses unsupervised learning to cluster jobs with similar I/O behavior, using information from high-level aggregated traces.

7.1.3. Online adaptation of the I/O stack to applications

I/O optimization techniques such as request scheduling can improve performance mainly for the access patterns they target, or they depend on the precise tune of parameters. In [19], we propose an approach to adapt the I/O forwarding layer of HPC systems to the application access patterns by tuning a request scheduler. Our case study is the TWINS scheduling algorithm, where performance improvements depend on the time window parameter, which depends on the current workload. Our approach uses a reinforcement learning technique to make the system capable of learning the best parameter value to each access pattern during its execution, without a previous training phase. Our approach can achieve a precision of 88% on the parameter selection in the first hundreds of observations of an access pattern. After having observed an access pattern for a few minutes (not necessarily contiguously), the system will be able to optimize its performance for the rest of the life of the system (years).

Such an auto-tuning approach requires a classification of application access patterns,to separate situations where the optimization techniques will have a different performance behavior. Such a classification is not available in the stateless server-side, hence it has to be estimated from metrics on recent accesses. In [8], we evaluate three machine learning techniques to automatically detect the I/O access pattern of HPC applications at run time: decision trees, random forests, and neural networks. We also proposed in [15] a pattern matching approach for server-side access pattern detection for the HPC I/O stack. The goal is to empower the system to learn a classification during the execution of the system, by representing access patterns by all relevant metrics. We build a time series to represent accesses spatiality, and use a pattern matching algorithm, in addition to an heuristic, to compare it to known patterns.

7.1.4. Data management for workflow execution

In [11], we studied a typical scenario in research facilities. Instrumental data is generated by lab equipment such as microscopes, collected by researchers into USB devices, and analyzed in their own computers. In this scenario, an instrumental data management framework could store data in a institution-level storage infrastructure and allow to execute tasks to analyze this data in some available processing nodes. This setup has the advantages of promoting reproducible research and the efficient usage of the expensive lab equipment (in addition to increasing researchers productivity). We detailed the requirements for such a framework regarding the needs of our case study of the CEA, analyzed performance limitations of the proposed architecture, and pointed to the connection between centralized storage and the processing nodes as the critical point.

In order to alleviate this bottleneck, we investigated using the storage devices of the processing nodes as a cache for the remote storage, and replication strategies to maximize data locality for tasks. A simulator called RepliSim was developed for this research.

7.2. Data Aware Batch Scheduling

We obtained in 2018 two important results on on-line scheduling using resource augmentation. The main idea is that the algorithm is applied to a more powerful environment than that of the adversary. We focused more specifically on the mechanism of rejection based on the concept of duality for mathematical programming applied for the analysis of the algorithm's performance. More precisely, we proposed a primal-dual algorithm for the online scheduling problem of minimizing the total weighted flow time of jobs on unrelated machines when the preemption of jobs is not allowed. This analysis concerned usual sequential jobs. These results have been distinguished among the most significant ones on the annual ACM review of on-line algorithms. We extended this work on a practical side by applying the analysis to actual batch schedulers with parallel jobs, rejection was interpreted as redirecting jobs to some predefined machines.

Machine Learning is a hot topic which received recently a great attention for dealing with the huge amount of data produced by the explosion of the digital applications and for dealing with uncertainties. The members of DataMove promoted a methodology based on simulation and machine learning to obtain efficient dynamic scheduling policies. The main idea is to focus the learning scheme targeting the policies them-selves, and not the specific parameters of the problem. Today, this methodology is mature and it is applied in several project like ANR Energumen (performances and replaced by energy saving). We also launched a new project at MIAI on edge Intelligence. The idea is to propose an alternative to the high-consuming classical IA by doing most of the computations close the the place where the data are produced. We are developing both an efficient task orchestration framework and distributed learning algorithms.

We wrote a survey [20] on scheduling on heterogeneous machines where we provided a complete benchmark suite and we recoded all existing algorithms and compared them.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

- **EDF R&D (2019)**. Integration of Melissa and OpenTurn.
- **TOTAL SA (2019)**. Proof of Concept for performing large scale sensibility analysis with Melissa on Total use-case.

8.2. Bilateral Grants with Industry

- **ATOS-BULL (2016-2019)**. Two PhD grants (Michael Mercier and Adrien Faure). Job and resource management algorithms.
- **Qarnot Computing (2019-2022)**. PhD grant (Angan Mitra).

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. 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.

9.1.2. Competitivity Clusters

- **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).

9.1.3. 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/

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

- **H2020 EoCoE-II (2019-2021)**
  - Energy oriented Center of Excellence on HPC.

  - H2020 RIA european project, call H2020-INFRAEDI-2018-1.
  - PI: CEA.
  - Partners: CEA, FZL, ENEA, BSC, CNRS, Inria, CERFACS, Max-Planck-Gesellschaft, FRAUNHOFER, FAU, CNR, UNITN, PSNC, ULB, UBAH, CIEMAT, IFPEN, DDN. Datamove is leading the WP5 (Ensemble Runs)
  - Summary: The EoCoE-II project will build on its unique, established role at the crossroads of HPC and renewable energy to accelerate the adoption of production, storage and distribution of clean electricity. How will we achieve this? In its proof-of-principle phase, the EoCoE consortium developed a comprehensive, structured support pathway for enhancing the HPC capability of energy-oriented numerical models, from simple entry-level parallelism to fully-fledged exascale readiness. At the top end of this scale, promising applications from each energy domain have been selected to form the basis of 5 new Energy Science Challenges:
    * Wind turbine modelling, from detailed understanding single turbine dynamics to flow across entire wind farms in complex terrain;
    * Energy Meteorology, where probabilistic forecasting is needed to predict the production efficiency of solar and wind parks and their impact on energy trading across the grid;
    * Design and study of new energy materials for photovoltaic cells, batteries and super-capacitors;
    * Water for energy to manage geothermal and hydro-power including the influence of climate change on these resources;
    * And fusion for energy, where the mandatory kinetic modelling of plasma turbulence and transport from the core to the edge of complex tokamak magnetic geometries requires exascale resources.

9.2.2. 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.

9.3. International Initiatives

9.3.1. Inria International Labs

9.3.1.1. 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).

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

9.3.2.1. UNIFY

- Title: Intelligent Unified Data Services for Hybrid Workflows Combining Compute-Intensive Simulations and Data-Intensive Analytics at Extreme Scales
- Partners:
  - Inria teams: KerData, DataMove
  - Argonne National Lab (Tom PETERKA)
- Duration: 2019-2021

9.3.3. Participation in Other International Programs

9.3.3.1. STIC AmSud SAQED

- Title: Scalable Approximate Query Evaluation on Document Inverted Files for GPU based Big-Data Applications
- International Partner:
  - Universidad Nacional de San Luis - UNSL, Argentina
  - Universidad de Santiago de Chile - USACH, Chile
  - Universidade Federal de São Carlos - UFSCAR, Brazil
- Duration: 2019-2021
- Develop efficient and scalable approximate search and document similarity evaluation on large datasets based on document inverted files using high performance computing and GPUs.

9.3.3.2. LICIA

- Title: International Laboratory in High Performance and Ubiquitous Computing
- International Partner (Institution - Laboratory - Researcher):
  - UFRGS (Brazil)
- Duration: Funded by CNRS in 2011-2018, by Univ Grenoble Alpes for 2019-2020.
- 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.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Professor visit: Alfredo Goldman, Professor at Universidade de São Paulo, visited Datamove from June to July 2019.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. 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 HeteroPar.

10.1.1.2. Member of the Organizing Committees

- Member of Organizing committee of HPML Workshop, co-located with CCGRID 2019
- Member of organization committee of Conv'2019

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- Vice-Chair of Algorithms track of IPDPS 2019
- Workshop Chair at Hetero-Par 2019
- Member of the steering committee of Euro-Par 2019
- Program co-chair of ISAV 2019 (workshop of SC)
- Member of steering committee of EGPV 2019

10.1.2.2. Member of the Conference Program Committees

- CCGrid 2019 (19th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing), May, Larcana, Cyprus.
- PARCO 2019 (International Conference on Parallel Computing), September, Prague, Czech Republic.
- PPAM 2019 (13th International Conference on Parallel Processing and Applied Mathematics), September, Bialystok, Poland.
- STACS 2019 (36th International Symposium on Theoretical Aspects of Computer Science), March, Berlin, Germany.
- Euro-Par 2019, August, Göttingen, Germany.
- IESM, September, Shanghai, China.
- ISAV 2019 (In Situ Infrastructures for Enabling Extreme-scale Analysis and Visualization), November, Denver, US.
- SBAC-PAD 2019 (International Symposium on Computer Architecture and High Performance Computing), October, Campo Grande, Brazil.
- HiPC 2019 (26th IEEE International Conference on High Performance Computing, Data, and Analytics), December, Hyderabad, India.

- ALGOCLOUD 2019 (5th International Symposium on Algorithmic Aspects of Cloud Computing), September, Munich, Germany.

10.1.3. Journal

10.1.3.1. 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.

10.1.4. Research Administration

- Director of Pôle MSTIC of COMUE Univ. Grenoble-Alpes.

10.2. Teaching - Supervision - Juries

10.2.1. 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. 90 hours per year. Combinatorial scientific computing in Master at ENS Lyon, Algorithmic in Licence at Grenoble INP and Algorithmic at Univ. Grenoble-Alpes.
- 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. 237 hours per year. Master (M1/2nd year and M2/3rd year) at Engineering school ENSIMAG, Grenoble-INP, Univ Grenoble Alpes.
- 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).

10.2.2. Supervision

- PhD: Danilo Carastan Dos Santos, Apprentissage sur heuristiques simples pour l'ordonnancement online de tâches parallèles, Univ Grenoble Alpes and Federal University of ABC, Brazil (co-tutelle), November, 27th 2019. Advisers: Denis Trystam and Raphael Yokoingawa De Camargo.
- PhD: Michael Mercier, Resource Management and Job Scheduling in HPC–Cloud environments towards the Big Data era, Univ. Grenoble Alpes. Started October 2016. July, 1st 2019. Advisers: Olivier Richard and Bruno Raffin.
- 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, Olivier Richard
- 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, Arnaud Legrand.
- PhD in progress: Sebastian Friedemann, Large Scale Data Assimilation, Adviser: Bruno Raffin.
- PhD in progress: Vincent Fagnon, Analyse de politique d'ordonnancement dynamique pour objets mobiles , Adviser: Denis Trystram
- PhD in progress: Angan Mitra, Theoretical and implementation challenges in Lifelong Learning and Edge Computing , Adviser: Denis Trystram
- PhD in progress: Paul Youssef, Coresets Analysis , Adviser: Denis Trystram
- PhD in progress: Ioannis Panagiotas, High performance algorithms for big data graph and hypergraph problems, Advisers: Bora Uçar (LIP) and Fanny Dufossé

10.2.3. Juries

- PhD Defense of Malin Rau, *Useful structures and how to find them*. Kiel (Allemagne). May, 24th, 2019. Referee.
- PhD Defense of Mohamad El Sayah, Random Generation for the Performance Evaluation of Scheduling Algorithms, November 20th 2019, Université de Bourgogne Franche-Comté. Member.
- PhD Defense of David Deibe, Geospatial Processing and Visualization of Point Clouds: from GPUs to Big Data Technologies, Univ. of A Coruna, December 2019. Referee
- PhD Defense of Charles Gueunet, Calcul Haute Performance pour l'Analyse Topologique de Données par Ensembles de Niveaux, Feburary 15th 2019, Sorbonne Université
- PhD Defense of Jorge Veiga Fachal, Evaluation and Optimization of Big Data Processing on High Performance Computing Systems, Univ. of A Coruna, Fev 2019. Jury.

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- Commission des emplois scientifiques du centre GRA

10.3.2. Interventions

- Fête de la Science, Grenoble
- CS unplugged for schoolchildren, Inria RA

11. 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

Articles in International Peer-Reviewed Journal

[2] G. AUPY, B. GOGLIN, V. HONORÉ, B. RAFFIN.*Modeling High-throughput Applications for in situ Analytics*, in "International Journal of High Performance Computing Applications", April 2019, vol. 33, n^{o} 6, p. 1185-1200, forthcoming [*DOI :* 10.1177/1094342019847263], https://hal.inria.fr/hal-02091340

[3] J. LUCA BEZ, A. RAMOS CARNEIRO, P. J. PAVAN, V. SOLDERA GIRELLI, F. ZANON BOITO, B. A. FAGUNDES, C. OSTHOFF, P. LEITE DA SILVA DIAS, J.-F. MÉHAUT, P. O. NAVAUX.*I/O Performance of the Santos Dumont Supercomputer*, in "International Journal of High Performance Computing Applications", 2019, p. 1-17, forthcoming [*DOI :* 10.1177/1094342019868526], https://hal.inria.fr/hal-02270908

[4] P. DE SOUZA, L. SILVESTER, A. DA SILVA, C. FERNANDES, T. RODRIGUES, S. PAUL, P. CAMARGO, R. WOJCIESZAK.*Exploiting the Synergetic Behavior of PtPd Bimetallic Catalysts in the Selective Hydrogenation of Glucose and Furfural*, in "Catalysts", February 2019, vol. 9, n° 2, p. 1-14 [*DOI :* 10.3390/CATAL9020132], https://hal.archives-ouvertes.fr/hal-02324316

Invited Conferences

[5] C. MOMMESSIN, G. LUCARELLI, D. TRYSTRAM.*Scheduling at the Edge*, in "14th Scheduling for Large Scale Systems Workshop", Bordeaux, France, June 2019, https://hal.inria.fr/hal-02459646

[6] D. TRYSTRAM, G. LUCARELLI, C. MOMMESSIN, Y. NGOKO.*Challenges for scheduling at the Edge*, in "Workshop on Mathematical Challenges in Scheduling Theory", Sanya, China, October 2019, https://hal.inria.fr/hal-02459551

International Conferences with Proceedings

[7] F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Effective heuristics for matchings in hypergraphs*, in "SEA2 2019 - International Symposium on Experimental Algorithms - Special Event", Kalamata, Greece, Springer, June 2019, p. 248-264 [*DOI :* 10.1007/978-3-030-34029-2_17], https://hal.inria.fr/hal-02417475

[8] J. LUCA BEZ, F. ZANON BOITO, R. NOU, A. MIRANDA, T. CORTES, P. O. NAVAUX.*Detecting I/O Access Patterns of HPC Workloads at Runtime*, in "SBAC-PAD 2019 - International Symposium on Computer Architecture and High Performance Computing", Campo Grande, Brazil, October 2019, https://hal.inria.fr/hal-02276191

[9] P. J. PAVAN, J. LUCA BEZ, M. S. SERPA, F. ZANON BOITO, P. O. NAVAUX.*An Unsupervised Learning Approach for I/O Behavior Characterization*, in "SBAC-PAD 2019 - International Symposium on Computer Architecture and High Performance Computing", Campo Grande, Brazil, October 2019, p. 1-8, https://hal.inria.fr/hal-02276230

[10] A. G. YABO, B. ROBU, O. RICHARD, B. BZEZNIK, E. RUTTEN.*A control-theory approach for cluster autonomic management: maximizing usage while avoiding overload*, in "CCTA 2019 - 3rd IEEE Conference on Control Technology and Applications", Hong-Kong, China, Proceedings of the 2019 IEEE Conference on Control Technology and Applications (CCTA), IEEE, August 2019, p. 189-195, https://hal.archives-ouvertes.fr/hal-02294272

[11] F. ZANON BOITO, J.-F. MÉHAUT, T. DEUTSCH, B. VIDEAU, F. DESPREZ.*Instrumental Data Management and Scientific Workflow Execution: the CEA case study*, in "IPDPSW 2019 - International Parallel and Distributed Processing Symposium Workshops", Rio de Janeiro, Brazil, IEEE, May 2019, p. 850-857 [*DOI :* 10.1109/IPDPSW.2019.00139], https://hal.inria.fr/hal-02076963

Conferences without Proceedings

[12] *Best Paper*
D. CARASTAN-SANTOS, R. Y. DE CAMARGO, D. TRYSTRAM, S. ZRIGUI.*One can only gain by replacing EASY Backfilling: A simple scheduling policies case study*, in "CCGrid 2019 - International Symposium in Cluster, Cloud, and Grid Computing", Larnaca, Cyprus, IEEE, May 2019, p. 1-10 [*DOI :* 10.1109/CCGRID.2019.00010], https://hal.archives-ouvertes.fr/hal-02237895.

[13] A. LEGRAND, D. TRYSTRAM, S. ZRIGUI.*Adapting Batch Scheduling to Workload Characteristics: What can we expect From Online Learning?*, in "IPDPS 2019 - 33rd IEEE International Parallel & Distributed Processing Symposium", Rio de Janeiro, Brazil, IEEE, May 2019, p. 686-695 [*DOI :* 10.1109/IPDPS.2019.00077], https://hal.archives-ouvertes.fr/hal-02044903

[14] Y. NGOKO, N. SAINTHERANT, C. CÉRIN, D. TRYSTRAM.*A Methodology for Handling Data Movements by Anticipation: Position Paper*, in "IPDPSW 2018 - International Parallel and Distributed Processing Symposium Workshops", Vancouver, Canada, IEEE, December 2019, p. 134-145 [*DOI :* 10.1109/IPDPSW.2018.00192], https://hal.archives-ouvertes.fr/hal-02066185

[15] *Best Paper*
F. ZANON BOITO, R. NOU, L. LIMA PILLA, J. LUCA BEZ, J.-F. MÉHAUT, T. CORTES, P. O. NAVAUX.*On server-side file access pattern matching*, in "HPCS 2019 - 17th International Conference on High Performance Computing & Simulation", Dublin, Ireland, IEEE, 2019, p. 1-8, outstanding paper award, https://hal.inria.fr/hal-02079899.

Research Reports

[16] A. BAUSKAR, A. DA SILVA, A. LEBRE, C. MOMMESSIN, P. NEYRON, Y. NGOKO, Y. RICORDEL, D. TRYSTRAM, A. VAN KEMPEN.*Investigating Placement Challenges in Edge Infrastructures through a Common Simulator*, DATAMOVE ; STACS ; DAPI IMT Atlantique, July 2019, n^{o} RR-9282, p. 1-16, https://hal.inria.fr/hal-02153203

[17] A. RIBES, T. TERRAZ, B. IOOSS, Y. FOURNIER, B. RAFFIN.*Large scale in transit computation of quantiles for ensemble runs*, EDF R&D, June 2019, https://arxiv.org/abs/1905.04180 , https://hal.inria.fr/hal-02016828

Other Publications

[18] M. KHATIRI, D. TRYSTRAM, F. WAGNER.*Work Stealing Simulator*, January 2020, https://arxiv.org/abs/1910.02803 - working paper or preprint [*DOI :* 10.02803], https://hal.archives-ouvertes.fr/hal-02444049

[19] J. LUCA BEZ, F. ZANON BOITO, R. NOU, A. MIRANDA, T. CORTES, P. O. NAVAUX.*Adaptive Request Scheduling for the I/O Forwarding Layer using Reinforcement Learning*, October 2019, working paper or preprint, https://hal.inria.fr/hal-01994677

[20] C. MOMMESSIN, O. BEAUMONT, L.-C. CANON, L. EYRAUD-DUBOIS, G. LUCARELLI, L. MARCHAL, B. SIMON, D. TRYSTRAM.*Scheduling on Two Types of Resources: a Survey*, January 2020, https://arxiv.org/abs/1909.11365 - working paper or preprint, https://hal.inria.fr/hal-02432381

[21] S. ZRIGUI, R. Y. DE CAMARGO, D. TRYSTRAM, A. LEGRAND.*Improving the Performance of Batch Schedulers Using Online Job Size Classification*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02334116

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

Frederick Douzet [Univ Vincennes-Saint Denis]
Olivier Hamant [INRA, HDR]
Kavé Salamatian [Univ Savoie Mont-Blanc, HDR]

PhD Student

Colin Gerard [Inria, PhD Student]

Administrative Assistant

Sylvie Boyer [Inria, Administrative Assistant]

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.

3.5. Cyberstrategy

- Geopolitics of BGP
- Cyberstrategy of infrastructures
- Internet content control

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. 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. New Software and Platforms

5.1. 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.

5.1.1. BGP Monitoring platform

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.

5.1.2. Atlas of Data

A platform to visualize data flows over the planet is under construction. It can be accessed online at https://theatlasofdata.earth/.

5.1.3. Observatory of foreign influence on social media

This observatory is monitoring on twitter and facebook the evolution of foreign influence. It is based on a twitter collection platform that is using an extensive database of foreign actors to detect and monitor foreign interference.

6. New Results

6.1. 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 [5], we showed how the movement above ground of the intermediation activity transforms territories.

6.2. 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 [4]. 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 specificity of Asia in these transformations [6].

6.3. Network data analytics

In collaboration with the Chinese Academy of Sciences, we worked on packet processing algorithmic for high speed network measurements. In [1] a packet capture archive system is developed and described. 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 [3].

6.4. Geopolitics of BGP

We have investigated the logical layer of cyberspace through an analysis of the structure of connectivity and the Border Gateway Protocol (BGP). This protocol has been leveraged by countries to control the flow of information or for active strategic purposes. We focused on several countries and characterized their strategies by linking them to current architectures and understanding their resilience in times of crisis. We focus on the case of Iran and uncovered the deep changes that has happened in the past 5 years. This study was premonitory as we observed in november 2019 the full extend of the strategy with the large scale internet disruptions. This generates a lot of mediatic coverage.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Grants with Industry

The PhD Thesis of Colin Gerard is funded through a contract with DGA (Ministry of Defense).

8. Partnerships and Cooperations

8.1. 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.

8.2. National Initiatives

- Chaire Castex, Ecole Militaire, Paris.
- AMNECYS (Alpine Multidisciplinary NEtwork on CYber-security Studies), University of Grenoble-Alpes.
- GEODE Research team on Geopolitics.
- Kavé Salamatian in co-leading the chair "AI and society" of the MIAI institute of University of Grenoble Alps.

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. 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.
- Nippon Institute of Computing Technology, Tokyo, Japan
- Cyber Civilisation Research Center at Keio University, Tokyo, Japan

8.4. International Research Visitors

8.4.1. Visits to International Teams

8.4.1.1. Research Stays Abroad

Stéphane Grumbach has been visiting scientist at the Research Institute on Humanity and Nature, RIHN, in Kyoto for a semester in 2018/2019.

9. Dissemination

9.1. 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.

9.1.1. Scientific Events: Selection

9.1.1.1. Member of the Conference Program Committees

Kavé Salamatian has been TPC member for the below conferences:

- Infocom 2020
- Infocom 2019
- Aintec 2019

9.1.2. Journal

9.1.2.1. Reviewer - Reviewing Activities

Kavé Salamatian has been reviewer for the below journals:

- IEEE/ACM Transactions on Networking
- IEEE Transactions on Communication
- IEEE Transactions on Networked Systems Management
- Computer Communications
- IEEE transactions on Information Theory

9.2. Teaching - Supervision - Juries

9.2.1. 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.

9.2.2. Supervision

PhD: Jingxiu Su, DNS data analysis, 2016-2019, directeur de thèse Kave Salamatian. Defense 30/12/2019

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, co-Directors: Frederick Douzet, Kavé Salamatian

Phd in progress: Xinyi Zhang, IA and cybersecurity, 2019-, Advisor: Kavé Salamatian

Phd in progress: Ali Marandi, Efficient diffusion in NDN, 2016-2020, Advisor: Kavé Salamatian

9.3. Popularization

Various publications have appeared in journals accessible to a larger audience. In particular, our research work has been featured in a full page of Le Monde newspaper in July 2019 (https://www.lemonde.fr/international/visuel/2018/07/23/cyberespace-la-guerre-mondiale-des-donnees_5334901_3210.html).

Kavé Salamatian did several radio intervention, in particular, he was invited in the "Methode Scientifique" broadcasting program in November 2019.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] J. Su, Z. Li, S. Grumbach, M. Ikram, K. Salamatian, G. Xie.*A cartography of web tracking using DNS records*, in "Computer Communications", January 2019, vol. 134, p. 83-95 [*DOI :* 10.1016/j.comcom.2018.11.008], https://hal.archives-ouvertes.fr/hal-01950449

Conferences without Proceedings

[2] H. Jiang, Y. Yang, H. Guan, G. Xie, K. Salamatian.*A Massively Multi-Tenant Virtualized Network Intrusion Prevention Service on NFV Platform*, in "ICCCN 2019 - 28th International Conference on Computer Communications and Networks", Valencia, Spain, IEEE, July 2019, p. 1-9 [*DOI :* 10.1109/ICCCN.2019.8846924], https://hal.archives-ouvertes.fr/hal-02425277

[3] A. MARANDI, T. BRAUN, K. SALAMATIAN, N. THOMOS.*Pull-based Bloom Filter-based Routing for Information-Centric Networks*, in "CCNC 2019 - IEEE Consumer Communications & Networking Conference", Las Vegas, United States, IEEE, January 2019, p. 1-6 [*DOI :* 10.1109/CCNC.2019.8651713], https://hal.archives-ouvertes.fr/hal-02425271

Scientific Books (or Scientific Book chapters)

[4] S. GRUMBACH.*L'Entropie du Numérocène Quelques reflexions sur la Révolution Numérique et l'Anthropocène*, in "Anthropocène, à l'école de l'indiscipline", S. GRUMBACH, O. HAMANT, J. L. GALL, I. NEGRUTIU (editors), January 2019, p. 1-13, https://hal.inria.fr/hal-02021716

[5] S. GRUMBACH.*New Intermediaries: Extra-territorial Platforms*, in "The Digital Era 2: Political Economy Revisited", J.-P. CHAMOUX (editor), January 2019, https://hal.inria.fr/hal-02411656

[6] S. GRUMBACH.*Digital control and the global ecosystem: can the governance of the anthropocene be designed in East Asia?*, in "M. Terada and D. Niles (eds.), Questioning the Anthropocene: From the view point of bottom up and region (tentative title), Kyoto: Kyoto University Press, forthcomig (to be published in 2020).(2020)", Kyoto University Press, 2020, forthcoming, https://hal.inria.fr/hal-02411654

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
B1.1.3. - Developmental biology
B1.1.4. - Genetics and genomics
B1.1.5. - Immunology
B1.1.6. - Evolutionnary biology
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]
Vincent Calvez [CNRS, Senior Researcher, from Jun 2019, HDR]
Olivier Gandrillon [CNRS, Senior Researcher, HDR]
Simon Girel [Univ de Lyon, Researcher, until Aug 2019]
Thomas Lepoutre [Inria, Researcher, HDR]
Vitaly Volpert [CNRS, Senior Researcher, HDR]

Faculty Members

Thibault Espinasse [Univ Claude Bernard, Associate Professor, from Jun 2019]
Laurent Pujo-Menjouet [Univ Claude Bernard, Associate Professor, HDR]
Léon Tine [Univ Claude Bernard, Associate Professor]

External Collaborator

Fabien Crauste [CNRS, Senior Researcher, HDR]

PhD Students

Ghada Abi Younes [Univ Claude Bernard, PhD Student, from Oct 2019]
Loïs Boullu [Univ Claude Bernard, PhD Student, until Aug 2019]
Aurélien Canet [PhD Student, granted by CIFRE]
Kyriaki Dariva [Univ Claude Bernard, PhD Student]
Léonard Dekens [Univ Claude Bernard, PhD Student, from Sep 2019]
Mete Demircigil [École Normale Supérieure de Paris, PhD Student, from Sep 2019]
Ronan Duchesne [École Normale Supérieure de Lyon, PhD Student]
Cheikh Gueye [PhD Student, from Oct 2019]
Alexey Koshkin [Inria, PhD Student]
Paul Lemarre [Univ Claude Bernard, PhD Student]
Elias Ventre [École Normale Supérieure de Lyon, PhD Student, from Oct 2019]

Administrative Assistant

Claire Sauer [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Presentation

Dracula is a joint research team between Inria, Université 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.

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.

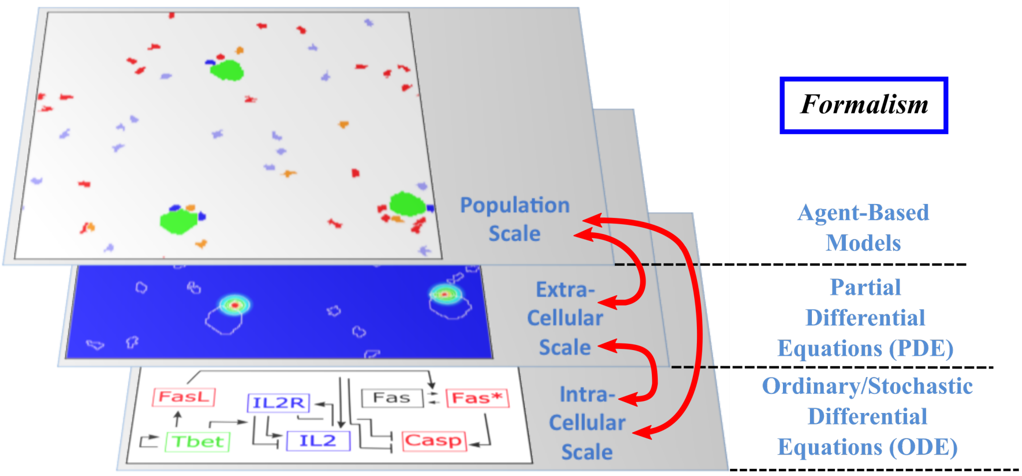


Figure 1. Scheme of multi-scale models of cell dynamics

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://team.inria.fr/mamba/fr/), 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 (François Morlé, 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 Grégoire 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://team.inria.fr/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 ([40]). 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 ([40]), 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 [41] 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 ([37], [38] and [39]). 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 [37]. 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. Mathematical models describing the interaction between cancer and immune cells in the lymph node

To study the interplay between tumor progression and the immune response, we develop in [5] two new models describing the interaction between cancer and immune cells in the lymph node. The first model consists of partial differential equations (PDEs) describing the populations of the different types of cells. The second one is a hybrid discrete-continuous model integrating the mechanical and biochemical mechanisms that define the tumor-immune interplay in the lymph node. We use the continuous model to determine the conditions of the regimes of tumor-immune interaction in the lymph node. While we use the hybrid model to elucidate the mechanisms that contribute to the development of each regime at the cellular and tissue levels. We study the dynamics of tumor growth in the absence of immune cells. Then, we consider the immune response and we quantify the effects of immunosuppression and local EGF concentration on the fate of the tumor.

5.2. WASABI: a dynamic iterative framework for gene regulatory network inference

Background Inference of gene regulatory networks from gene expression data has been a long-standing and notoriously difficult task in systems biology. Recently, single-cell transcriptomic data have been massively used for gene regulatory network inference, with both successes and limitations. In the work [8], we propose an iterative algorithm called WASABI, dedicated to inferring a causal dynamical network from time-stamped single-cell data, which tackles some of the limitations associated with current approaches. We first introduce the concept of waves, which posits that the information provided by an external stimulus will affect genes one-by-one through a cascade, like waves spreading through a network. This concept allows us to infer the network one gene at a time, after genes have been ordered regarding their time of regulation. We then demonstrate the ability of WASABI to correctly infer small networks, which have been simulated in silico using a mechanistic model consisting of coupled piecewise-deterministic Markov processes for the proper description of gene expression at the single-cell level. We finally apply WASABI on in vitro generated data on an avian model of erythroid differentiation. The structure of the resulting gene regulatory network sheds a new light on the molecular mechanisms controlling this process. In particular, we find no evidence for hub genes and a much more distributed network structure than expected. Interestingly, we find that a majority of genes are under the direct control of the differentiation-inducing stimulus. Conclusions Together, these results demonstrate WASABI versatility and ability to tackle some general gene regulatory networks inference issues. It is our hope that WASABI will prove useful in helping biologists to fully exploit the power of time-stamped single-cell data.

5.3. A multiscale model of platelet-fibrin thrombus growth in the flow

Thrombosis is a life-threatening clinical condition characterized by the obstruction of blood flow in a vessel due to the formation of a large thrombus. The pathogenesis of thrombosis is complex because the type of formed clots depends on the location and function of the corresponding blood vessel. To explore this phenomenon, we develop in [9] a novel multiscale model of platelet-fibrin thrombus growth in the flow. In this model, the regulatory network of the coagulation cascade is described by partial differential equations. Blood flow is introduced using the Navier–Stokes equations and the clot is treated as a porous medium. Platelets are represented as discrete spheres that migrate with the flow. Each platelet can attach to the thrombus, aggregate, become activated, express proteins on its surface, detach, and/or become non-adhesive. The interaction of platelets with blood flow is captured using the Immersed Boundary Method (IBM). We use the model to investigate the role of flow conditions in shaping the dynamics of venous and arterial thrombi. We describe the formation of red and white thrombi under venous and arterial flow respectively and highlight the main characteristics of each type. We identify the different regimes of normal and pathological thrombus formation depending on flow conditions.

5.4. Mathematical modeling of platelet production

- In [10], we analyze the existence of oscillating solutions and the asymptotic convergence for a nonlinear delay differential equation arising from the modeling of platelet production. We consider four different cell compartments corresponding to different cell maturity levels: stem cells, megakaryocytic progenitors, megakaryocytes, and platelets compartments, and the quantity of circulating thrombopoietin (TPO), a platelet regulation cytokine.
- In [11], we analyze the stability of a differential equation with two delays originating from a model for a population divided into two subpopulations, immature and mature, and we apply this analysis to a model for platelet production. The dynamics of mature individuals is described by the following nonlinear differential equation with two delays: $x'(t) = -\lambda x(t) + g(x(t-\tau_1)) - g(x(t-\tau_1-\tau_2))e^{-\lambda\tau_2}$. The method of D-decomposition is used to compute the stability regions for a given equilibrium. The center manifold theory is used to investigate the steady-state bifurcation and the Hopf bifurcation. Similarly, analysis of the center manifold associated with a double bifurcation is used to identify a set of parameters such that the solution is a torus in the pseudo- phase space. Finally, the results of the local stability analysis are used to study the impact of an increase of the death rate γ or of a decrease of the survival time τ_2 of platelets on the onset of oscillations. We show that the stability is lost through a small decrease of survival time (from 8.4 to 7 days), or through an important increase of the death rate (from 0.05 to 0.625 days^{-1}).
- In [12], we analyze the stability of a system of differential equations with a threshold-defined delay arising from a model for platelet production. We consider a maturity-structured population of megakaryocyte progenitors and an age-structured population of platelets, where the cytokine thrombopoietin (TPO) increases the maturation rate of progenitors. Using the quasi-steady-state approximation for TPO dynamics and the method of characteristics, partial differential equations are reduced to a system of two differential equations with a state-dependent delay accounting for

the variable maturation rate. We start by introducing the model and proving the positivity and boundedness of the solutions. Then we use a change of variables to obtain an equivalent system of two differential equations with a constant delay, from which we prove existence and uniqueness of the solution. As linearization around the unique positive steady state yields a transcendental characteristic equation of third degree, we introduce the main result, a new framework for stability analysis on models with fixed delays. This framework is then used to describe the stability of the megakaryopoiesis with respect to its parameters. Finally, with parameters being obtained and estimated from data, we give an example in which oscillations appear when the death rate of progenitors is increased 10-fold.

5.5. Nonlinear analysis of a model for yeast cell communication

In [13], we study the non-linear stability of a coupled system of two non-linear transport-diffusion equations set in two opposite half-lines. This system describes some aspects of yeast pairwise cellular communication, through the concentration of some protein in the cell bulk and at the cell boundary. We show that it is of bistable type, provided that the intensity of active molecular transport is large enough. We prove the non-linear stability of the most concentrated steady state, for large initial data, by entropy and comparison techniques. For small initial data we prove the self-similar decay of the molecular concentration towards zero. Informally speaking, the rise of a dialog between yeast cells requires enough active molecular transport in this model. Besides, if the cells do not invest enough in the communication with their partner, they do not respond to each other; but a sufficient initial input from each cell in the dialog leads to the establishment of a stable activated state in both cells.

5.6. Alzheimer's disease and prion: An in vitro mathematical model

Alzheimer's disease (AD) is a fatal incurable disease leading to progressive neuron destruction. AD is caused in part by the accumulation in the brain of $A\beta$ monomers aggregating into oligomers and fibrils. Oligomers are amongst the most toxic structures as they can interact with neurons via membrane receptors, including PrP^c proteins. This interaction leads to the misconformation of PrP^c into pathogenic oligomeric prions, PrP^{ol}. In [14], we develop a model describing in vitro $A\beta$ polymerization process. We include interactions between oligomers and PrP^c, causing the misconformation of PrP^c into PrP^{ol}. The model consists of nine equations, including size structured transport equations, ordinary differential equations and delayed differential equations. We analyse the well-posedness of the model and prove the existence and uniqueness of solutions of our model using Schauder fixed point theorem and Cauchy-Lipschitz theorem. Numerical simulations are also provided to give an illustration of the profiles that can be obtained with this model.

5.7. Calibration, Selection and Identifiability Analysis of a Mathematical Model of the in vitro Erythropoiesis in Normal and Perturbed Contexts

The in vivo erythropoiesis, which is the generation of mature red blood cells in the bone marrow of whole organisms, has been described by a variety of mathematical models in the past decades. However, the in vitro erythropoiesis, which produces red blood cells in cultures, has received much less attention from the modelling community. In the paper [15], we propose the first mathematical model of in vitro erythropoiesis. We start by formulating different models and select the best one at fitting experimental data of in vitro erythropoietic differentiation obtained from chicken erythroid progenitor cells. It is based on a set of linear ODE, describing 3 hypothetical populations of cells at different stages of differentiation. We then compute confidence intervals for all of its parameters estimates, and conclude that our model is fully identifiable. Finally, we use this model to compute the effect of a chemical drug called Rapamycin, which affects all states of differentiation in the culture, and relate these effects to specific parameter variations. We provide the first model for the kinetics of in vitro cellular differentiation which is proven to be identifiable. It will serve as a basis for a model which will better account for the variability which is inherent to the experimental protocol used for the model calibration.

5.8. Model-based assessment of the role of uneven partitioning of molecular content on heterogeneity and regulation of differentiation in CD8 T-cell immune responses

Activation of naive CD8 T-cells can lead to the generation of multiple effector and memory subsets. Multiple parameters associated with activation conditions are involved in generating this diversity that is associated with heterogeneous molecular contents of activated cells. Although naive cell polarisation upon antigenic stimulation and the resulting asymmetric division are known to be a major source of heterogeneity and cell fate regulation, the consequences of stochastic uneven partitioning of molecular content upon subsequent divisions remain unclear yet. In [16], we aim at studying the impact of uneven partitioning on molecular-content heterogeneity and then on the immune response dynamics at the cellular level. To do so, we introduce a multiscale mathematical model of the CD8 T-cell immune response in the lymph node. In the model, cells are described as agents evolving and interacting in a 2D environment while a set of differential equations, embedded in each cell, models the regulation of intra and extracellular proteins involved in cell differentiation. Based on the analysis of in silico data at the single cell level, we 1 show that immune response dynamics can be explained by the molecular-content heterogeneity generated by uneven partitioning at cell division. In particular, uneven partitioning acts as a regulator of cell differentiation and induces the emergence of two coexisting sub-populations of cells exhibiting antagonistic fates. We show that the degree of unevenness of molecular partitioning, along all cell divisions, affects the outcome of the immune response and can promote the generation of memory cells.

5.9. Spatial lymphocyte dynamics in lymph nodes predicts the cytotoxic T-Cell frequency needed for HIV infection control

The surveillance of host body tissues by immune cells is central for mediating their defense function. In vivo imaging technologies have been used to quantitatively characterize target cell scanning and migration of lymphocytes within lymph nodes (LNs). The translation of these quantitative insights into a predictive understanding of immune system functioning in response to various perturbations critically depends on computational tools linking the individual immune cell properties with the emergent behavior of the immune system. By choosing the Newtonian second law for the governing equations, we developed in [17] a broadly applicable mathematical model linking individual and coordinated T-cell behaviors. The spatial cell dynamics is described by a superposition of autonomous locomotion, intercellular interaction, and viscous damping processes. The model is calibrated using in vivo data on T-cell motility metrics in LNs such as the translational speeds, turning angle speeds, andmeandering indices. Themodel is applied to predict the impact of T-cell motility on protection against HIV infection, i.e., to estimate the threshold frequency of HIV-specific cytotoxic T cells (CTLs) that is required to detect productively infected cells before the release of viral particles starts. With this, it provides guidance for HIV vaccine studies allowing for the migration of cells in fibrotic LNs.

5.10. Drugs modulating stochastic gene expression affect the erythroid differentiation process

To better understand the mechanisms behind cells decision-making to differentiate, we assessed in [18] the influence of stochastic gene expression (SGE) modulation on the erythroid dif- ferentiation process. It has been suggested that stochastic gene expression has a role in cell fate decision-making which is revealed by single-cell analyses but studies dedicated to demonstrate the consistency of this link are still lacking. Recent observations showed that SGE significantly increased during differentiation and a few showed that an increase of the level of SGE is accompanied by an increase in the differentiation process. However, a consistent relation in both increasing and decreasing directions has never been shown in the same cellular system. Such demonstration would require to be able to experimentally manipulate simultaneously the level of SGE and cell differentiation in order to observe if cell behavior matches with the current theory. We identified three drugs that modulate SGE in primary erythroid progenitor cells. Both Artemisinin and Indomethacin decreased SGE and reduced the amount of differentiated cells. On the contrary, a third component called MB-3 simultaneously increased the level of SGE and the amount of differentiated cells. We then used a dynamical modelling approach which confirmed that differentiation rates were indeed affected by the drug treatment. Using single-cell analysis and modeling tools, we provide experimental evidence that, in a physiologically relevant cellular system, SGE is linked to differentiation.

5.11. Stochastic gene expression with a multistate promoter: breaking down exact distributions

We consider in [19] a stochastic model of gene expression in which transcription depends on a multistate promoter, including the famous two-state model and refractory promoters as special cases, and focus on deriving the exact stationary distribution. Building upon several successful approaches, we present a more unified viewpoint that enables us to simplify and generalize existing results. In particular, the original jump process is deeply related to a multivariate piecewise-deterministic Markov process that may also be of interest beyond the biological field. In a very particular case of promoter configuration, this underlying process is shown to have a simple Dirichlet stationary distribution. In the general case, the corresponding marginal distributions extend the well-known class of Beta products, involving complex parameters that directly relate to spectral properties of the promoter transition matrix. Finally, we illustrate these results with biologically plausible examples.

5.12. Cell generation dynamics underlying naive T-cell homeostasis in adult humans

Thymic involution and proliferation of naive T-cells both contribute to shaping the naive T-cell repertoire as humans age, but a clear understanding of the roles of each throughout a human life span has been difficult to determine. By measuring nuclear bomb test-derived 14 C in genomic DNA, we determined in [22] the turnover rates of CD4 + and CD8 + naive T-cell populations and defined their dynamics in healthy individuals ranging from 20 to 65 years of age. We demonstrate that naive T-cell generation decreases with age because of a combination of declining peripheral division and thymic production during adulthood. Concomitant decline in T-cell loss compensates for decreased generation rates. We investigated putative mechanisms underlying age-related changes in homeostatic regulation of CD4+ naive T-cell turnover, using mass cytometry to profile candidate signaling pathways involved in T-cell activation and proliferation relative to CD31 expression, a marker of thymic proximity for the CD4+ naive T-cell population. We show that basal nuclear factor kB (NF-kB) phosphory-lation positively correlated with CD31 expression and thus is decreased in peripherally expanded naive T-cell clones. Functionally, we found that NF-kB signaling was essential for naive T-cell proliferation to the homeostatic growth factor interleukin (IL)-7, and reduced NF-kB phosphorylation in CD4 + CD31 - naive T cells is linked to reduced homeostatic proliferation potential. Our results reveal an age-related decline in naive T-cell turnover as a putative regulator of naive T-cell diversity and identify a molecular pathway that restricts proliferation of peripherally expanded naive T-cell clones that accumulate with age.

5.13. Erythroid differentiation displays a peak of energy consumption concomitant with glycolytic metabolism rearrangements

Our previous single-cell based gene expression analysis pointed out significant variations of LDHA level during erythroid differentiation. Deeper investigations highlighted that a metabolic switch occurred along differentiation of erythroid cells. More precisely we showed in [26] that self-renewing progenitors relied mostly upon lactate-productive glycolysis, and required LDHA activity, whereas differentiating cells, mainly involved mitochondrial oxidative phosphorylation (OXPHOS). These metabolic rearrangements were coming along with a particular temporary event, occurring within the first 24h of erythroid differentiation. The activity of glycolytic metabolism and OXPHOS rose jointly with oxgene consumption dedicated to ATP production at 12-24h of the differentiation process before lactate-productive glycolysis sharply fall down and energy needs decline. Finally, we demonstrated that the metabolic switch mediated through LDHA drop and OXPHOS upkeep might be necessary for erythroid differentiation. We also discuss the possibility that metabolism, gene expression and epigenetics could act together in a circular manner as a driving force for differentiation.

6. Partnerships and Cooperations

6.1. Regional Initiatives

- The Région ARA project INGERENCE dedicated to "INferring GEne REgulatory NEtworks from single CEll Data to improve vaccine design", 2018-2021.
  Participants: Olivier Gandrillon, Fabien Crauste [Coordinator].

6.2. National Initiatives

6.2.1. ANR

- ANR SinCity "Single cell transcriptomics on genealogically identified differentiating cells" (https://anr.fr/Projet-ANR-17-CE12-0031), 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.

6.2.2. Other projects

- Thomas Lepoutre is a member of the ERC MESOPROBIO (head Vincent Calvez) dedicated to "Mesoscopic models for propagation in biology", 2015-2020: (http://vcalvez.perso.math.cnrs.fr/mesoprobio.html).

6.3. European Initiatives

6.3.1. FP7 & H2020 Projects

- Olivier Gandrillon and Alexey Koshkin participate 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

6.4. International Initiatives

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

6.4.1.1. MathModelingHematopoiesis

Title: Mathematical modeling of hematopoietic stem cell dynamics in normal and pathological hematopoiesis with optimal control for drug therapy

International Partner (Institution - Laboratory - Researcher):

Presidency University, Kolkata (India) - Subhas Khajanchi

Start year: 2019

The project proposes to develop and analyse new mathematical models of Hematopoietic Stem Cell population dynamics in normal and pathological hematopoiesis. Two important questions will be explored in this project: i) the biological data concerning the hematopoiesis process evolves constantly, and new understanding modifies the established mathematical models, ii) modeling constraints us to simplify the complicated biological scenarios, which moving away from the reality, but enabling us to reach a certain comprehension of the hematopoiesis process.

The project will shed new light on the different physiological mechanisms that converge toward the continuous regeneration of blood cells, for example: the behavior of hematopoietic stem cells under stress conditions, the understanding of deregulation of erythropoiesis under drug treatments (this can lead to lack of red blood cells (anemia), or a surplus of red blood cells (erythrocytoses)), the appearance of oscillations in patients with Chronic Myeloid Leukemia (CML); Or, the overproduction of blasts in patients with Acute Myeloid Leukemia (AML)). The effect of the immune system and drug therapy in the presence of CML or AML will be included in the model and optimal control method will also be used.

6.4.2. Participation in Other International Programs

6.4.2.1. Indo-French Center of Applied Mathematics

Title: Mathematical modeling of hematopoiesis process in application to chronic and acute myelogenous leukemia

International Partner (Institution - Laboratory - Researcher):

Department of Mathematics - Presidency University, Kolkata (India) - Subhas Khajanchi

Duration: 2018 - 2021

Start year: 2018

6.5. International Research Visitors

6.5.1. Visits of International Scientists

Jairo Gomes da Silva, PhD student at Institute of Biosciences, São Paulo State University (UNESP), Botucatu, Brazil, visiting the team for 6 months (from September 2019 to February 2020).

6.5.2. Visits to International Teams

Paul Lemarre is visiting University of California, Merced, USA, in 2019-2020.

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 2019, Lyon, http://www.ens-lyon.fr/evenement/recherche/icsb-2018-19th-international-conference-systems-biology

7.1.1.2. Member of the organizing Committees

- Laurent Pujo-Menjouet, The Society for Mathematical Biology, Annual Meeting and Conference, Montreal, Canada, mini-symposium co-organizer, 21-26 July 2019

7.1.2. Journal

7.1.2.1. Member of the Editorial Boards

- Mostafa Adimy, Journal of Nonlinear Systems and Applications
- Olivier Gandrillon, Associate editor for BMC research notes
- Vincent Calvez, Journal of Mathematical Biology; SIAM Journal of Mathematical Analysis
- Laurent Pujo-Menjouet, Associate editor of PLOS ONE; Journal of Thoerical Biology; Mathematical modelling of natural phenomena

7.1.2.2. Reviewer - Reviewing Activities

- Mostafa Adimy, Canadian Journal of Mathematics; PLOS Neglected Tropical Diseases; Journal of Mathematical Modeling of Natural Phenomena
- Olivier Gandrillon, Plos Computational Biology; Progress in Biophysics and Molecular Biology; Journal of the Royal Society Interface; Journal of theoretical biology; Systems Biology and Applications; Genes
- Vincent Calvez, Inventiones Mathematicae; Duke Math Journal; Journal of the EMS; Nature Physics; Ecology Letters
- Laurent Pujo-Menjouet, Journal of Theoretical Biology; Plos One, Plos Computational Biology; Journal of Mathematical Modeling of Natural Phenomena; Journal of Mathematical Biology; Bulletin of Mathematical Biology; Computational and applied mathematics; Mathematical Biosciences and Engineering

7.1.3. Invited Talks

- Mostafa Adimy, International Conference on Mathematical Modelling in Biomedicine, Moscow - Russia, 30 Septembre - 04 Octobre, 2019, http://eng.rudn.ru/science/conferences/9370
- Mostafa Adimy, 9ème colloque, Tendances dans les Applications Mathématiques (TAMTAM), Tlemcen - Algeria, 24 - 27 February 2019, https://tamtam2019-univ-tlemcen.weebly.com/accueil.html
- Vincent Calvez, BIOMAT Granada, Patterns in Life and Social Sciences, Granada - Spain, June 17-19, http://www.ugr.es/~kinetic/biomat/
- Vincent Calvez, Colloque d'ouverture 50 ans du Laboratoire J.L. Lions, Roscoff - France, 4-8 mars 2019 https://ljll-roscoff.sciencesconf.org
- Vincent Calvez, Journées Aussois Chaire MMB, Aussois - France, 19 - 23 May http://www.cmap.polytechnique.fr/chaire-mmb/Aussois2019.html
- Vincent Calvez, LMS Research School, PDEs in Mathematical Biology: Modelling and Analysis, ICMS, Edinburgh - UK, 29 April – 3 May 2019, https://www.icms.org.uk/LMS_PDEsmathbio.php
- Laurent Pujo-Menjouet, Mathe´matiques e´tonnantes, SMF conference, Lyon, 9 December, https://smf.emath.fr/conference-coeur

7.1.4. Leadership within the Scientific Community

- Olivier Gandrillon, Director of BioSyL, the Federative Research Structure for Systems Biology attached to University of Lyon, http://www.biosyl.org
- Thomas Lepoutre, Head of the Groupe de Recherches CNRS MAMOVI on applied mathematical modelling in Life Sciences

7.1.5. Research Administration

- Léon Tine, Membre conseil du département de Mathématiques, UCBL 1
- Léon Tine, Co-responsable de l'enseignement TMB (Techniques Mathématiques de Base) du portail PCSI, UCBL 1
- Mostafa Adimy, Comité scientifique (COS) du centre Inria Rhône-Alpes
- Mostafa Adimy, Comité scientifique (CS) de l'Institut Camille Jordan, UCBL 1

- Mostafa Adimy, Comité des thèses de l'Institut Camille Jordan, UCBL 1
- Laurent Pujo-Menjouet, Responsable de la filière mathématiques pour la biologie et la médecine pour le master M2 Maths en Actions à l'UCBL 1
- Laurent Pujo-Menjouet, Correspondant mobilité international pour le département de mathématiques à l'UCBL 1
- Laurent Pujo-Menjouet, Directeur du portail mathématiques et informatique à l'UCBL 1
- Thomas Lepoutre member of the CORDI-S commission of Inria.

7.2. Teaching - Supervision - Juries

7.2.1. Teaching

- Licence: Laurent Pujo-Menjouet, Fondamentaux des mathématiques, 138h EQTD, L1, UCBL 1
- Licence: Laurent Pujo-Menjouet, 3ème année biosciences BIM, Systèmes Dynamiques et EDP, 45h EQTD, INSA Lyon
- Licence : Léon Tine, Techniques mathématiques de base, 53h (EqTD), niveau L0, UCBL 1
- Licence : Léon Tine, Techniques mathématiques de base, 62h (EqTD), niveau L1, UCBL 1
- Licence : Léon Tine, Initiation LaTex+ stage, 12h (EqTD), niveau L3, UCBL 1
- Licence : Vincent Calvez, Cours de math pour étudiants médecins, cursus Médecine-Sciences (2e année fac de médecine), 40h, Lyon Est/Sud
- Master : Samuel Bernard, Population Dynamics, 36h ETD, M2, UCBL 1
- Master : Mostafa Adimy, Population Dynamics, 9h ETD, M2, UCBL, UCBL 1
- Master : Mostafa Adimy, Epidemiology, 21h ETD, M2, UCBL, UCBL 1
- Master: Thomas Lepoutre, préparation à l'option pour l'agrégation, 45 h eq TD, M2 UCBL 1
- Master: Laurent Pujo-Menjouet, Systèmes Dynamiques, 72 h EQTD, M1, UCBL1
- Master: Laurent Pujo-Menjouet, Systèmes complexes: modelling biology and medicine, M2, 9h EQTD, ENS-Lyon
- Master: Laurent Pujo-Menjouet, 4ème année biosciences BIM: ED-EDP, 24h EQTD, INSA Lyon
- Master: Léon Tine, Maths en action, Remise à niveau analyse, 12h (EqTD), niveau M2, UCBL 1
- Master: Léon Tine, Maths en action, épidémiologie, 18h (EqTD), niveau M2, UCBL 1
- Master: Vincent Calvez, Modèles mathématiques et analyse pour Ecologie et Evolution, 24h, niveau M2 avancé, ENS-Lyon
- Master: Olivier Gandrillon, Systems Biology, 8h, niveau M2 Génpopath, UCBL 1

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, supervisors: 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, since 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, Décembre 2019, 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 in progress: Léonard Dekens, “Adaptation d’une population avec une reproduction sexuée à un environnement hétérogène : modélisation mathématique et analyse asymptotique dans un régime de petite variance”, Université Lyon 1, since September 2019, supervisor: Vincent Calvez.
- PhD in progress: Mete Demircigil, “Etude du mouvement collectif chez Dictyostelium Discoideum et autres espèces. Modélisation, Analyse et Simulation”, ENS-Lyon, since September 2019, supervisor: Vincent Calvez.
- PhD in progress: Ghada Abi Younes, “Modélisation mathématique des maladies inflammatoires”, Université Lyon, since November 2019, supervisor: Vitaly Volpert.
- PhD in progress: Elias Ventre, “Paysage et trajectoire - des réseaux de régulation génique aux trajectoires cellulaires", ENS - Lyon, since October 2019, supervisors: Olivier Gandrillons and Thibault Espinasse.
- PhD in progress: Cheikh Gueye, “Problèmes inverses pour l’estimation de paramètre de modèles mathématiques”, Université Lyon, since October 2019, supervisors: Laurent Pujo-Menjouet and Léon Tine.

7.2.3. Juries

- Laurent Pujo-Menjoeu: PhD of Hugo Martin, Étude de données et analyse de modèles intégro-différentiels en biologie cellulaire, Laboratoire Jacques-Louis Lions - Sorbonne Université, examiner.
- Vincent Calvez: PhD of Pierre Roux, Équations aux dérivées partielles de type Keller-Segel en dynamique des populations et de type Fokker-Planck en neurosciences, Laboratoire de mathématiques d’Orsay - Université Paris-Sud, reviewer
- Vincent Calvez: PhD of Xiaoming Fu, Reaction-diffusion Equations with Nonlinear and Nonlocal Advection Applied to Cell Co-culture, Institut de Mathématiques de Bordeaux, Université de Bordeaux, examiner
- Olivier Gandrillon: PhD of Ronan Duchesne, Erythroid differentiation in vitro under the lens of mathematical modelling, ENS-Lyon, examiner.

7.3. Popularization

7.3.1. Articles and contents

- Laurent Pujo-Menjouet : Radio (France Inter), la tête au carré, https://www.franceinter.fr/emissions/la-tete-au-carre/la-tete-au-carre-10-mai-2019
- Laurent Pujo-Menjouet : Radio (Europe 1), Wendy Bouchard, https://www.europe1.fr/emissions/on-fait-le-tour-de-la-question-avec-wendy-bouchard/wendy-bouchard-aujourdhui-alchimie-equation-formule-magique-lamour-est-il-le-fruit-du-hasard-3905606
- Laurent Pujo-Menjouet : Radio (Québec), https://www.qub.radio/balado?id=6bb8beb9-9bb3-493f-abe3-aa7100f60a73&episode=3f6e7a5c-b11a-45ff-aa28-aa96015f517b&fbclid=IwAR2DUNgo-QvKh3z9oGeslTJOnCa3f5GzTYWlIGrYSe5hrm7ydSgc1gW6Z5g
- Laurent Pujo-Menjouet : Radio Canada, https://ici.radio-canada.ca/premiere/emissions/region-zero-8/episodes/439145/audio-fil-du-vendredi-26-juillet-2019/10
- Laurent Pujo-Menjouet : Presse écrite (L’Express), https://www.lexpress.fr/actualite/sciences/le-livre-qui-veut-modeliser-l-amour-par-les-mathematiques_2075520.html
- Laurent Pujo-Menjouet : Presse écrite (Le Figaro), http://www.lefigaro.fr/sciences/le-bonheur-conjugal-c-est-simple-comme-une-equation-differentielle-20190506

- Laurent Pujo-Menjouet : Presse écrite (Slate), http://www.slate.fr/story/178140/maths-livre-jeu-amour-hasard-laurent-pujo-menjouet
- Laurent Pujo-Menjouet : Presse écrite (Le Parisien), http://www.leparisien.fr/societe/et-si-l-amour-n-etait-pas-le-fruit-du-hasard-10-06-2019-8090055.php
- Laurent Pujo-Menjouet : Presse écrite (Montreal Gazette), https://montrealgazette.com/news/local-news/equations-can-predict-the-success-of-relationships-mathematician-says
- Laurent Pujo-Menjouet : Presse écrite (Society), http://math.univ-lyon1.fr/~pujo/Society.pdf
- Laurent Pujo-Menjouet : Presse écrite (Le matin dimanche, Suisse), https://twitter.com/lpmenjouet/status/1130474006312890373/photo/1
- Laurent Pujo-Menjouet : Presse écrite (Le journal de Montréal), https://www.journaldemontreal.com/2019/07/30/et-si-lamour-se-maintenait-avec-des-equations-mathematiques

7.3.2. Interventions

- Thomas Lepoutre: Visite de l'inria pour les stagiaires de 3e de l'ICJ (plusieurs fois dans l'année, une matinée en général).
- Thomas Lepoutre participation to MathaLyon intervention
- Laurent Pujo-Menjouet participation to "Math en Jeans", Collège-Lycée Lacassagne, Lyon and Collège Gratte-Ciel, Villeurbanne
- Laurent Pujo-Menjouet participation to "Mathématiques étonnantes, SMF conférence", Lyon
- Laurent Pujo-Menjouet participation to TEDx, Montrouge https://www.youtube.com/watch?v=xJ8wE_wmtUY

8. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] R. DUCHESNE.*Erythroid differentiation in vitro under the lens of mathematical modelling*, Université de Lyon, December 2019, https://hal.archives-ouvertes.fr/tel-02440831

Articles in International Peer-Reviewed Journal

[2] M. ADIMY, A. CHEKROUN, T. KUNIYA.*Coupled reaction-diffusion and difference system of cell-cycle dynamics for hematopoiesis process with Dirichlet boundary conditions*, in "Australian Journal of Mathematical Analysis and Applications", November 2019, vol. 479, n° 1, p. 1030-1068 [*DOI :* 10.1016/J.JMAA.2019.06.066], https://hal.inria.fr/hal-02448738

[3] M. ADIMY, A. CHEKROUN, C. PIO FERREIRA.*Global dynamics of a differential-difference system: a case of Kermack-McKendrick SIR model with age-structured protection phase*, in "Mathematical Biosciences and Engineering", 2020, vol. 17, n° 2, p. 1329-1354 [*DOI :* 10.3934/MBE.2020067], https://hal.inria.fr/hal-02431660

[4] P. ARNER, S. BERNARD, L. APPELSVED, K.-Y. FU, D. ANDERSSON, M. SALEHPOUR, A. THORELL, M. RYDÉN, K. SPALDING.*Adipose lipid turnover and long-term changes in body weight*, in "Nature Medicine", September 2019, vol. 25, n° 9, p. 1385-1389 [*DOI :* 10.1038/S41591-019-0565-5], https://hal.archives-ouvertes.fr/hal-02317491

[5] M. A. BENCHAIB, A. BOUCHNITA, V. VOLPERT, A. MAKHOUTE.*Mathematical Modeling Reveals That the Administration of EGF Can Promote the Elimination of Lymph Node Metastases by PD-1/PD-L1 Blockade*, in "Frontiers in Bioengineering and Biotechnology", May 2019, vol. 7, p. 1-18 [*DOI :* 10.3389/FBIOE.2019.00104], https://hal.archives-ouvertes.fr/hal-02363404

[6] N. BESSONOV, A. BEUTER, S. TROFIMCHUK, V. VOLPERT.*Cortical waves and post-stroke brain stimulation*, in "Mathematical Methods in the Applied Sciences", June 2019, vol. 42, n^o^ 11, p. 3912-3928 [*DOI :* 10.1002/MMA.5620], https://hal.archives-ouvertes.fr/hal-02363412

[7] 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

[8] A. BONNAFFOUX, U. HERBACH, A. RICHARD, A. GUILLEMIN, S. GONIN-GIRAUD, P.-A. GROS, O. GANDRILLON.*WASABI: a dynamic iterative framework for gene regulatory network inference*, in "BMC Bioinformatics", April 2019, vol. 20, n^o^ 1, p. 1-19 [*DOI :* 10.1186/S12859-019-2798-1], https://hal.archives-ouvertes.fr/hal-02159694

[9] A. BOUCHNITA, V. VOLPERT.*A multiscale model of platelet-fibrin thrombus growth in the flow*, in "Computers & Fluids", April 2019, vol. 184, p. 10-20 [*DOI :* 10.1016/J.COMPFLUID.2019.03.021], https://hal.archives-ouvertes.fr/hal-02363406

[10] 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", 2019, vol. 24, n^o^ 6, p. 2417-2442 [*DOI :* 10.3934/DCDSB.2018259], https://hal.inria.fr/hal-01835435

[11] L. BOULLU, L. PUJO-MENJOUET, J. BÉLAIR.*Stability analysis of an equation with two delays and application to the production of platelets*, in "Discrete and Continuous Dynamical Systems - Series S", 2019, p. 1-24, forthcoming [*DOI :* 10.3934/DCDSS.2020131], https://hal.inria.fr/hal-02109546

[12] L. BOULLU, L. PUJO-MENJOUET, J. WU.*A model for megakaryopoiesis with state-dependent delay*, in "SIAM Journal on Applied Mathematics", July 2019, vol. 79, n^o^ 4, p. 1218–1243 [*DOI :* 10.1137/18M1201020], https://hal.inria.fr/hal-02123621

[13] V. CALVEZ, T. LEPOUTRE, N. MEUNIER, N. MULLER.*Non-linear analysis of a model for yeast cell communication*, in "ESAIM: Mathematical Modelling and Numerical Analysis", 2019, p. 1-32, forthcoming [*DOI :* 10.1051/M2AN/2019065], https://hal.inria.fr/hal-02372511

[14] I. S. CIUPERCA, M. DUMONT, A. LAKMECHE, P. MAZZOCCO, L. PUJO-MENJOUET, H. REZAEI, L. TINE.*Alzheimer's disease and prion: An in vitro mathematical model*, in "Discrete and Continuous Dynamical Systems - Series B", October 2019, vol. 24, n^o^ 10, p. 5225-5260 [*DOI :* 10.3934/DCDSB.2019057], https://hal.archives-ouvertes.fr/hal-02414635

[15] R. DUCHESNE, A. GUILLEMIN, F. CRAUSTE, O. GANDRILLON.*Calibration, Selection and Identifiability Analysis of a Mathematical Model of the in vitro Erythropoiesis in Normal and Perturbed Contexts*, in "In Silico Biology", May 2019, vol. 13, n^o^ 1-2, p. 55-69 [*DOI :* 10.3233/ISB-190471], https://hal.archives-ouvertes.fr/hal-02309925

[16] S. GIREL, C. ARPIN, J. MARVEL, O. GANDRILLON, F. CRAUSTE.*Model-based assessment of the Role of Uneven Partitioning of Molecular Content on Heterogeneity and Regulation of Differentiation in CD8 T-cell Immune Responses*, in "Frontiers in Immunology", 2019, p. 1-41, https://arxiv.org/abs/1903.01233 , forthcoming [*DOI :* 10.3389/FIMMU.2019.00230], https://hal.archives-ouvertes.fr/hal-02008705

[17] D. GREBENNIKOV, A. BOUCHNITA, V. VOLPERT, N. BESSONOV, A. MEYERHANS, G. BOCHAROV.*Spatial Lymphocyte Dynamics in Lymph Nodes Predicts the Cytotoxic T Cell Frequency Needed for HIV Infection Control*, in "Frontiers in Immunology", June 2019, vol. 10, p. 1-15 [*DOI :* 10.3389/FIMMU.2019.01213], https://hal.archives-ouvertes.fr/hal-02363399

[18] A. GUILLEMIN, R. DUCHESNE, F. CRAUSTE, S. GONIN-GIRAUD, O. GANDRILLON.*Drugs modulating stochastic gene expression affect the erythroid differentiation process*, in "PLoS ONE", November 2019, vol. 14, n° 11, p. 1-19 [*DOI :* 10.1371/JOURNAL.PONE.0225166], https://hal.archives-ouvertes.fr/hal-02375801

[19] U. HERBACH.*Stochastic Gene Expression with a Multistate Promoter: Breaking Down Exact Distributions*, in "SIAM Journal on Applied Mathematics", June 2019, vol. 79, n° 3, p. 1007-1029 [*DOI :* 10.1137/18M1181006], https://hal.archives-ouvertes.fr/hal-02159699

[20] T. LEPOUTRE.*Improved duality estimates: time discrete case and applications to a class of cross-diffusion systems*, in "Communications in Mathematical Sciences", 2019, vol. 17, n° 2, p. 1-17, https://arxiv.org/abs/1812.01296 , https://hal.inria.fr/hal-01943576

[21] M. MARION, V. VOLPERT.*Existence of Pulses for the System of Competition of Species*, in "Journal of Dynamics and Differential Equations", September 2019, vol. 31, n° 3, p. 1653-1670 [*DOI :* 10.1007/S10884-017-9582-6], https://hal.archives-ouvertes.fr/hal-02363423

[22] J. E. MOLD, P. RÉU, A. OLIN, S. BERNARD, J. MICHAELSSON, S. RANE, A. YATES, A. KHOSRAVI, M. SALEHPOUR, G. POSSNERT, P. BRODIN, J. FRISÉN.*Cell generation dynamics underlying naïve T cell homeostasis in adult humans*, in "PLoS Biology", October 2019, vol. 17, n° 10, e3000383 [*DOI :* 10.1101/635995], https://hal.archives-ouvertes.fr/hal-02160888

[23] N. RATTO, M. MARION, V. VOLPERT.*Existence of pulses for a reaction-diffusion system of blood coagulation in flow*, in "Mathematical Biosciences and Engineering", 2019, vol. 16, n° 5, p. 4196-4212, forthcoming [*DOI :* 10.3934/MBE.2019209], https://hal.archives-ouvertes.fr/hal-02084120

[24] N. RATTO, M. MARION, V. VOLPERT.*Existence of pulses for a reaction-diffusion system of blood coagulation*, in "Topological Methods in Nonlinear Analysis", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-02089011

[25] N. RATTO, M. MARION, V. VOLPERT.*Existence of Waves for a Bistable Reaction–Diffusion System with Delay*, in "Journal of Dynamics and Differential Equations", 2019, vol. 16, n° 5, p. 4196-4212 [*DOI :* 10.1007/S10884-019-09751-4], https://hal.archives-ouvertes.fr/hal-02363403

[26] A. RICHARD, E. VALLIN, C. ROMESTAING, D. ROUSSEL, O. GANDRILLON, S. GONIN-GIRAUD.*Erythroid differentiation displays a peak of energy consumption concomitant with glycolytic metabolism rearrangements*, in "PLoS ONE", September 2019, vol. 14, n° 9, e0221472 [*DOI :* 10.1371/JOURNAL.PONE.0221472], https://hal-univ-lyon1.archives-ouvertes.fr/hal-02318812

[27] 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

[28] S. TROFIMCHUK, V. VOLPERT.*Global continuation of monotone waves for bistable delayed equations with unimodal nonlinearities*, in "Nonlinearity", July 2019, vol. 32, n^o^ 7, p. 2593-2632 [*DOI :* 10.1088/1361-6544/AB0E23], https://hal.archives-ouvertes.fr/hal-02363393

[29] M. YEUNG, M. DJELLOUL, E. STEINER, S. BERNARD, M. SALEHPOUR, G. POSSNERT, L. BRUNDIN, J. FRISÉN.*Dynamics of oligodendrocyte generation in multiple sclerosis*, in "Nature", January 2019, n^o^ 566, p. 538–542 [*DOI :* 10.1038/S41586-018-0842-3], https://hal.archives-ouvertes.fr/hal-02002274

Conferences without Proceedings

[30] M. ANDRADE-RESTREPO, P. LEMARRE, L. PUJO-MENJOUET, L. TINE, S. I. CIUPERCA.*Modeling the spatial propagation of Aβ oligomers in alzheimer's disease*, in "CEMRACS 2018 - Numerical and mathematical modeling for biological and medical applications: deterministic, probabilistic and statistical descriptions", Marseille, France, 2019, p. 1 - 10, https://hal.archives-ouvertes.fr/hal-02414660

Scientific Books (or Scientific Book chapters)

[31] M. CALÉRO, I. S. CIUPERCA, L. PUJO-MENJOUET, L. MATAR TINE.*Alzheimer Disease: Convergence Result from a Discrete Model Towardsa Continuous One*, in "Trends in Biomathematics: Mathematical Modeling for Health, Harvesting, and Population Dynamics", Springer, 2019, p. 397-431 [*DOI :* 10.1007/978-3-030-23433-1_25], https://hal.archives-ouvertes.fr/hal-02414726

[32] K. GRZESIK, K. ENG, F. CRAUSTE, S. BATTAGLIA.*In Silico Methods for Studying T Cell Biology*, in "Int. Rev. Cell. Mol. Biol. 342", Elsevier, 2019, p. 265-304 [*DOI :* 10.1016/BS.IRCMB.2018.07.005], https://hal.archives-ouvertes.fr/hal-02237891

[33] L. PUJO-MENJOUET.*Le jeu de l'amour sans le hasard Mathématiques du couple*, Des Equateurs, May 2019, p. 1-248, https://hal.archives-ouvertes.fr/hal-02414777

[34] L. PUJO-MENJOUET.*Passer notre amour à la machine*, in "Futures of Love", August 2019, https://hal.archives-ouvertes.fr/hal-02414758

Scientific Popularization

[35] L. PUJO-MENJOUET.*Est-ce par hasard s'il a le cœur qui s'égare*, in "Pour la science", November 2019, https://hal.archives-ouvertes.fr/hal-02414752

Other Publications

[36] C. ROCABERT, G. BESLON, C. KNIBBE, S. BERNARD.*Phenotypic Noise and the Cost of Complexity*, November 2019, EvoLyon 2019, Poster, https://hal.archives-ouvertes.fr/hal-02402443

References in notes

[37] 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

[38] 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

[39] 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

[40] 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

[41] 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

Project-Team ELAN

modELing the Appearance of Nonlinear phenomena

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Numerical schemes and simulations

Table of contents

Project-Team ELAN

Creation of the Project-Team: 2019 August 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
A6.1.5. - Multiphysics modeling
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.5. - Numerical Linear Algebra
A6.2.6. - Optimization
A6.2.7. - High performance computing
A6.2.8. - Computational geometry and meshes
A6.3.1. - Inverse problems
A6.5. - Mathematical modeling for physical sciences
A6.5.1. - Solid mechanics
A6.5.2. - Fluid mechanics
A6.5.3. - Transport
A9.2. - Machine learning

Other Research Topics and Application Domains:

B1.1.2. - Molecular and cellular biology
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 Scientists

Florence Descoubes [Team leader, Inria, Researcher, HDR]
Thibaut Metivet [Inria, Researcher, from Sep 2019]

Technical Staff

Laurence Boissieux [Inria, Engineer]
Victor Romero Gramegna [Inria, Engineer]

PhD Students

Raphael Charrondiere [Univ Grenoble Alpes, PhD Student]
Mickael Ly [Inria, PhD Student]
Abdullah Haroon Rasheed [Inria, PhD Student]

Post-Doctoral Fellow

Gauthier Rousseau [Inria, Post-Doctoral Fellow, from Apr 2019]

Visiting Scientist

Julien Duron [École normale supérieure de Rennes, from Jun 2019 until Jul 2019]

2. Overall Objectives

2.1. Overall Objectives

ELAN is a young research 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 buckling of an elastic ribbon, the flowing of sand, 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 and frictional contact, 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 [10], [12], [15]. 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 [37]. We note that adopting the multibody system dynamics point of view helped us formulate a linear-time integration scheme [11], 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 [25], [34]. 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., [29], [16]). However, no real consensus has been obtained so far about a unified nonlinear shell theory able to support large displacements. In [13] 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 [14], 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 [26] are not able to deal with non-persistent or sliding contact.

3.2. Discrete and continuous 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 [17], [33]. 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 [9]. 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., [36], [28]). 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 [18]. This justifies solving the reduced discrete frictional contact problem where primary unknowns are forces, as usually advocated in nonsmooth mechanics [31]. 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 [32]. 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 granular and fibrous media

Though we have recently made progress on the continuum formulation and solving of granular materials in Gilles Daviet's PhD thesis [22], [20], [19], 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 [21]. 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 [24], [23]. 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 [13].

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$^{\text{th}}$ 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. Highlights of the Year

[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 [27].

[0] It is however amusing to observe that research in these areas is quite successful in obtaining the IG Nobel prize [8], [30], thus still being considered as an exotic research topic by physicists.

5.1. Highlights of the Year

5.1.1. Creation of Graphyz, a new graphics-physics workshop

- F. Bertails-Descoubes, together with B. Audoly (École Polytechnique), has founded, chaired and organized the first graphics-physics workshop, Graphyz, held at Inria Montbonnot on October 24-25 2019. An outstanding scientific program, gathering 15 international experts from both Computer Graphics and Physics, originally combined talks from both communities around various topics ranging from viscous thread coiling to snow avalanches. The workshop was entirely funded by the ERC GEM. Being a high success, it will be organized again in 2021, in Paris.

5.1.2. Keynote at Eurographics 2019

- Florence Bertails-Descoubes was a Keynote speaker at Eurographics 2019 held in May 2019 in Genova, Italy.

5.1.3. Awards

BEST PAPERS AWARDS :

[7]

R. CHARRONDIÈRE, F. BERTAILS-DESCOUBES, S. NEUKIRCH, V. ROMERO.*Modélisation numérique de rubans en éléments de haut degré*, in "JF.IG.RV 2019 - Journées Françaises d'Informatique Graphique et de Réalité Virtuelle", Marseille, France, November 2019, p. 1-7, https://hal.archives-ouvertes.fr/hal-02384170

6. New Software and Platforms

6.1. Argus-distribution

KEYWORDS: Frictional contact - Cloth dynamics - Mesh adaptation

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). The Argus-distribution code was awarded in 2019 the [Graphics Replicability stamp](http://www.replicabilitystamp.org/), which acknowledges its reproducibility.

- Participants: Jie Li, Gilles Daviet, Rahul Narain, Florence Bertails, Matthew Overby, George Brown and Laurence Boissieux
- Partners: Department of Computer Science and Engineering, University of Minnesota - IIT Delhi
- Contact: Florence Bertails
- Publication: An Implicit Frictional Contact Solver for Adaptive Cloth Simulation
- URL: http://www-users.cselabs.umn.edu/~lixx4611/contact_friction.html

6.2. Feel++

KEYWORDS: High order finite elements - Discontinuous Galerkin - High-Performance Computing

FUNCTIONAL DESCRIPTION: Feel++ is a high-performance C++ library for the resolution of general variational formulations, including continuous and discontinuous Galerkin methods, finite element or spectral element methods, reduced basis formulations, etc. It features a high-level domain specific embedded language (DSEL) for Galerkin methods, space dimension-agnostic computation kernels and seamless and automatic parallelism. It also includes applicative toolboxes to solve physics problems in fluid mechanics, solid mechanics, thermal conduction, and the corresponding multi-physics coupling.

- Partners: Université de Strasbourg - UGA - Inria
- Contact: Thibaut Metivet
- URL: http://www.feelpp.org

7. New Results

7.1. Static simulation of thin elastic ribbons

Participants: Raphaël Charrondière, Florence Bertails-Descoubes, Victor Romero.

In collaboration with Sébastien Neukirch (Sorbonne Université, Institut Jean le Rond d'Alembert), we have proposed a robust and efficient numerical model to compute stable equilibrium configurations of thin elastic ribbons featuring arbitrarily curved natural shapes. Our spatial discretization scheme relies on elements characterized by a linear normal curvature and a quadratic geodesic torsion with respect to arc length. Such a high-order discretization allows for a great diversity of kinematic representations, while guaranteeing the ribbon to remain perfectly inextensible. Stable equilibria are calculated by minimizing gravitational and elastic energies of the ribbon, under a developability constraint. This work is currently under review in a journal of Mechanics. Some preliminary results have already been communicated about in two French congresses, one in Mechanics [6] and one in Computer Graphics [7] (best paper award).

7.2. Video-based measurement of the friction coefficient between cloth and a substrate

Participants: Haroon Rasheed, Victor Romero, Florence Bertails-Descoubes.

In collaboration with Arnaud Lazarus (Sorbonne Université, Institut Jean le Rond d'Alembert), Jean-Sébastien Franco and Stefanie Wuhrer (Inria, Morphéo team), we have investigated a first non-invasive measurement network for estimating cloth friction at contact with a substrate. Our network was trained on data exclusively generated by the solver ARGUS co-developed by the ELAN team, which we have carefully validated against real experiments under controlled conditions. We have shown promising friction measurement results on multiple real cloth samples contacting various kinds of substrates, by comparing our estimations based on a simple video acquisition protocol against standard measurements. This work has been submitted in late 2019 for publication in a Computer Vision conference, and some preliminary results have been communicated about in a mechanical congress [4].

7.3. Willmore flow simulation with diffusion-redistanciation numerical schemes

Participant: Thibaut Metivet.

In collaboration with Arnaud Sengers (Université Claude Bernard), Emmanuel Maitre (Laboratoire Jean Kuntzmann, Grenoble INP) and Mourad Ismail (Laboratoire Interdisciplinaire de Physique, UGA), we have proposed original diffusion-redistanciation numerical schemes to compute the static shapes of elastic membranes with bending stiffness under constant area constraints. This numerical method relies on an implicit representation of the surface which is used as an initial condition for diffusion-like equations. This allows to circumvent the usual difficulties pertaining to the high geometrical order and non-linearities of the bending energy and to benefit from the robustness of discretised diffusion operators. We have implemented the schemes within the finite element library Feel++ and studied the numerical convergence properties in 2D and 3D. We have also validated our method using comparative benchmarks computed with standard approaches. This work has led to the PhD defense of Arnaud Sengers [35] and a publication is under preparation.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. National Collaborations

- Long-term collaboration with Christophe Prud'homme and Vincent Chabannes (Université de Strasbourg and Centre de modélisation et de simulation de Strasbourg).

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. 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.

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. 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).

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. General Chair, Scientific Chair

- Florence Bertails-Descoubes was co-founder and co-chair (together with Basile Audoly, École Polytechnique) of the new graphics-physics worskhop Graphyz, held at Inria in Montbonnot on October 24-25 2019. The ELAN team was the local organizer of the event.

9.1.1.2. Member of the Organizing Committees

- Florence Bertails-Descoubes, together with the help of Inria and of the ELAN team, has organized the new graphics-physics workshop Graphyz at Inria in Montbonnot. See above.

9.1.2. Scientific Events: Selection

9.1.2.1. Member of the Conference Program Committees

- Florence Bertails-Descoubes was member of the ACM SIGGRAPH Technical Program Committee in 2019, and of the Eurographics Technical Program Committee in 2019.

9.1.3. Journal

9.1.3.1. Reviewer - Reviewing Activities

- Florence Bertails-Descoubes was reviewer in 2019 for ACM Transaction on Graphics, ACM SIGGRAPH 2019, ACM SIggraph Asia 2019, ACM-EG Symposium on Computer Animation, UIST 2019, Nonlinear Dynamics, Computer Graphics Forum, SIAM Journal on Scientific Computing, Elsevier Computer Methods in Applied Mechanics and Engineering.
- Thibaut Metivet was reviewer in 2019 for Elsevier Computer Methods in Applied Mechanics and Engineering and Coupled Systems Mechanics.

9.1.4. Invited Talks

- Florence Bertails-Descoubes, Keynote speaker at Eurographics 2019, Genova, Italy, May 2019.
- Florence Bertails-Descoubes, invited talk at Matherials, a series of cross-disciplinary seminars jointly organized by Laboratoire Jean Kuntzmann, SIMAP and Institut Fourier, June 2019.
- Florence Bertails-Descoubes, Exposé invité aux Journées Françaises d'Informatique Graphique, Marseille, November 2019.
- Florence Bertails-Descoubes, invited talk at the HCERES evaluation of Laboratoire Jean Kuntzmann, Inria Montbonnot, December 2019.
- Thibaut Metivet, invited poster at RheoSuNN, a workshop on the numerical simulation of suspensions organised at Ecole Polytechnique in March 2019.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence : Raphaël Charrondière, TP Projet Logiciel, 13h éq TD, L2 STG Grenoble, Université Grenoble Alpes.

Licence : Florence Bertails-Descoubes, Méthodes Numériques, 18h éq TD, L3, ENSIMAG 1A, Grenoble INP.

License : Thibaut Metivet, Analyse, 33h éq TD, L3, ENSIMAG 1A, Grenoble INP.

Master : Raphaël Charrondière, Complexité Algorithmique Des Problèmes, 15h éq TD, M1, Université Grenoble Alpes.

Master : Florence Bertails-Descoubes, Special Course for M2 at École Normale Supérieure de Lyon, entitled “Numerical Mechanics: From Lagrangian mechanics to simulation tools for computer graphics“, 19h éq TD.

Master : Mickaël Ly, Special Course for M2 at École Normale Supérieure de Lyon, entitled “Numerical Mechanics: From Lagrangian mechanics to simulation tools for computer graphics“, 14h éq TD.

9.2.2. 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.

PhD in progress : François Der Hovsepian, Modélisation et simulation d’écoulement de cellules tumorales dans le sang et de l’adhésion aux parois, 19 October 2017, Christophe Prud’homme, Vincent Chabannes (Université de Strasbourg) and Thibaut Metivet.

9.2.3. Juries

Florence Bertails-Descoubes, member (Examinatrice) of Ph.D. Thesis committee of A. Sibellas (8 March 2019), INSA Lyon (directeur de thèse : E. Maire)

Florence Bertails-Descoubes, member (Présidente du jury) of Ph.D. Thesis committee of V. Leroy (17 October 2019), Inria Rhône-Alpes Montbonnot (directeur de thèse : E. Boyer, co-encadrant : J.-S. Franco).

Thibaut Metivet, member (Encadrant, invité) of Ph.D. Thesis committee of A. Sengers (19 July 2019), Université Grenoble-Alpes (directeur de thèse : E. Maitre).

9.3. Popularization

9.3.1. Articles and contents

- Focus on our cloth simulator Argus, on a video prepared by A. Aftalion and published in February 2019 in http://video.math.cnrs.fr.

10. Bibliography

Major publications by the team in recent years

[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^{o} 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^{o} 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^{o} 6, p. 1-16 [*DOI :* 10.1145/3272127.3275036], https://hal.inria.fr/hal-01883655

Publications of the year

International Conferences with Proceedings

[4] 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 "APS 2019 - American Physical Society March Meeting", Boston, United States, March 2019, 1, https://hal.inria.fr/hal-01982257

[5] V. Romero, F. Bertails-Descoubes, A. Derouet-Jourdan, A. Lazarus.*Inverse design of a suspended Kirchhoff rod: From theory to practice*, in "APS 2019 - American Physical Society March Meeting", Boston, United States, March 2019, https://hal.inria.fr/hal-01981923

Conferences without Proceedings

[6] R. Charrondière, F. Bertails-Descoubes, S. Neukirch.*Modélisation numérique de rubans en éléments de haut degré*, in "14ème Colloque National en Calcul des Structures (CSMA 2019)", Giens, France, May 2019, p. 1-8, https://hal.archives-ouvertes.fr/hal-02384085

[7] *Best Paper*
R. Charrondière, F. Bertails-Descoubes, S. Neukirch, V. Romero.*Modélisation numérique de rubans en éléments de haut degré*, in "JF.IG.RV 2019 - Journées Françaises d'Informatique Graphique et de Réalité Virtuelle", Marseille, France, November 2019, p. 1-7, https://hal.archives-ouvertes.fr/hal-02384170.

References in notes

[8] 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

[9] 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

[10] 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

[11] 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/

[12] 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

[13] 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

[14] A. BLUMENTALS.*Numerical modelling of thin elastic solids in contact*, Université de Grenoble Alpes, July 2017

[15] 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

[16] 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]

[17] 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

[18] 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

[19] 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

[20] 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

[21] 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

[22] 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

[23] 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

[24] 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

[25] 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

[26] 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

[27] 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

[28] 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]

[29] A. L. GOL'DENVEIZER.*Theory of Elastic Thin Shells*, Pergamon Press, 1961

[30] 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

[31] 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

[32] 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

[33] M. MOORE, J. WILHELMS.*Collision detection and response for computer animation*, in "Computer Graphics Proceedings (Proc. ACM SIGGRAPH'88)", 1988, p. 289–298

[34] 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

[35] A. SENGERS.*Semi-implicit and diffusion-redistanciation schemes for the dynamic of red blood cells*, Université Grenoble Alpes, July 2019, https://tel.archives-ouvertes.fr/tel-02341602

[36] 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

[37] 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

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

Solon Pissis [CWI, The Netherlands, Senior Researcher, since Mar 2019]
Marie-France Sagot [Inria, Team leader, Senior Researcher, HDR]
Alexander Schönhuth [CWI, The Netherlands, Senior Researcher, also since Oct 2017 part-time Professor at Univ of Utrecht]
Blerina Sinaimeri [Inria, Researcher]
Fabrice Vavre [CNRS, Researcher, HDR]
Alain Viari [Inria, Senior Researcher]

Faculty Members

Hubert Charles [INSA Lyon, Full Professor, HDR]
Roberto Grossi [Univ Pisa, Italy, Full Professor]
Giuseppe Francesco Italiano [LUISS Univ, Rome, Italy, Professor]
Vincent Lacroix [Univ de Claude Bernard, Associate Professor]
Alberto Marchetti Spaccamela [Sapienza Univ Rome, Italy, Full Professor]
Arnaud Mary [Univ de Claude Bernard, Associate Professor]
Nadia Pisanti [Univ Pisa, Italy, Associate 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

Eric Cumunel [Univ Claude Bernard Lyon, from Nov 2019]

PhD Students

Marianne Borderes [MaatPharma and Claude Bernard Lyon]
Audric Cologne [Inria, until Sep 2019]
Nicolas Homberg [Inra and Inria, from Nov 2019]
Leandro Ishi Soares de Lima [CNPq Brazil & Univ Claude Bernard Lyon, until Apr 2019]
Carol Moraga Quinteros [Conicyt & Univ Claude Bernard Lyon]
Henri Pusa [Inria, until Feb 2019]
Camille Sessegolo [Univ Claude Bernard Lyon]
Yishu Wang [Univ Claude Bernard Lyon]
Irene Ziska [Inria]

Post-Doctoral Fellows

Ricardo de Andrade Abrantes [Univ. of São Paulo, Brazil, since Nov 2019]
Audric Cologne [CNRS, since Nov 2019]
Alex Di Genova [Inria, until Nov 2019]
Mariana Galvão Ferrarini [Insa Lyon & Univ Claude Bernard Lyon]
Scheila Mucha [Inria, from Nov 2019]
Henri Pusa [Inria, until May 2019]

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 in 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 (co)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 main health areas currently addressed that are intrinsically inter-related.

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 on 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

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 (https://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.

- Participants: Alex Di Genova, Marie-France Sagot, Alejandro Maass and Gonzalo Ruz Heredia
- 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 P. 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 : The 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+.

KISSPLICE was applied to a new application field, virology, through a collaboration with the group of Nadia Naffakh at Institut Pasteur. The goal is to understand how a virus (in this case influenza) manipulates the splicing of its host. This led to new developments in KISSPLICE. Taking into account the strandedness of the reads was required, in order not to mis-interpret transcriptional readthrough. We now use BCALM instead of DBG-V4 for the de Bruijn graph construction and this led to major improvements in memory and time requirements of the pipeline. We still cannot scale to very large datasets like in cancer, the time limiting step being the quantification of bubbles.

- 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

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.

- Participants: Ricardo Luiz de Andrade Abrantes, Nuno Mira, Susana Vinga and Marie-France Sagot
- Contact: Marie-France Sagot
- URL: http://momo-sysbio.gforge.inria.fr

5.19. Moomin

Mathematical explOration of Omics data on a MetabolIc Network

KEYWORDS: Metabolic networks - Transcriptomics

FUNCTIONAL DESCRIPTION:MOOMIN is a tool for analysing differential expression data. It takes as its input a metabolic network and the results of a DE analysis: a posterior probability of differential expression and a (logarithm of a) fold change for a list of genes. It then forms a hypothesis of a metabolic shift, determining for each reaction its status as "increased flux", "decreased flux", or "no change". These are expressed as colours: red for an increase, blue for a decrease, and grey for no change. See the paper for full details: https://doi.org/10.1093/bioinformatics/btz584

- Participants: Henri Taneli Pusa, Mariana Ferrarini, Ricardo Luiz de Andrade Abrantes, Arnaud Mary, Alberto Marchetti-Spaccamela, Leen Stougie and Marie-France Sagot
- Contact: Marie-France Sagot
- URL: https://github.com/htpusa/moomin

5.20. MultiPus

KEYWORDS: Systems Biology - Algorithm - Graph algorithmics - Metabolic networks - Computational biology

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/

5.21. 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.

- Participants: Vicente Acuna Aguayo, Paulo Vieira Milreu, Alberto Marchetti-Spaccamela, Leen Stougie, Martin Wannagat and Marie-France Sagot
- Contact: Marie-France Sagot
- URL: http://gforge.inria.fr/projects/pitufo/

5.22. Sasita

KEYWORDS: Bioinformatics - Graph algorithmics - Systems Biology

FUNCTIONAL DESCRIPTION:SASITA is a software for the exhaustive enumeration of minimal precursor sets in metabolic networks.

- Participants: Vicente Acuna Aguayo, Ricardo Luiz de Andrade Abrantes, Paulo Vieira Milreu, Alberto Marchetti-Spaccamela, Leen Stougie, Martin Wannagat and Marie-France Sagot
- Contact: Marie-France Sagot
- URL: http://sasita.gforge.inria.fr/

5.23. 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

5.24. 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

5.25. Rime

KEYWORDS: Bioinformatics - Genomics - Sequence alignment

FUNCTIONAL DESCRIPTION: Detects long similar fragments occurring at least twice in a set of biological sequences.

- Participants: Nadia Pisanti and Marie-France Sagot
- Contact: Nadia Pisanti

5.26. 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.

- Participants: Alice Julien-Laferrière, Ricardo Luiz de Andrade Abrantes, Arnaud Mary, Mariana Ferrarini, Susana Vinga, Irene Ziska and Marie-France Sagot
- Contact: Marie-France Sagot
- URL: http://hyperstories.gforge.inria.fr/

5.27. VG-Flow

Viral haplotype reconstruction from contigs using variation graphs

KEYWORD: Haplotyping

FUNCTIONAL DESCRIPTION: The goal of haplotype-aware genome assembly is to reconstruct all individual haplotypes from a mixed sample and to provide corresponding abundance estimates. VG-FLOW provides a reference-genome-independent solution based on the construction of a variation graph, capturing all quasispecies diversity present in the sample. We solve the contig abundance estimation problem and propose a greedy algorithm to efficiently build full-length haplotypes. Finally, we obtain accurate frequency estimates for the reconstructed haplotypes through linear programming techniques.

- Contact: Alexander Schonhuth
- URL: https://bitbucket.org/jbaaijens/vg-flow

5.28. Virus-VG

Viral haplotype reconstruction from contigs using variation graphs

KEYWORD: Haplotyping

FUNCTIONAL DESCRIPTION: 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. VIRUS-VG aims to reconstruct full-length haplotypes together with their abundances from such contigs, represented as a genome variation graph.

- Contact: Alexander Schonhuth
- URL: https://bitbucket.org/jbaaijens/virus-vg

5.29. Wengan

Making the path

KEYWORD: Genome assembly

FUNCTIONAL DESCRIPTION:WENGAN is a new genome assembler that unlike most of the current long-reads assemblers avoids entirely the all-vs-all read comparison. The key idea behind WENGAN is that long-read alignments can be inferred by building paths on a sequence graph. To achieve this, WENGAN builds a new sequence graph called the Synthetic Scaffolding Graph. The SSG is built from a spectrum of synthetic mate-pair libraries extracted from raw long-reads. Longer alignments are then built by performing a transitive reduction of the edges. Another distinct feature of WENGAN is that it performs self-validation by following the read information. WENGAN identifies miss-assemblies at differents steps of the assembly process.

- Participants: Alex Di Genova and Marie-France Sagot
- Contact: Marie-France Sagot
- URL: https://github.com/adigenova/wengan

5.30. 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 2019.

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 chose not to detail here the results on more theoretical aspects of computer science when these are initially addressed in contexts not directly related to computational biology even though those on string [11], [36], [40], [41], [23], [45] and graph algorithms in general [35], [39], [38], [17], [43] are relevant for life sciences, such as for instance pan-genome analysis, or could become more specifically so in a near future. One important example of the latter concerns enumeration algorithms that has always been at the heart of the computer science and mathematics interests of the team. In such context, the so-called reconfiguration problem which asks whether one solution can be transformed into the other in a step-by-step fashion such that each intermediate solution is also feasible is of particular relevance. This was explored in the context of a perfect matching problem [37].

A few other results of 2019 are not mentioned in this report, not because the corresponding work is not important, but because it was likewise more specialised [8], [9], [12], [44]. In the same way, also for space reasons, we chose not to detail the results presented in some biological papers of the team when these did not require a mathematical or algorithmic input [16], [22].

On the other hand, we do mention a couple of works that were in preparation or about to be submitted towards the end of 2018.

6.2. Axis 1: Genomics

Transcriptome profiling using Nanopore sequencing Our vision of DNA transcription and splicing has changed dramatically with the introduction of short-read sequencing. These high-throughput sequencing technologies promised to unravel the complexity of any transcriptome. Generally gene expression levels are well-captured using these technologies, but there are still remaining caveats due to the limited read length and the fact that RNA molecules had to be reverse transcribed before sequencing. Oxford Nanopore Technologies has recently launched a portable sequencer which offers the possibility of sequencing long reads and most importantly RNA molecules. In [28], we generated a full mouse transcriptome from brain and liver using such Oxford Nanopore device. As a comparison, we sequenced RNA (RNA-Seq) and cDNA (cDNA-Seq) molecules using both long and short reads technologies and tested the TeloPrime preparation kit, dedicated to the enrichment of full-length transcripts. Using spike-in data, we confirmed in [28] that expression levels are efficiently captured by cDNA-Seq using short reads. More importantly, Oxford Nanopore RNA-Seq tends to be more efficient, while cDNA-Seq appears to be more biased. We further showed that the cDNA library preparation of the Nanopore protocol induces read truncation for transcripts containing internal runs of T's. This bias is marked for runs of at least 15 T's, but is already detectable for runs of at least 9 T's and therefore concerns more than 20% of the expressed transcripts in mouse brain and liver. Finally, we outlined that bioinformatic challenges remain ahead for quantifying at the transcript level, especially when reads are not full-length. Accurate quantification of repeat-associated genes such as processed pseudogenes also remains difficult, and we show in the paper that current mapping protocols which map reads to the genome largely over-estimate their expression, at the expense of their parent gene.

Genotyping and variant detection The amount of genetic variation discovered and characterised in human populations is huge, and is growing rapidly with the widespread availability of modern sequencing technologies. Such a great deal of variation data, that accounts for human diversity, leads to various challenging computational tasks, including variant calling and genotyping of newly sequenced individuals. The standard pipelines for addressing these problems include read mapping, which is a computationally expensive procedure. A few mapping-free tools were proposed in recent years to speed up the genotyping process. While such tools have highly efficient run-times, they focus on isolated, bi-allelic SNPs, providing limited support for multi-allelic SNPs, indels, and genomic regions with high variant density. To address these issues, we introduced MALVA, a fast and lightweight mapping-free method to genotype an individual directly from a sample of reads [10]. MALVA is the first mapping-free tool that is able to genotype multi-allelic SNPs and indels, even in high density genomic regions, and to effectively handle a huge number of variants such as those provided by the 1000 Genome Project. An experimental evaluation on whole-genome data shows that MALVA requires one order of magnitude less time to genotype a donor than alignment-based pipelines, providing similar accuracy. Remarkably, on indels, MALVA provides even better results than the most widely adopted variant discovery tools.

Still on the issue of SNP detection, in [25], we developed the positional clustering theory that (i) describes how the extended Burrows–Wheeler Transform (eBWT) of a collection of reads tends to cluster together bases that cover the same genome position, (ii) predicts the size of such clusters, and (iii) exhibits an elegant and precise LCP array based procedure to locate such clusters in the eBWT. Based on this theory, we designed and implemented an alignment-free and reference-free SNP calling method, and we devised a SNP calling pipeline. Experiments on both synthetic and real data show that SNPs can be detected with a simple scan of the eBWT and LCP arrays as, in agreement with our theoretical framework, they are within clusters in the eBWT of the reads. Finally, our tool intrinsically performs a reference-free evaluation of its accuracy by returning the coverage of each SNP. Based on the results of the experiments on synthetic and real data, we conclude that the positional clustering framework can be effectively used for the problem of identifying SNPs, and it appears to be a promising approach for calling other types of variants directly on raw sequencing data.

Finally, variant detection and various related algorithmic problems were extensively explored in the PhD of Leandro I. S. de Lima [2] defended in April 2019.

Bubble generator Bubbles are pairs of internally vertex-disjoint (s, t)-paths in a directed graph, which have many applications in the processing of DNA and RNA data such as variant calling as presented above. Listing and analysing all bubbles in a given graph is usually unfeasible in practice, due to the exponential number of bubbles present in real data graphs. In [4], we proposed a notion of bubble generator set, *i.e.*, a polynomial-sized subset of bubbles from which all the other bubbles can be obtained through a suitable application of a specific symmetric difference operator. This set provides a compact representation of the bubble space of a graph. A bubble generator can be useful in practice, since some pertinent information about all the bubbles can be more conveniently extracted from this compact set. We provided a polynomial-time algorithm to decompose any bubble of a graph into the bubbles of such a generator in a tree-like fashion. Finally, we presented two applications of the bubble generator on a real RNA-seq dataset.

Genome assembly The continuous improvement of long-read sequencing technologies along with the development of ad-doc algorithms has launched a new *de novo* assembly era that promises high-quality genomes. However, it has proven difficult to use only long reads to generate accurate genome assemblies of large, repeat-rich human genomes. To date, most of the human genomes assembled from long error-prone reads add accurate short reads to further improve the consensus quality (polishing). In a paper to be submitted before the end of 2019 (with as main authors A. di Genova and M.-F. Sagot), we report the development of an algorithm for hybrid assembly, WENGAN, and its application to hybrid sequence datasets from four human samples. WENGAN implements efficient algorithms that exploit the sequence information of short and long reads to tackle assembly contiguity as well as consensus quality. We show that the resulting genome assemblies have high contiguity (contig NG50:16.67-62.06 Mb), few assembly errors (contig NGA50:10.9-45.91 Mb), good consensus quality (QV:27.79-33.61), high gene completeness (BUSCO complete: 94.6-95.1%), and consume few computational resources (CPU hours:153-1027). In particular, the WENGAN assembly of the haploid CHM13 sample achieved a contig NG50 of 62.06 Mb (NGA50:45.91 Mb), which surpasses the contiguity of the current human reference genome (GRCh38 contig NG50:57.88 Mb). Because of its lower cost, WENGAN is an important step towards the democratisation of the *de novo* assembly of human genomes. WENGAN is available at https://github.com/adigenova/wengan.

On assembly still, although haplotype-aware genome assembly plays an important role in genetics, medicine and various other disciplines, the generation of haplotype-resolved *de novo* assemblies remains a major challenge. Beyond distinguishing between errors and true sequential variants, one needs to assign the true variants to the different genome copies. Recent work has pointed out that the enormous quantities of traditional NGS read data have been greatly underexploited in terms of haplotig computation so far, which reflects the fact that the methodology for reference independent haplotig computation has not yet reached maturity. We presented in [7] a new approach, called POLYploid genome fitTEr (POLYTE) for a *de novo* generation of haplotigs for diploid and polyploid genomes of known ploidy. Our method follows an iterative scheme where in each iteration reads or contigs are joined, based on their interplay in terms of an underlying haplotype-aware overlap graph. Along the iterations, contigs grow while preserving their haplotype identity. Benchmarking experiments on both real and simulated data demonstrate that POLYTE establishes new standards in terms of error-free reconstruction of haplotype-specific sequences. As a consequence, POLYTE outperforms state-of-the-art approaches in various relevant aspects, notably in polyploid settings.

Others Besides the above, we have also explored a proteogenomics workflow for the expert annotation of eukaryotic genomes [18], as well as a technology- and species-independent simulator of sequencing data and genomic variants [42].

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, who was until the end of 2018 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 Analysis of differential expression of genes is often performed to understand how the metabolic activity of an organism is impacted by a perturbation. However, because the system of metabolic regulation is complex and all changes are not directly reflected in the expression levels, interpreting these data can be difficult. In [26], we presented a new algorithm and computational tool that uses a genome-scale metabolic reconstruction to infer metabolic changes from differential expression data. Using the framework of constraint-based analysis, our method produces a qualitative hypothesis of a change in metabolic activity. In other words, each reaction of the network is inferred to have increased, decreased, or remained unchanged in flux. In contrast to similar previous approaches, our method does not require a biological objective function and does not assign on/off activity states to genes. An implementation is provided and is available online at the address https://github.com/htpusa/moomin. We applied the method to three published datasets to show that it successfully accomplishes its two main goals: confirming or rejecting metabolic changes suggested by differentially expressed genes based on how well they fit in as parts of a coordinated metabolic change, as well as inferring changes in reactions whose genes did not undergo differential expression. The above work was also part of the PhD of Taneli Pusa [3] defended in February 2019.

Metabolic games Game theory is a branch of applied mathematics originally developed to describe and reason about situations where two or more rational agents, the "homo economicus", are faced with choices and have potentially conflicting goals. All participants want to maximise their own well-being, but are doing so taking into account that everyone else is doing the same. Thus paradoxical, suboptimal, outcomes are possible and even common. Evolutionary game theory was born out of the realisation that rational choice can be replaced by natural selection: in the course of evolution the strategy (phenotype) that would "win" the game would prevail by simply proliferating more successfully thanks to its success in the "game". It turns out that phenotype prediction in the context of metabolic networks is exactly the type of problem that evolutionary game theory was meant to answer: given a set of choices (as defined by a metabolic network reconstruction), what will be the actual metabolism observed? In other words, if we culture a set of organisms together in a given medium, which are the phenotype(s) that emerge as winners? In [27], we sought to provide a short introduction to both evolutionary game theory and its use in the context of metabolic modelling. This work was also part of the PhD of Taneli Pusa [3].

6.4. Axis 3: (Co)Evolution

Modelling invasion Nowadays, the most used model in studies of the coevolution of hosts and symbionts is phylogenetic tree reconciliation. A crucial issue in this model is that from a biological point of view, reasonable cost values for an event-based reconciliation are not easily chosen. Different methods have been developed to infer such cost values for a given pair of host and symbiont trees, including one we established in the past. However, a major limitation of these methods is their inability to model the "invasion" of different host species by a same symbiont species (referred to as a spread event), which is often observed in symbiotic relations. Indeed, many symbionts are generalist. For instance, the same species of insect may pollinate different species of plants. In a paper currently in preparation, we propose a method, called AmoCoala, which for a given pair of host and symbiont trees, estimates the frequency of the cophylogenetic events, in presence of spread events, based on an approximate Bayesian computation (ABC) approach that may be more efficient than a classical likelihood method. The algorithm that we propose on one hand provides more confidence in the set of costs to be used for a given pair of host and symbiont trees, while on the other hand, it allows to estimate the frequency of the events even in the case of large datasets. We evaluated our method on both synthetic and real datasets.

Co-divergence and tree topology In reconstructing the common evolutionary history of hosts and symbionts, the current method of choice is the phylogenetic tree reconciliation. In this model, we are given a host tree H, a symbiont tree S, and a function σ mapping the leaves of S to the leaves of H and the goal is to find, under some biologically motivated constraints, a reconciliation, that is a function from the vertices of S to the vertices of H that respects σ and allows the identification of biological events such as co-speciation, duplication and host switch. The maximum co-divergence problem consists in finding the maximum number of co-speciations in a reconciliation. This problem is NP-hard for arbitrary phylogenetic trees and no approximation algorithm is known. In [14], we considered the influence of tree topology on the maximum co-divergence problem. In particular, we focused on a particular tree structure, namely caterpillar, and showed that in this case the heuristics that are mostly used in the literature provide solutions that can be arbitrarily far from the optimal value. We then proved that finding the max co-divergence is equivalent to computing the maximum length of a subsequence with certain properties of a given permutation. This equivalence leads to two consequences: (i) it shows that we can compute efficiently in polynomial time the optimal time-feasible reconciliation, and (ii) it can be used to understand how much the tree topology influences the value of the maximum number of co-speciations.

6.5. Axis 4: Human and animal health

Rare disease studies Minor intron splicing plays a central role in human embryonic development and survival. Indeed, biallelic mutations in RNU4ATAC, transcribed into the minor spliceosomal U4atac snRNA, are responsible for three rare autosomal recessive multimalformation disorders named Taybi-Linder (TALS/MOPD1), Roifman (RFMN), and Lowry-Wood (LWS) syndromes, which associate numerous overlapping signs of varying severity. Although RNA-seq experiments have been conducted on a few RFMN patient cells, none have been performed in TALS, and more generally no in-depth transcriptomic analysis of the 700 human genes containing a minor (U12-type) intron had been published as yet. We thus sequenced RNA from cells derived from five skin, three amniotic fluid, and one blood biosamples obtained from seven unrelated TALS cases and from age- and sex-matched controls. This allowed us to describe for the first time the mRNA expression and splicing profile of genes containing U12-type introns, in the context of a functional minor spliceosome. Concerning RNU4ATAC-mutated patients, we showed in [15] that as expected, they display distinct U12-type intron splicing profiles compared to controls, but that rather unexpectedly the mRNA expression levels are mostly unchanged. Furthermore, although U12-type intron missplicing concerns most of the expressed U12 genes, the level of U12-type intron retention is surprisingly low in fibroblasts and amniocytes, and much more pronounced in blood cells. Interestingly, we found several occurrences of introns that can be spliced using either U2, U12, or a combination of both types of splice site consensus sequences, with a shift towards splicing using preferentially U2 sites in TALS patients' cells compared to controls.

This work is part of the PhD of Audric Cologne [1] defended in October 2019.

Cancer studies Circular RNAs (circRNAs) are a class of RNAs that is under increasing scrutiny, although their functional roles are debated. In [30], we analysed RNA-seq data of 348 primary breast cancers and developed a method to identify circRNAs that does not rely on unmapped reads or known splice junctions. We identified 95,843 circRNAs, of which 20,441 were found recurrently. Of the circRNAs that match exon boundaries of the same gene, 668 showed a poor or even negative ($R < 0.2$) correlation with the expression level of the linear gene. An *In silico* analysis showed that only a minority (8.5%) of circRNAs could be explained by known splicing events. Both these observations suggest that specific regulatory processes for circRNAs exist. We confirmed the presence of circRNAs of CNOT2, CREBBP, and RERE in an independent pool of primary breast cancers. We identified circRNA profiles associated with subgroups of breast cancers and with biological and clinical features, such as amount of tumour lymphocytic infiltrate and proliferation index. siRNA-mediated knockdown of circCNOT2 was shown to significantly reduce viability of the breast cancer cell lines MCF-7 and BT-474, further underlining the biological relevance of circRNAs. Furthermore, we found that circular, and not linear, CNOT2 levels are predictive for progression-free survival time to aromatase inhibitor (AI) therapy in advanced breast cancer patients, and found that circCNOT2 is detectable in cell-free RNA from plasma. We showed that circRNAs are abundantly present, show characteristics of being specifically regulated, are associated with clinical and biological properties, and thus are relevant in breast cancer.

Other cancer studies have concerned the automatic discovery of the 100-miRNA signature for cancer classification [21], an Integrative and comparative genomic analysis to identify clinically relevant pulmonary carcinoid groups and unveil the supra-carcinoids [5], [complete with 2 papers not yet entered in Hal], and finally the investigation of new therapeutic interventions that are needed to increase the immunogenicity of tumours and overcome the resistance to these immuno-therapies [29].

Infection studies *Mycoplasma hyopneumoniae* is an economically devastating pathogen in the pig farming industry, however little is known about its relation with the swine host. To improve our understanding on this interaction, we infected epithelial cells with *M. hyopneumoniae* to identify the effects of the infection on the expression of swine genes and miRNAs. In addition, we identified miRNAs differentially expressed (DE) in the extracellular milieu and in exosome-like vesicles released by infected cells. A total of 1,268 genes and 170 miRNAs were DE post-infection ($p<0.05$). We identified the up-regulation of genes related to redox homeostasis and antioxidant defense, most of them putatively regulated by the transcription factor NRF2. Down-regulated genes were enriched in cytoskeleton and ciliary function, which could partially explain *M. hyopneumoniae* induced ciliostasis. Our predictions showed that DE miRNAs could be regulating the aforementioned functions, since we detected down-regulation of miRNAs predicted to target antioxidant genes and up-regulation of miRNAs targeting ciliary and cytoskeleton genes. Based on these observations, *M. hyopneumoniae* seems to elicit an antioxidant response induced by NRF2 in infected cells; in addition, we propose that ciliostasis caused by this pathogen might be related to down-regulation of ciliary genes. The paper presenting these results has been submitted and is in revision.

Others Besides the above, a first step towards deep learning assisted genotype-phenotype association in whole genome-sized data has been explored in the context of predicting amyotrophic lateral sclerosis [34].

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 (2018-2020).
- 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), Audric Cologne, Eric Cumunel, 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. 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/

8.2.1.3. GrR

- Title: Graph Reconfiguration
- Coordinator: N. Bousquet
- ERABLE participant(s): A. Mary
- Type: ANR JCJC (2019-2021)
- Web page: Not available

8.2.1.4. 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

8.2.1.5. 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

8.2.1.6. Networks

- Title: Networks
- Coordinator: Michel Mandjes, University of Amsterdam
- ERABLE participant(s): S. Pissis, L. Stougie
- Type: NWO Gravity Program (2014-2024)
- Web page: https://www.thenetworkcenter.nl/

8.2.1.7. 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

8.2.1.8. 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

8.2.1.9. 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), 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. 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.

8.2.3.1. AHeAD

- Title: efficient Algorithms for HArnessing networked Data
- Coordinator: G. Italiano
- ERABLE participant(s): R. Grossi, G. Italiano
- Type: MUIR PRIN, Italian Ministry of Education, University and Research (2019-2022)
- Web page: https://sites.google.com/view/aheadproject

8.2.3.2. CMACBioSeq

- Title: Combinatorial Methods for analysis and compression of biological sequences
- Coordinator: G. Rosone
- ERABLE participant(s): N. Pisanti
- Type: SIR, MIUR PRIN, Italian Ministry of Research National Projects (2015-2019)
- Web page: http://pages.di.unipi.it/rosone/CMACBioSeq.html

8.2.3.3. MyOwnResearch

- Title: MyOwnResearch: Homogeneous subgroup identification in fatigue management across chronic immune diseases through single subject research design
- Coordinator: A. Schönhuth
- ERABLE participant(s): A. Schönhuth
- Type: Health Holland project (2018-2021)
- Web page: Not available

8.2.3.4. 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

8.3. European Initiatives

8.3.1. Collaborations in European Programs, Except FP7 & H2020

8.3.1.1. Pangaia

- Title: Pan-genome Graph Algorithms and Data Integration
- Coordinator: Paola Bonizzoni, University of Milan, Italy
- ERABLE participant(s): S. Pissis, A. Schönhuth, L. Stougie
- Type: H2020 MSCA-RISE (2020-2022)
- Web page: Not available

8.3.2. 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 then became a European Inria Team.

8.4. International Initiatives

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

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 Other 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. This year (2019) is the final one. 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 participates in Network for Organismal Interactions Research (NOIR), a project funded by Conicyt in Chile within the call Internation Networking between Research Centers. The project started in 2019 and will last until the end of 2020. The coordinator on the Chilean side is Elena Vida from the Universidad Mayor, Santiago, Chile, and the Erable participants are Carol Moraga Quinteros, Mariana Ferrarini and Marie-France Sagot.

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 Portuguese partner. Perseids ended in December 2019.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

In 2019, ERABLE greeted the following International scientists:

- In France: Alexandra Carvalho and Susana Vinga, Assistant and Associate professors resp., Instituto Superior Técnico, Lisbon, Portugal; Helisson Faoro, researcher, Instituto Carlos Chagas, Fiocruz, Paraná, Brazil; Ariel Silber, professor, Universidade de São Paulo, Brazil; Arnaldo Zaha, professor at Universidade Federal do Rio Grande do Sul, Brazil.
- In Italy: Travis Gaggie, Associate professor, Dalhousie University; Nicola Prezza, postdoc, University of Pisa; Elena Arseneva, Assistant professor, St Petersburg State University, Blerina Sinaimeri, Junior Researcher, Inria (see below); Marie-France Sagot, Senior researcher, Inria (see below).
- In the Netherlands: Wiktor Zuba, PhD student, University of Warsaw; Lorraine Ayad, Lecturer, King's College London; Grigorios Loukides, Lecturer, King's College London; Martin Farach-Colton, Professor, Rutgers University; Grigorios Loukides, Lecturer, King's College London; Martin Dyer, Professor, University of Leeds.

8.5.1.1. Internships

In 2019, ERABLE in France greeted the following Internships:

- Phablo Moura, postdoc, University of Campinas, Brazil.
- Diego Pérez and Evelyn Sanchéz, PhD students of Elena Vidal, Universidad Mayor, Santiago, Chile.

In the Netherlands, ERABLE greeted the following Internships: Luca Denti, University Bicocca of Milano, Italy, from October 2018 to January 2019, Mick van Dijk, TU Delft, from May 2018 to January 2019, Giulia Barnardini, University Bicocca of Milano, Italy, from September 2018 to November 2019.

8.5.2. Visits to International Teams

8.5.2.1. Sabbatical programme

From July 2019 to June 2020, Blerina Sinaimeri was on Sabbatical at Luiss University to work with Giuseppe Italiano, member of Erable.

8.5.2.2. Research Stays Abroad

In 2019, Marie-France Sagot visited Luiss University for 11 days as Visiting Professor from LUISS University to work with Blerina Sinaimeri who is on Sabbatical at Luiss University from July 2019 to June 2020, and with Giuseppe Italiano, member of Erable. While there, M.-F. Sagot also worked with Alberto Marchetti-Spaccamela from Sapienza University of Rome and from Erable.

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 of 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

- Leen Stougie was co-organiser of MAPSP 2019, Jun 2019, Hotel Zeeuwse Stromen, Renesse; and of the Networks Workshop on Random graphs, counting and sampling, Sep 2019, CWI, Amsterdam.

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 *APF*, *ATMOS*, and *CIAC*.
- Arnaud Mary was a member of the Program Committee of *MFCS*, and *WEPA*.
- Nadia Pisanti was a member of the Program Committee of *BIOINFORMATICS*, *CPM*, *ICCS*, *ISBRA*, *IWOCA*, and *WABI*.
- Marie-France Sagot was a member of the Program Committee of *BIBM*, *CIAC*, *CPM*, *PSC*, *RecombCG*, d an*WABI*.

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

- 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*.
- Vincent Lacroix is recommender for *Peer Community in Genomics*, see https://genomics.peercommunityin.org/.
- 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: invited talk on “2-Connectivity on Directed Graphs", 14th Computer Science Symposium in Russia (CSR 2019), Novosibirsk, Russia.
- Nadia Pisanti: Invited talk on “Mapping Reads on a Pan-Genome: Pattern Matching on Degenerate Texts", 1st Workshop on Computational Pan-Genomics, Bielefeld, Germany; Invited talk on “Online (approximate) Pattern Matching on Degenerate Texts and Applications", 14th Workshop on Compression, Text and Algorithms (WCTA), Segovia, Spain.
- Solon Pissis: Invited talk on “Even Faster Elastic-Degenerate String Matching via Fast Matrix Multiplication", Algorithms Group Seminar Series, 24 Oct 2019, University of Warsaw, Warsaw, Poland; Invited talk on “When linear space is impractical: computing absent words in output-sensitive space", Bonsai Bioinformatics Seminar Series, 11 Jun 2019, Université de Lille, Lille, France; Invited talk on “Elastic-degenerate strings: a new representation for pattern matching in a collection of similar texts", Computer Science Seminar Series, Feb 12 2019, University of Pisa, Pisa, Italy.
- Leen Stougie: Invited talk on “Fixed-Order Scheduling on parallel machines", Workshop on Combinatorial Optimization, 26-27 September 2019, TU Berlin, Germany.
- Cristina Vieira: Invited talk on “Contribution of Transposable element to gene expression in Drosophila (and other)", XI Symposium of Ecology, Genetic and Drosophila Evolution, November 2019, Pelotas, Brazil.

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.

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. He is member of the General Board of the Dutch Network on the Mathematics of Operations Research (Landelijk Netwerk Mathematische Besliskunde (LNMB)), and member of the Management Team of the Gravity project Networks.

Alain Viari 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 CoNRS8.

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 the M1 master in bioinformatics (https://www.bioinfo-lyon.fr/) and of the following courses (L3: Advanced Bioinformatics, M1: Methods for Data Analysis in Genomics, M1: Methods for Data Analysis in Transcriptomics, M1: Bioinformatics Project, M2: Ethics). He taught 96 hours in 2018-2019 and 192 hours in 2019-2020. Arnaud Mary is responsible for three courses of the Bioinformatics Curriculum at the University (L2: Introduction to Bioinformatics and Biostatistics, M1: Object Oriented Programming, M2: new course on Advanced Algorithms for Bioinformatics) and one at Insa (Discrete Mathematics). He taught 198 hours in 2019. Blerina Sinaimeri taught 36 hours in 2019 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 and Audric Cologne were instructors in the NGS data analysis training for the CNRS Formation, a course coordinated by Annabelle Haudry, LBBE (https://cnrsformation.cnrs.fr/stage-19026-Bioinformatique-pour-le-traitement-de-donnees-de-sequencage-%28NGS%29—Lyon.html).

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 60 and 100 hours per year, again at the undergraduate and Master levels, in applied mathematics (*e.g.* Operational Research, Advanced Linear Programming), machine learning (Deep Learning) and computational biology (*e.g.* Biological Network Analysis, Algorithms for Genomics).

9.2.2. Supervision

The following PhDs were defended in ERABLE in 2019:

- Jasmijn Baaijens, CWI (supervisor: Alexander Schönhuth), Sep 2019
- Annelieke Baller, Vrije Universiteit Amsterdam (co-supervisor: Leen Stougie), Nov 2019
- Thomas Bosman, Vrije Universiteit Amsterdam (co-supervisor: Leen Stougie), Nov 2019
- Audric Cologne, University of Lyon 1 (funded by Inserm and Inria, co-supervisors: Patrick Edery – Federation of Health Research of Lyon-Est, Vincent Lacroix), Oct 2019
- 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), Apr 2019
- Nikos Parotsidis, University of Rome Tor Vergata, supervisor: Giuseppe Italiano, Mar 2019
- 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), Feb 2019

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)
- Nicolas Homberg, Inra, Inria & University of Lyon 1 (funded by Inra & Inria, co-supervisors: Christine Gaspin at Inra; 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)
- Camille Sessegolo, University of Lyon 1 (funded by ANR Aster; co-supervisors: Vincent Lacroix, Arnaud Mary)
- Michelle Sweering, CWI (co-supervisors: Solon Pissis and Leen Stougie)
- 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)

9.2.3. Juries

The following are the PhD or HDR juries to which members of ERABLE participated in 2019.

- Vincent Lacroix: External reviewer of the PhD of Patricia Sieber, supervised by Stefan Schuster at Friedrich-Schiller University of Jena, Germany; external reviewer of the PhD of Luca Denti, supervised by Paola Bonizzoni at University Bicocca of Milano, Italy.
- Arnaud Mary: External reviewer of the PhD of Karima Ennaoui, supervised by Lhouari Nourine at University of Clermont-Ferrand, France.
- Marie-France Sagot: External Reviewer of the PhD of Pierre Marijon, University of Lille, France, Dec 2019.
- Leen Stougie: Reading Committee of the PhD of Teun Janssen, TU Delft, Mar 2019; Chair Reading Committee of the PhD of Pieter Kleer, Vrije Universiteit Amsterdam, Sep 2019; Reading Committee of the PhD of Peter van der Gulik, Univ. of Amsterdam, Sep 2019; Chair Reading Committee of the PhD of Maaike Hoogeboom, Vrije Universiteit Amsterdam, Dec 2019.
- Cristina Vieira: Member of the PhD Committee of Olivier Tabone, Faculté de Médecine Rockfeller, Jan 2019; Member of the PhD Committee of Sébastien Lemaire, ENS Lyon, Mar 2019; External Reviewer of the PhD of Natalia Martinez, Université Paris Sud, Oct 2019.

9.3. Popularization

9.3.1. Interventions

Carol Moraga Quinteros participated in the contest “DESCRYPThèse” of the doctoral school E2M2 of the University of Lyon 1, winning a prize for one of the best presentations in April 2019. The title of the talk was “BrumiR: un algorithme *de novo* pour prédire les petits ARNs sans génome de référence”.

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] A. COLOGNE.*Exploration of minor splicing function during embryonic development with the Taybi-Linder Syndrome (TALS) model*, Université de Lyon, October 2019, https://tel.archives-ouvertes.fr/tel-02363211

[2] L. ISHI SOARES DE LIMA.*De novo algorithms to identify patterns associated with biological events in de Bruijn graphs built from NGS data*, Université de Lyon ; Università degli studi di Roma "Tor Vergata", April 2019, https://tel.archives-ouvertes.fr/tel-02280110

[3] T. PUSA.*Modélisation mathématique des impacts de l'environnement à l'aide de réseaux métaboliques et de la théorie des jeux*, Université de Lyon ; Università degli studi La Sapienza (Rome), February 2019, https://tel.archives-ouvertes.fr/tel-02096971

Articles in International Peer-Reviewed Journal

[4] V. ACUÑA, R. GROSSI, G. F. ITALIANO, L. LIMA, R. RIZZI, G. SACOMOTO, M.-F. SAGOT, B. SINAIMERI.*On Bubble Generators in Directed Graphs*, in "Algorithmica", 2019, p. 1-19 [*DOI :* 10.1007/S00453-019-00619-Z], https://hal.inria.fr/hal-02284946

[5] N. ALCALA, N. LEBLAY, A. GABRIEL, L. MANGIANTE, D. HERVÁS, T. GIFFON, A.-S. SERTIER, A. FERRARI, J. DERKS, A. GHANTOUS, T. DELHOMME, A. CHABRIER, C. CUENIN, B. ABEDI-ARDEKANI, A. BOLAND, R. OLASO, V. MEYER, J. ALTMULLER, F. LE CALVEZ-KELM, G. DURAND, C. VOEGELE, S. BOYAULT, L. MOONEN, N. LEMAÎTRE, P. LORIMIER, A.-C. TOFFART, A. SOLTERMANN, J. CLEMENT, J. SAENGER, J. FIELD, M. BREVET, C. BLANC-FOURNIER, F. GALATEAU-SALLÉ, N. LE STANG, P. RUSSELL, G. WRIGHT, G. SOZZI, U. PASTORINO, S. LACOMME, J. VIGNAUD, V. HOFMAN, P. HOFMAN, O. T. BRUSTUGUN, M. LUND-IVERSEN, V. THOMAS DE MONTPREVILLE, L. A. MUSCARELLA, P. GRAZIANO, H. H. POPPER, J. STOJSIC, J.-F. DELEUZE, Z. HERCEG, A. VIARI, P. NUERNBERG, G. PELOSI, A.-M. C. DINGEMANS, M. MILIONE, L. ROZ, L. BRCIC, M. VOLANTE, M. PAPOTTI, C. CAUX, J. SANDOVAL, H. HERNANDEZ-VARGAS, E. BRAMBILLA, E. SPEEL, N. GIRARD, S. LANTUEJOUL, J. MCKAY, M. FOLL, L. FERNANDEZ-CUESTA.*Integrative and comparative genomic analyses identify clinically relevant pulmonary carcinoid groups and unveil the supra-carcinoids*, in "Nature Communications", December 2019, vol. 10, n° 1, p. 1-21 [*DOI :* 10.1038/S41467-019-11276-9], https://hal.inria.fr/hal-02339242

[6] 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", 2019, vol. 27, n° 3, p. 268-281 [*DOI :* 10.1016/J.TIM.2018.11.004], https://hal.inria.fr/hal-01964983

[7] J. A. BAAIJENS, A. SCHÖNHUTH.*Overlap graph-based generation of haplotigs for diploids and polyploids*, in "Bioinformatics", November 2019, vol. 35, n° 21, p. 4281-4289 [*DOI :* 10.1093/BIOINFORMATICS/BTZ255], https://hal.inria.fr/hal-02344853

[8] A. C. BALLER, M. VAN EE, M. HOOGEBOOM, L. STOUGIE.*Complexity of inventory routing problems when routing is easy*, in "Networks", October 2019, p. 1-11 [*DOI :* 10.1002/NET.21908], https://hal.inria.fr/hal-02422721

[9] S. K. BARUAH, V. BONIFACI, R. BRUNI, A. MARCHETTI-SPACCAMELA.*ILP models for the allocation of recurrent workloads upon heterogeneous multiprocessors*, in "Journal of Scheduling", April 2019, vol. 22, n° 2, p. 195-209 [*DOI :* 10.1007/S10951-018-0593-X], https://hal.inria.fr/hal-02339161

[10] G. BERNARDINI, P. BONIZZONI, L. DENTI, M. PREVITALI, A. SCHÖNHUTH.*MALVA: Genotyping by Mapping-free ALlele Detection of Known VAriants*, in "iScience", 2019, vol. 18, p. 20-27 [*DOI :* 10.1016/J.ISCI], https://hal.inria.fr/hal-02344254

[11] G. BERNARDINI, N. PISANTI, S. PISSIS, G. ROSONE.*Approximate pattern matching on elastic-degenerate text*, in "Theoretical Computer Science", August 2019, p. 1-30 [*DOI :* 10.1016/J.TCS.2019.08.012], https://hal.inria.fr/hal-02298622

[12] V. BONIFACI, A. WIESE, S. K. BARUAH, A. MARCHETTI-SPACCAMELA, S. STILLER, L. STOUGIE.*A Generalized Parallel Task Model for Recurrent Real-time Processes*, in "ACM Transactions on Parallel Computing", June 2019, vol. 6, n^o^ 1, p. 63-72 [*DOI :* 10.1109/RTSS.2012.59], https://hal.inria.fr/hal-02347467

[13] M. BONNIN, N. FARES, B. TESTONI, Y. ESTORNES, K. WEBER, B. VANBERVLIET, L. LEFRANCOIS, A. GARCIA, A. KFOURY, F. PEZ, I. COSTE, P. SAINTIGNY, A. VIARI, K. LANG, B. GUEY, V. HERVIEU, B. BANCEL, B. BARTOCH, D. DURANTEL, T. RENNO, P. MERLE, S. LEBECQUE.*Toll-like receptor 3 downregulation is an escape mechanism from apoptosis during hepatocarcinogenesis*, in "Journal of Hepatology", October 2019, vol. 71, n^o^ 4, p. 763-772 [*DOI :* 10.1016/J.JHEP.2019.05.031], https://hal.inria.fr/hal-02367101

[14] T. CALAMONERI, A. MONTI, B. SINAIMERI.*Co-divergence and tree topology*, in "Journal of Mathematical Biology", August 2019, vol. 79, n^o^ 3, p. 1149-1167 [*DOI :* 10.1007/S00285-019-01385-W], https://hal.inria.fr/hal-02298643

[15] A. COLOGNE, C. BENOIT-PILVEN, A. BESSON, A. PUTOUX, A. CAMPAN-FOURNIER, M. B. BOBER, C. E. DE DIE-SMULDERS, A. PAULUSSEN, L. PINSON, A. TOUTAIN, C. M. ROIFMAN, A.-L. LEUTENEGGER, S. MAZOYER, P. EDERY, V. LACROIX.*New insights into minor splicing-a transcriptomic analysis of cells derived from TALS patients*, in "RNA", 2019, p. 1-21 [*DOI :* 10.1261/RNA.071423.119], https://hal.inria.fr/hal-02305628

[16] M. FABLET, A. JACQUET, R. REBOLLO, A. A. HAUDRY, C. REY, J. SALCES-ORTIZ, P. BAJAD, N. BURLET, M. JANTSCH, M. P. GARCIA GUERREIRO, C. VIEIRA.*Dynamic Interactions Between the Genome and an Endogenous Retrovirus: Tirant in Drosophila simulans Wild-Type Strains*, in "G3", March 2019, vol. 9, n^o^ 3, p. 855-865 [*DOI :* 10.1534/G3.118.200789], https://hal.archives-ouvertes.fr/hal-02004384

[17] L. GEORGIADIS, G. F. ITALIANO, A. KARANASIOU.*Approximating the smallest 2-vertex connected spanning subgraph of a directed graph*, in "Theoretical Computer Science", September 2019, p. 1-16 [*DOI :* 10.1016/J.TCS.2019.09.040], https://hal.inria.fr/hal-02335015

[18] L. GUILLOT, L. DELAGE, A. VIARI, Y. VANDENBROUCK, E. COM, A. A. RITTER, R. LAVIGNE, D. MARIE, P. PETERLONGO, P. POTIN, C. PINEAU.*Peptimapper: proteogenomics workflow for the expert annotation of eukaryotic genomes*, in "BMC Genomics", January 2019, vol. 20, n^o^ 1, 56 [*DOI :* 10.1186/S12864-019-5431-9], https://hal.inria.fr/hal-01987197

[19] E. LERAT, J. CASACUBERTA, C. CHAPARRO, C. VIEIRA.*On the Importance to Acknowledge Transposable Elements in Epigenomic Analyses*, in "Genes", March 2019, vol. 10, n^o^ 4, 258 [*DOI :* 10.3390/GENES10040258], https://hal.archives-ouvertes.fr/hal-02093613

[20] L. LIMA, C. MARCHET, S. CABOCHE, C. DA SILVA, B. ISTACE, J.-M. AURY, H. TOUZET, R. CHIKHI.*Comparative assessment of long-read error correction software applied to Nanopore RNA-sequencing data*, in "Briefings in Bioinformatics", June 2019, p. 1-18 [*DOI :* 10.1093/BIB/BBZ058], https://hal.archives-ouvertes.fr/hal-02394395

[21] A. Lopez-Rincon, M. Martinez-Archundia, G. U. Martinez-Ruiz, A. Schönhuth, A. Tonda.*Automatic discovery of 100-miRNA signature for cancer classification using ensemble feature selection*, in "BMC Bioinformatics", December 2019, vol. 20, n° 1, p. 1-17 [*DOI :* 10.1186/s12859-019-3050-8], https://hal.inria.fr/hal-02344257

[22] P. Marin, J. Genitoni, D. D. Barloy, S. Maury, P. Gibert, C. K. Ghalambor, C. Vieira.*Biological Invasion: The Influence of the Hidden Side of the (Epi)Genome*, in "Functional Ecology", 2019, p. 1-55, forthcoming [*DOI :* 10.1111/1365-2435.13317], https://hal-agrocampus-ouest.archives-ouvertes.fr/hal-02063295

[23] A. Marino, L. Versari, A. Conte, R. Grossi, G. Punzi, T. Uno.*Polynomial-Delay Enumeration of Maximal Common Subsequences*, in "SIAM Journal on Discrete Mathematics", October 2019, vol. 33, n° 2, p. 189-202 [*DOI :* 10.1007/978-3-030-32686-9_14], https://hal.inria.fr/hal-02338458

[24] A. Mary, Y. Strozecki.*Efficient enumeration of solutions produced by closure operations*, in "Discrete Mathematics and Theoretical Computer Science", June 2019, p. 1-30, https://arxiv.org/abs/1712.03714 - 30 pages, 1 figure. Long version of the article arXiv:1509.05623 of the same name which appeared in STACS 2016. Final version for DMTCS journal [*DOI :* 10.23638/DMTCS-21-3-22], https://hal.inria.fr/hal-02373737

[25] N. Prezza, N. Pisanti, M. Sciortino, G. Rosone.*SNPs detection by eBWT positional clustering*, in "Algorithms for Molecular Biology", December 2019, vol. 14, n° 1, p. 1-13 [*DOI :* 10.1186/s13015-019-0137-8], https://hal.inria.fr/hal-02335605

[26] T. Pusa, M. G. Ferrarini, R. Andrade, A. Mary, A. Marchetti-Spaccamela, L. Stougie, M.-F. Sagot.*MOOMIN – Mathematical explOration of 'Omics data on a MetabolIc Network*, in "Bioinformatics", August 2019, p. 1-10 [*DOI :* 10.1093/bioinformatics/btz584], https://hal.inria.fr/hal-02284835

[27] T. Pusa, M. Wannagat, M.-F. Sagot.*Metabolic Games*, in "Frontiers in Applied Mathematics and Statistics", April 2019, vol. 5, p. 1-13 [*DOI :* 10.3389/fams.2019.00018], https://hal.inria.fr/hal-02336595

[28] C. Sessegolo, C. Cruaud, C. Da Silva, A. Cologne, M. Dubarry, T. Derrien, V. Lacroix, J.-M. Aury.*Transcriptome profiling of mouse samples using nanopore sequencing of cDNA and RNA molecules*, in "Scientific Reports", December 2019, vol. 9, n° 1, p. 1-12 [*DOI :* 10.1038/s41598-019-51470-9], https://hal.inria.fr/hal-02335574

[29] T. Shekarian, E. Sivado, A.-C. Jallas, S. Depil, J. Kielbassa, I. Janoueix-Lerosey, G. Hutter, N. Goutagny, C. Bergeron, A. Viari, S. Valsesia-Wittmann, C. Caux, A. Marabelle.*Repurposing rotavirus vaccines for intratumoral immunotherapy can overcome resistance to immune checkpoint blockade*, in "Science Translational Medicine", October 2019, vol. 11, n° 515, eaat5025 [*DOI :* 10.1126/scitranslmed.aat5025], https://hal.inria.fr/hal-02347852

[30] M. Smid, S. M. Wilting, K. Uhr, F. G. Rodríguez-González, V. de Weerd, W. J. Prager-Van der Smissen, M. van der Vlugt-Daane, A. van Galen, S. Nik-Zainal, A. P. Butler, S. Martin, H. R. Davies, J. Staaf, M. J. van de Vijver, A. L. Richardson, G. MacGrogan, R. Salgado, G. G. Van den Eynden, C. A. Purdie, A. M. Thompson, C. Caldas, P. N. Span, F. C. Sweep, P. T. Simpson, S. R. Lakhani, S. Van Laere, C. Desmedt, A. Paradiso, J. Eyfjord, A. Broeks, A. Vincent-Salomon, A. P. Futreal, S. Knappskog, T. King, A. Viari, A.-L. Børresen-Dale, H. G. Stunnenberg, M. Stratton, J. A. Foekens, A. M. Sieuwerts, J. W.

MARTENS.*The circular RNome of primary breast cancer*, in "Genome Research", March 2019, vol. 29, n° 3, p. 356-366 [*DOI :* 10.1101/GR.238121.118], https://hal.inria.fr/hal-02338656

[31] 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", January 2019, vol. 75, n° 1, p. 11-15 [*DOI :* 10.1016/J.EURURO.2018.08.026], https://hal.inria.fr/hal-01921597

[32] O. TRÉDAN, Q. WANG, D. PISSALOUX, P. CASSIER, A. DE LA FOUCHARDIÈRE, J. FAYETTE, F. DESSEIGNE, I. RAY-COQUARD, C. DE LA FOUCHARDIÈRE, D. FRAPPAZ, P. E. HEUDEL, A. BONNEVILLE-LEVARD, A. FLECHON, M. SARABI, P. GUIBERT, T. BACHELOT, M. PÉROL, B. YOU, N. BONNIN, O. COLLARD, C. LEYRONNAS, V. ATTIGNON, C. BAUDET, E. SOHIER, J.-P. VILLEMIN, A. VIARI, S. BOYAULT, S. LANTUEJOUL, S. PAINDAVOINE, I. TREILLLEUX, C. RODRIGUEZ, V. AGRAPART, V. CORSET, G. GARIN, S. CHABAUD, D. PEROL, J.-Y. BLAY.*Molecular screening program to select molecular-based recommended therapies for metastatic cancer patients: analysis from the ProfiLER trial*, in "Annals of Oncology", May 2019, vol. 30, n° 5, p. 757-765 [*DOI :* 10.1093/ANNONC/MDZ080], https://hal.inria.fr/hal-02367110

[33] 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", 2019, p. 1-11 [*DOI :* 10.1109/TCBB.2018.2838667], https://hal.inria.fr/hal-01842451

[34] B. YIN, M. BALVERT, R. VAN DER SPEK, B. E. DUTILH, S. BOHTE, J. VELDINK, A. SCHÖNHUTH.*Using the structure of genome data in the design of deep neural networks for predicting amyotrophic lateral sclerosis from genotype*, in "Bioinformatics", July 2019, vol. 35, n° 14, p. i538-i547 [*DOI :* 10.1093/BIOINFORMATICS/BTZ369], https://hal.inria.fr/hal-02344253

International Conferences with Proceedings

[35] A. ABBOUD, L. GEORGIADIS, G. F. ITALIANO, R. KRAUTHGAMER, N. PAROTSIDIS, O. TRABELSI, P. UZNAŃSKI, D. WOLLEB-GRAF.*Faster Algorithms for All-Pairs Bounded Min-Cuts*, in "ICALP 2019 - 46th International Colloquium on Automata, Languages and Programming", Patras, Greece, July 2019, p. 1-15 [*DOI :* 10.4230/LIPICS.ICALP.2019.7], https://hal.inria.fr/hal-02335025

[36] G. BERNARDINI, P. GAWRYCHOWSKI, N. PISANTI, S. PISSIS, G. ROSONE.*Even Faster Elastic-Degenerate String Matching via Fast Matrix Multiplication*, in "ICALP 2019 - 46th International Colloquium on Automata, Languages and Programming", Patras, Greece, July 2019, p. 1-15 [*DOI :* 10.4230/LIPICS.ICALP.2019.21], https://hal.inria.fr/hal-02298621

[37] M. BONAMY, N. BOUSQUET, M. HEINRICH, T. ITO, Y. KOBAYASHI, A. MARY, M. MÜHLENTHALER, K. WASA.*The Perfect Matching Reconfiguration Problem*, in "MFCS 2019 - 44th International Symposium on Mathematical Foundations of Computer Science", Aachen, Germany, August 2019, p. 1-14 [*DOI :* 10.4230/LIPICS.MFCS.2019.80], https://hal.inria.fr/hal-02335588

[38] A. CONTE, R. GROSSI, M. M. KANTÉ, A. MARINO, T. UNO, K. WASA.*Listing Induced Steiner Subgraphs as a Compact Way to Discover Steiner Trees in Graphs*, in "MFCS 2019 - 44th International Symposium on Mathematical Foundations of Computer Science", Aachen, Germany, August 2019, p. 1-14 [*DOI :* 10.4230/LIPICS.MFCS.2019.73], https://hal.inria.fr/hal-02335601

[39] A. CONTE, R. GROSSI, A. MARINO, L. TATTINI, L. VERSARI.*A Fast Discovery Algorithm for Large Common Connected Induced Subgraphs*, in "WEPA 2019 - Workshop on Enumeration Problems & Applications", Awaji Island, Japan, October 2019, p. 1-26, https://hal.inria.fr/hal-02338435

[40] A. CONTE, R. GROSSI, G. PUNZI, T. UNO.*Polynomial-Delay Enumeration of Maximal Common Subsequences*, in "SPIRE 2019 - 26th International Symposium on String Processing and Information Retrieval", Segovia, Spain, October 2019, p. 189-202 [*DOI :* 10.1007/978-3-030-32686-9_14], https://hal.inria.fr/hal-02338437

[41] M. EQUI, R. GROSSI, V. MAKINEN.*On the Complexity of Exact Pattern Matching in Graphs: Binary Strings and Bounded Degree*, in "ICALP 2019 - 46th International Colloquium on Automata, Languages and Programming", Patras, Greece, July 2019, p. 1-15, https://hal.inria.fr/hal-02338498

[42] F. GERACI, R. MASSIDDA, N. PISANTI.*Technology and Species independent Simulation of Sequencing data and Genomic Variants*, in "BIBE 2019 - 19th annual IEEE International Conference on BioInformatics and BioEngineering", Athens, Greece, IEEE, October 2019, p. 1-8, https://hal.inria.fr/hal-02336600

[43] N. HJULER, G. F. ITALIANO, N. PAROTSIDIS, D. SAULPIC.*Dominating Sets and Connected Dominating Sets in Dynamic Graphs*, in "STACS 2019 - 36th International Symposium on Theoretical Aspects of Computer Science", Berlin, Germany, March 2019, p. 1-17 [*DOI :* 10.4230/LIPICS.STACS.2019.35], https://hal.inria.fr/hal-02335028

[44] G. F. ITALIANO, S. LATTANZI, V. MIRROKNI, N. PAROTSIDIS.*Dynamic Algorithms for the Massively Parallel Computation Model*, in "SPAA 2019 - 31st ACM Symposium on Parallelism in Algorithms and Architectures", Phoenix, France, ACM Press, June 2019, p. 49-58, https://arxiv.org/abs/1905.09175 [*DOI :* 10.1145/3323165.3323202], https://hal.inria.fr/hal-02335011

Scientific Books (or Scientific Book chapters)

[45] N. PISANTI.*Algorithms Foundations*, in "Encyclopedia of Bioinformatics and Computational Biology", Elsevier, 2019, vol. 1, p. 1-4 [*DOI :* 10.1016/B978-0-12-809633-8.20315-4], https://hal.inria.fr/hal-01964689

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, HDR]
Aline Marguet [Inria, Researcher, from Oct 2019]
Delphine Ropers [Inria, Researcher]

Faculty Members

Johannes Geiselmann [Univ Grenoble Alpes, Professor, HDR]
Stéphan 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, Engineer]
Corinne Pinel [CNRS, Engineer]
Célia Boyat [Univ Grenoble Alpes]

PhD Students

Thibault Etienne [INRA, PhD Student]
Antrea Pavlou [Inria, PhD Student]
Maaike Sangster [Inria, PhD Student, from Oct 2019]

Post-Doctoral Fellows

Aline Marguet [Inria, Post-Doctoral Fellow, until Sep 2019]
Marco Mauri [Inria, Post-Doctoral Fellow, until Apr 2019]

Visiting Scientist

Tomas Gedeon [Professor in Mathematics, Montana State University, from May 2019 until Aug 2019]

Administrative Assistants

Alexandra Fitzgerald [Inria, Administrative Assistant, until Sep 2019]
Diane Courtiol [Inria, Administrative Assistant, from Oct 2019]

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.

[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.

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 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 80 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.

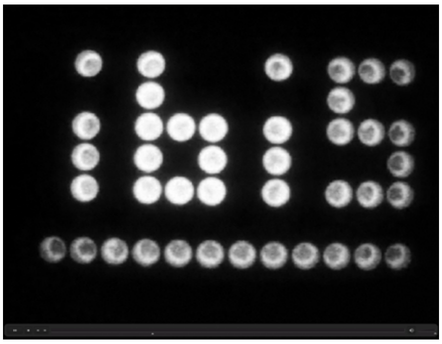


Figure 1. Display of the project-team name on a "bacterial billboard" (see http://team.inria.fr/ibis for the corresponding movie). A microplate containing a minimal medium (with glucose and acetate) is filmed during 36 hours. Wells contain E. coli bacteria which are transformed with a reporter plasmid 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.

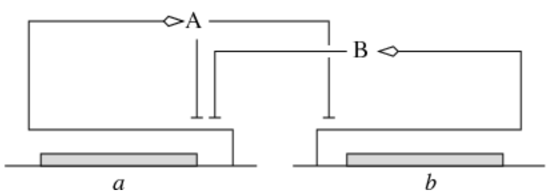
The dynamics of gene regulatory networks can be modeled by means of ordinary differential equations (ODEs), describing the rate of synthesis and degradation of the gene products as well as regulatory interactions

[0] See http://team.inria.fr/ibis for a complete list.

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.

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, Stéphan 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.

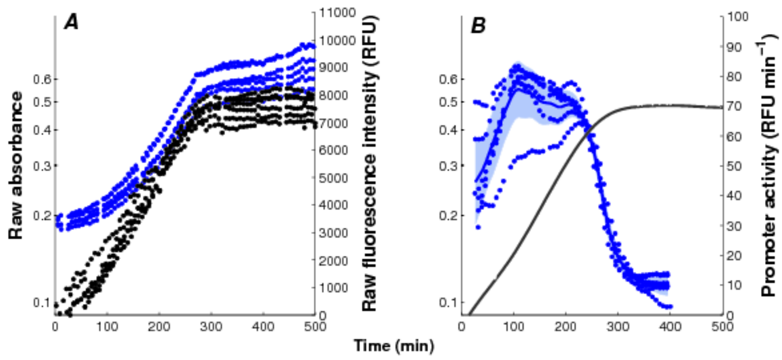


Figure 3. Monitoring of bacterial gene expression in vivo using fluorescent reporter genes (Stefan et al., PLoS Computational Biology, 11(1):e1004028, 2015). The plots show the primary data obtained in a kinetic experiment with E. coli cells, focusing on the expression of the motility gene tar in a mutant 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 ± 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, Stéphan 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], Stéphan Lacour, Marco Mauri, Tamas Muszbek, Michel Page, Antrea Pavlou, Delphine Ropers, Maaike Sangster.

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.

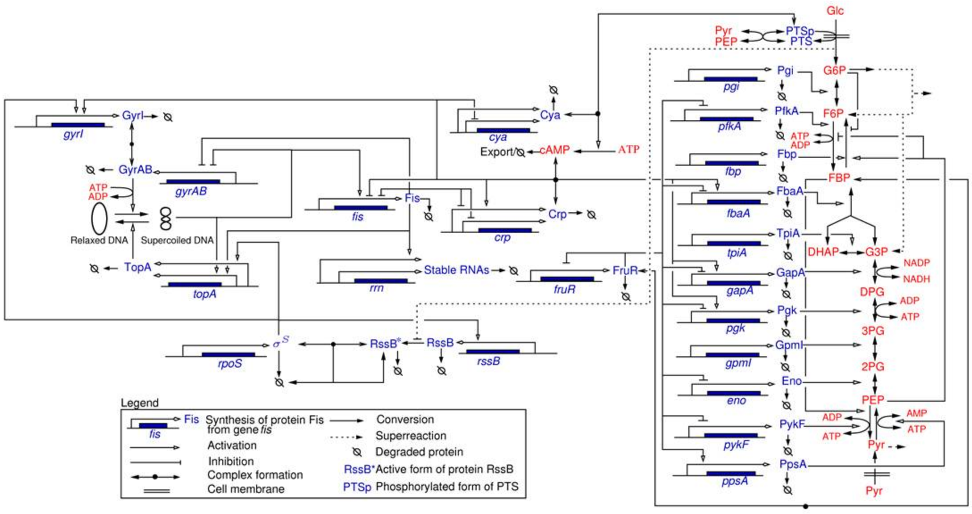


Figure 4. Network of key genes, proteins, and regulatory interactions involved in the carbon assimilation network in E. coli (Baldazzi et al., PLoS Computational Biology, 6(6):e1000812, 2010). The metabolic part includes the glycolysis/gluconeogenesis pathways as well as a simplified description of the 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.

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 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.

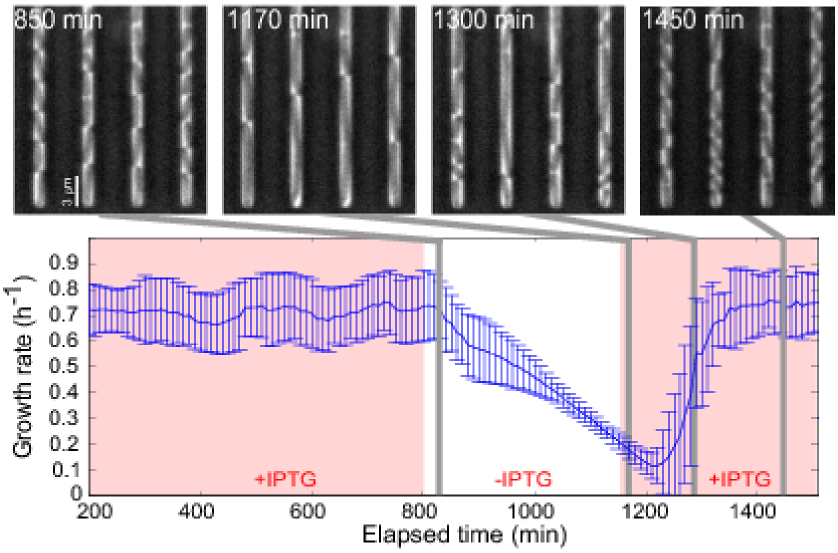


Figure 5. Growth arrest by external control of the gene expression machinery (Izard, Gomez Balderas et al., Molecular Systems Biology, 11:840, 2015). An E. coli strain in which an essential component of the gene expression machinery, the $\beta\beta'$ subunits of RNA polymerase, was put under the control 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

A publication on the use of mixed-effects models for the analysis of the inheritance and variability of gene expression parameters along lineage trees was published in a special issue of *Bioinformatics* and presented at the major bioinformatics conference ISMB/ECCB 2020. A publication in *BMC Bioinformatics* accompanied the release of the new version of the web application WELLINVERTER for the analysis of fluorescent reporter gene data. IBIS member Michel Page launched his start-up ProLeads (https://proleads.fr/), a specialized business search engine.

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: Publication in BMC Bioinformatics describing the new version of 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). Publication in BMC Bioinformatics 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. 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. Tutorial on the use of the model formalism for analyzing synthetic genetic circuits.

- 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 [13], 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. 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. An article 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, was published in *BMC Bioinformatics* this year [22]. We notably show that the expression pattern in different growth media, supporting different growth rates, corresponds to the pattern expected for a constitutive gene.

6.2. 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 6.1. 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. The investigation of the problem of reconstructing promoter activity statistics from reporter gene population snapshot data has led 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. The method has been demonstrated on simulated data. Theoretical as well as simulation resuts have been published in *Automatica* this year [15], and will be the object of application to real data.

One of the key limitations of the method is the assumption of stationary promoter activation statistics. In the context of controlled gene expression processes, this may hamper applicability of the method. In response to this, an extension of the method for so-called modulated processes (stationary processes reshaped by a time-varying control input), has been developed and demonstrated on simulations of controlled gene expression. Results were submitted for possible presentation and publication in the proceedings of the IFAC world congress 2020.

6.3. Mathematical analysis of structured branching populations

The investigation of cellular populations at the single-cell level has led to the discovery of important phenomena, such as the co-occurrence of different phenotypes in an isogenic population. Novel experimental techniques, such as time-lapse fluorescence microscopy combined with the use of microfluidic devices (Section 3.2), enable one to take the investigation further by providing time-course profiles of the dynamics of individual cells over entire lineage trees. The development of models that take into account the genealogy of individual cells is an important step in the study of inheritance in bacterial population. As a prerequisite, the efficient analysis of single-cell data relies on the mathematical analysis of those models.

Structured branching processes allow for the study of populations, where the lifecycle of each cell is governed by a given characteristic or trait, such as the concentration of a specific protein inside the cell. The dependence of bacterial phenotypes like cell division times or ageing on such characteristics has been investigated by Aline Marguet using mathematical analysis of the underlying processes. To understand the long-time behavior of structured branching populations, the process describing the trait of a typical individual along its ancestral lineage, called auxiliary process [21] and its asymptotic behavior play a key role. In a publication in *ESAIM: Probability and Statistics* that appeared this year [20], we proved that the empirical measure of the structured branching process converges to the mean value of this auxiliary process. The approach relies on ergodicity arguments for the time-inhomogeneous auxiliary Markov process. The novelty compared to existing spectral methods is that our method allows to consider processes with time-varying rates for the modeling of changing environments. For example, we studied the case of a size-structured population in a varying environment and proved the convergence of the empirical measure in this specific case.

In collaboration with Charline Smadi (IRSTEA Grenoble), Aline Marguet also investigated the long-time behavior of a general class of branching Markov processes. This work, which has been submitted for publication [27], aims at understanding the link between the dynamic of the trait and the dynamic of the population. In the case of a trait modelling the proliferation of a parasite infection in a cellular population, we exhibit conditions on the dynamics of the parasites to survive in the population, despite the cellular divisions that dilute the number of parasites in each cell.

The study of the asymptotic behavior of general non-conservative semigroups is important for several aspects of branching processes, especially to prove the efficiency of statistical procedures. Vincent Bansaye from École Polytechnique, Bertrand Cloez from INRA Montpellier, Pierre Gabriel from Université Versailles Saint-Quentin, and Aline Marguet obtained necessary and sufficient conditions for uniform exponential contraction in weighted total variation norm of non-conservative semigroups. It ensures the existence of Perron eigenelements and provides quantitative estimates of spectral gaps, complementing Krein-Rutman theorems and generalizing recent results relying on probabilistic approaches. This work was submitted for publication this year [26].

6.4. Inference of gene expression parameters on lineage trees

As explained in the previous section, recent technological developments have made it possible to obtain time-course single-cell measurements of gene expression as well as the associated lineage information. However, most of the existing methods for the identification of mathematical models of gene expression are not well-suited to single-cell data and make the simplifying assumptions that cells in a population are independent, thus ignoring cell lineages. The development of statistical tools taking into account the correlations between individual cells will allow in particular for the investigation of inheritance of traits in bacterial populations.

In the framework of structured branching processes, we studied the statistical reconstruction of parameters. We considered the problem of estimating the division rate from the observations of the trait of the cells at birth. Previous works on the subject considered deterministic dynamics for the evolution of the trait. In collaboration with Marc Hoffmann (Université Paris Dauphine), Aline Marguet investigated the case of a trait evolving according to a diffusion process. The study of the asymptotic behavior of the tagged-chain, corresponding to the trait of a uniformly chosen individual, allowed us to prove the convergence of the empirical measure of the branching process, and the asymptotic minmax efficiency of nonparametric estimators for the density of the transition kernel and the invariant measure of the tagged-chain. For the estimation of the division rate, we proved in a parametric framework the asymptotic efficiency of a standard maximum likelihood proxy estimation. Finally, we demonstrate the validity of our approach on simulated datasets. The results of this work were published in *Stochastic Processes and their Applications* [17].

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], Eugenio Cinquemani, Marc Lavielle (XPOP, Inria Saclay–Île-de-France) and Aline Marguet developed a new model and a method for inference from data for gene expression along tree where the kinetic expression parameters are assumed to be inherited from the mother cell in an autoregressive way. This model generalizes the state-of-the-art mixed-effect models to the case of lineage trees. We implemented the inference procedure in Julia and proved that it provides unbiased estimates of the parameters. The application to the data of osmotic shock response by yeast show that the correlation between the parameter of a cell and its daughter is of 0.6 according to our model, leading to new biological questions such as the understanding of the origin of this inheritance. The results of this study were presented at the major bioinformatics conference ISMB/ECCB 2020 and published in the associated special issue of *Bioinformatics* [19].

6.5. Modeling and inference of RNA degradation

The ability to rapidly respond to changing nutrient availability is crucial for *E. coli* to survive in many environments including the gut. Reorganization of gene expression is the first step for bacteria to adjust their metabolism accordingly. It involves fine-tuning of both transcription and mRNA stability by dedicated regulatory interactions. While transcriptional regulation has been largely studied, the role of mRNA stability during a metabolic switch is poorly understood.

This question was addressed in the framework of the PhD thesis of Manon Morin funded by an INRA-Inria grant. Using combined genome-wide transcriptome and mRNA decay analyses, Manon Morin, Delphine Ropers and colleagues from the Toulouse Biotechnology Institute (ex-LISBP, INRA/INSA Toulouse) investigated the role of mRNA stability in the response of *E. coli* to nutrient changes. They demonstrated that transcript stability increases along metabolic transitions representative of the carbon source fluctuations, the glucose-acetate-starvation transition [9], [10]. Most of the stabilization occurs at glucose-acetate transition when glucose is exhausted. Stabilized mRNAs remain stable during acetate consumption and carbon starvation. Meanwhile, expression of most genes is downregulated. Metabolic control analysis showed that most of gene expression regulation is driven by changes in transcription. Post-transcriptional regulations appear to be important for genes involved in bacterial response to nutrient starvation. These results have been further developed in a paper recently submitted to a biology journal.

The observation of a global stabilization of cellular mRNAs during adaptation to carbon source depletion raises questions about the regulatory mechanisms at work. Known regulators of mRNA stability such as the protein Hfq, the carbon storage regulator Csr, and several small regulatory RNAs, specifically target mRNAs. Are these regulatory mechanisms sufficient to explain the systematic adjustment of mRNA half-lives? The collaboration with Muriel Cocaign-Bousquet and colleagues from the Toulouse Biotechnology Institute has been pursued to answer these questions, in the context of the PhD thesis of Thibault Etienne, funded by an INRA-Inria PhD grant. The objective is to develop models able to explain how cells coordinate their physiology and the functioning of the degradation machinery following environmental changes. In a paper submitted this year, Thibault Etienne, Delphine Ropers and Muriel Cocaign-Bousquet investigate the possibility that competition between mRNAs for their binding to the degradation machinery is an important mechanism for the regulation of mRNA half-lives. They develop a mathematical model of mRNA degradation and assess the role of competitive effects on mRNA degradation kinetics by numerical simulation and sensitivity analysis. Competition appears to globally increase the stability of cellular mRNAs and to amplify the effect of post-transcriptional regulation. In a follow-up study, the model is currently being used to interpret large data sets corresponding to the degradation kinetics of 4254 mRNAs in *E. coli* cells growing in four different environmental conditions.

6.6. 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 in combination with methods from optimal control theory. In a paper published in the *Journal of Mathematical Biology* this year [24], 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, were presented at the European Control Conference (ECC 2019) in Naples this year and published in the proceedings [25].

6.7. Bacterial growth inhibition by acetate

High concentrations of organic acids such as acetate inhibit growth of *Escherichia coli* and other bacteria. This phenomenon is of interest for understanding bacterial physiology but is also of practical relevance. Growth inhibition by organic acids underlies food preservation and causes problems during high-density fermentation in biotechnology. The development of new approaches for the relief of growth inhibition by acetate during high-density fermentation of *E. coli* is one of the motivating assumptions for the work of IBIS in the IPL project COSY (Sections 7.2 and 6.8 below).

What causes growth inhibition by acetate? Classical explanations invoke the uncoupling effect of acetate and the establishment of an anion imbalance. During his PhD thesis, Stéphane Pinhal investigated an alternative hypothesis: the perturbation of acetate metabolism due to the inflow of excess acetate. In an experimental and modelling study published in the *Journal of Bacteriology* [23], Stéphane Pinhal, Delphine Ropers, Hans Geiselmann, and Hidde de Jong developed a set of isogenic strains that remove different parts of the metabolic network involved in acetate metabolism. Analysis of these strains revealed that the inflow of acetate accounts for 20% of the growth-inhibitory effect through a modification of the acetyl phosphate concentration. While the study does not provide a definite answer to the question of what accounts for the remaining 80% of the reduction in growth rate, some of the observations argue against a prominent role of uncoupling in growth inhibition by acetate in the conditions tested.

6.8. Modeling synthetic microbial communities for improving productivity

Modelling, analysis and control of microbial community dynamics is a fast-developing subject with great potential implications in the understanding of natural processes and the enhancement of biootechnological processes. Within the IPL COSY (Section 7.2), we picked up the challenge to design and investigate the dynamics of synthetically engineered microbial communities with a consortium of Inria partners. In IBIS, in particular, we are addressing the design of a bacterial community of two *E.coli strains*, mimicking mutualistic relationships found in nature, and with the potential to outperform a single producer strain in the production of a heterologous protein. During the post-doctoral stay of Marco Mauri, we developed an ODE model of the key growth phenotypes of the community and their interactions, calibrated the model on literature data, and analysed the model for an in-depth understanding of the conditions supporting coexistence and of the tradeoffs encountered in this production process. The results are presented in a paper submitted for publication this year and will be tested experimentally in the framework of the recently-started PhD project of Maaike Sangster. Analysis of optimal community control problems as well as design and deployment of optimal control strategies will follow in synergy with other IPL COSY partners.

6.9. Detection of small non-coding RNAs

Small non-coding RNAs (sRNAs) regulate numerous cellular processes in all domains of life. Several approaches have been developed to identify them from RNA-seq data, which are efficient for eukaryotic sRNAs but remain inaccurate for the longer and highly structured bacterial sRNAs. Together with colleagues from INSA de Lyon, Stéphan Lacour developed APERO, a new algorithm to detect small transcripts from paired-end bacterial RNA-seq data. This algorithm is based on a novel approach, which does not start from the read coverage distribution, but analyzes boundaries of individual sequenced fragments to infer the 5´ and 3´ ends of all transcripts. Validation of the algorithm on *Escherichia coli* and *Salmonella enterica* datasets, based on experimentally validated sRNAs, showed it to outperform all existing methods in terms of sRNA detection and boundary precision. Moreover, APERO was able to identify the small transcript repertoire of *Dickeya dadantii* including putative intergenic RNAs, 5´ UTR or 3´ UTR-derived RNA products and antisense RNAs. This work was published in *Nucleic Acids Research* this year [18]. APERO is freely available as an open source R package (https://github.com/Simon-Leonard/APERO). In other work, together with colleagues from the University of Salento, Lecce (Italy), Stéphan Lacour contributed to RhoTermPredict, an algorithm for predicting Rho-dependent transcription terminators in bacterial genomes [16].

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	G. Batt
IBIS participants	E. Cinquemani, A. Marguet, D. Ropers
Type	ANR project (2016-2020)

Project name	ENZINVIVO – Détermination in vivo des paramètres enzymatiques dans une voie métabolique synthétique
Coordinator	G. Truan
IBIS participants	J. Geiselmann, H. de Jong
Type	ANR project (2016-2020)

Project name	MAXIMIC: Optimal control of microbial cells by natural and synthetic strategies
Coordinator	H. de Jong
IBIS participants	C. Boyat, E. Cinquemani, J. Geiselmann, H. de Jong, A. Pavlou, C. Pinel, D. Ropers
Type	ANR project (2017-2021)
Web page	https://project.inria.fr/maximic

Project name	RIBECO (RIBonucleotide ECOnomy): Engineering RNA life cycle to optimize economy of microbial energy
Coordinator	M. Cocaign-Bousquet
IBIS participants	E. Cinquemani, T. Etienne, D. Ropers
Type	ANR project (2018-2022)
Web page	https://project.inria.fr/ribeco/

Project name	COSY: real-time COntrol of SYnthetic microbial communities
Coordinator	E. Cinquemani
IBIS participants	E. Cinquemani, H. de Jong, J. Geiselmann, M. Mauri, T. Muszbek, C. Pinel, D. Ropers, M. Sangster
Type	Inria Project Lab (2017-2021)
Web page	https://project.inria.fr/iplcosy/

Project name	OPTICO : OPTImal COntrol software for microbial communities in a system of minibioreactors
Coordinator	E. Cinquemani
IBIS participants	E. Cinquemani, H. de Jong, J. Geiselmann, T. Muszbek
Type	Inria ADT (2019-2021)

Project name	AlgeaInSilico: Prédire et optimiser la productivité des microalgues en fonction de leur milieu de croissance
Coordinator	O. Bernard
IBIS participants	H. de Jong
Type	Inria Project Lab (2015-2019)
Web page	https://project.inria.fr/iplalgaeinsilico/

Project name	Analyse intégrative de la coordination entre stabilité des ARNm et physiologie cellulaire chez Escherichia coli
Coordinators	D. Ropers, M. Cocaign-Bousquet (Inra, LISBP)
IBIS participants	T. Etienne, D. Ropers
Type	Contrat Jeune Scientifique Inra-Inria (2016-2019)

7.3. International Research Visitors

7.3.1. Visits of International Scientists

Tomas Gedeon, professor in Mathematics at Montana State University (USA), visited the IBIS project-team during two months (May-July 2019) to work on modeling and analysis of resource allocation in microorganisms. His stay at Inria was funded by the Visiting researcher program of the research center Grenoble - Rhône-Alpes.

7.3.1.1. Internships

Emmanouil Sideris, enrolled in the MSc program in Computer Science at the University of Patras (Greece), did a Master internship with Eugenio Cinquemani.

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
Aline Marguet	Biohasard: Stochastic models for biology, Grenoble	Jun 2020
Delphine Ropers	Séminaire de Modélisation du Vivant (SeMoVi), Lyon and Grenoble	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 and 2020)	Associate editor

8.1.2.2. Member of conference program committees

IBIS member	Conference, workshop, program
Eugenio Cinquemani	ECC 2019 and 2020, CMSB 2018 and 2019, HSB 2019 and 2020, SASB 2019, IEEE CIBCB 2019
Hidde de Jong	CMSB 2019, FOSBE 2019, HSB 2019 and 2020
Delphine Ropers	CSBio 2019

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
Hidde de Jong	International Human Frontier Science Program (HFSP)	Member grant selection committee
Delphine Ropers	INRA-Inria	Member selection committee PhD grants
Delphine Ropers	Elixir	Member of Microbial Biotechnology Community

8.1.5. Invited talks and other presentations

Eugenio Cinquemani

Title	Event and location	Date
Estimation and control of microbial dynamics: From single-cell to microbial communities	Presentation at Inria-KAIST workshop on Applied mathematics, Paris	Mar 2019
Optimal control of bacterial growth for metabolite production: The role of timing and costs of control	Presentation at ECC 2019, Naples (Italy)	Jun 2019
Enhanced production of heterologous proteins by a synthetic microbial consortium: Conditions and tradeoffs	Invited talk at Workshop Advanced Physics for Medicine, CNR, Rome (Italy)	Sep 2019

Hidde de Jong

Title	Event and location	Date
Integrated models of the cell: metabolism, gene expression, signalling	Tutorial at CompSysBio: Advanced Lecture Course on Computational Systems Biology, Aussois	Apr 2019
Optimal control of bacterial growth: biotechnological applications	Invited talk at Workshop New Vistas in Computational Systems and Synthetic Biology, Vigo (Spain)	May 2019
Optimal control of microbial growth and bioproduction	Invited talk at SYMER workshop, Global Challenge Week, Univ Grenoble Alpes, Grenoble	Jun 2019
Reengineering bacterial metabolism using synthetic biology and optimal control theory	Invited talk in special session on Predictive approaches for biological systems engineering, JOBIM 2019, Nantes	Jul 2019
Production of heterologous proteins by a synthetic microbial consortium: conditions and tradeoffs	Invited talk at Biocore seminar, Peyresq	Sep 2019

Stéphan Lacour

Title	Event and location	Date
New insights into the regulation of curli production by E. coli	BactoGre workshop, Grenoble	Avr 2019

Aline Marguet

Title	Event and location	Date
Inheritance and variability of kinetic gene expression parameters in microbial cells: Modelling and inference from lineage tree data	Invited talk at École de printemps de la chaire MMB, Aussois	May 2019
Inheritance and variability of kinetic gene expression parameters in microbial cells: Modelling and inference from lineage tree data	Presentation at ISMB/ECCB conference, Basel (Switzerland)	Jul 2019
Inheritance and variability of kinetic gene expression parameters in microbial cells: Modelling and inference from lineage tree data	Invited talk at Biohasard 2019 workshop, Rennes	Aug 2019
Inheritance and variability of kinetic gene expression parameters in microbial cells: Modelling and inference from lineage tree data	Presentation at Journée Inria-Bio, Lyon	Oct 2019
Inheritance and variability of kinetic gene expression parameters in microbial cells: Modelling and inference from lineage tree data	Seminar Master Mathématiques pour les Sciences du Vivant, Ecole Polytechnique, Palaiseau	Oct 2019
Modelling and statistics of branching processes	Workshop on Growth and division in mathematics and medicine, London (UK)	Nov 2019
Inheritance and variability of kinetic gene expression parameters in microbial cells: Modelling and inference from lineage tree data	Seminar équipe-projet DRACULA, Lyon	Nov 2019
Inheritance and variability of kinetic gene expression parameters in microbial cells: Modelling and inference from lineage tree data	Seminar unité MaIage, Jouy-en-Josas	Dec 2019
Long time behavior of a general class of branching Markov processes	Seminar Institut Fourier, Grenoble	Dec 2019

Marco Mauri

Title	Event and location	Date

Thibault Etienne

Title	Event and location	Date
Regulatory mechanisms underlying coordination of mRNA degradation and cell physiology in Escherichia coli	Seminar during Journée annuelle du LIPhy, Grenoble	Jun 2019
Large scale modelling of mRNA decay in Escherichia coli	Seminar during Journée InriaBio, ENS de Lyon	Oct 2019

Antrea Pavlou

Title	Event and location	Date
Experimental and computational analysis of bacterial self-replicators	Poster at CompSysBio: Advanced Lecture Course on Computational Systems Biology, Aussois	Apr 2019
Experimental and computational analysis of bacterial self-replicators	Seminar for GdT MathBio, Grenoble	Nov 2019

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 Grenoble - Rhône-Alpes	Référente chercheurs
Delphine Ropers	Inria Grenoble - Rhône-Alpes	Co-coordinator of mentoring program

8.1.7. Recruitment committees

IBIS member	Organism	Recruitment
Johannes Geiselmann	University of Barcelona	Assistant professor
Delphine Ropers	Inria Bordeaux Sud-Ouest	Chargés de recherche (jury d'admissibilité)
Delphine Ropers	Inria national	Chargés de recherche (jury d'admissibilité)
Delphine Ropers	Univ de Rennes	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 (4 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)

Aline Marguet

Practicals: Biostatistics, M2, Univ Grenoble Alpes (24 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)

Thibault Etienne

Course: Construction et analyse de plasmides in silico, M1, Master ingéniérie de la santé, Univ Grenoble Alpes (4 h)

Antrea Pavlou

Practicals: Cellular biology, L1, Chemistry and biochemistry, Univ Grenoble Alpes (28 h)

8.2.2. Supervision

HdR: **Eugenio Cinquemani**, Identification, estimation and control of gene expression and metabolic network dynamics [14], Univ Grenoble Alpes, Nov 2019

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

PhD in progress: **Maaike Sangster**, Development, characterization and control of E. coli communities on an automated experimental platform. Supervisors: Eugenio Cinquemani and Johannes Geiselmann

8.2.3. Juries

PhD thesis committees

IBIS member	Role	PhD student	University	Date
Johannes Geiselmann	Président	Vanni Petrolli	Univ Grenoble Alpes	Nov 2019
Hidde de Jong	Rapporteur	Pauline Trébulle	AgroParisTech	Oct 2019
Hidde de Jong	Rapporteur	Firas Hammami	Univ d'Aix-Marseille	Dec 2019
Delphine Ropers	Examinateur	Marianyela Petrizzelli	Univ Paris Sud	Jul 2019
Delphine Ropers	Examinateur	Ronan Duchesne	Univ de Lyon	Dec 2019

Habilitation (HDR) committees

IBIS member	Role	HDR candidate	University	Date
Johannes Geiselmann	Président	Marianne Weidenhaupt	Univ Grenoble Alpes	Jan 2019
Hidde de Jong	Président	Arnaud Tonnelier	Univ Grenoble Alpes	Jun 2019
Hidde de Jong	Rapporteur	Stefan Müller	University of Vienna	Sep 2019
Hidde de Jong	Examinateur	Eugenio Cinquemani	Univ Grenoble Alpes	Nov 2019

PhD advisory committees

IBIS member	PhD student	University
Eugenio Cinquemani	Arthur Carcano	Institut Pasteur/Inria
Eugenio Cinquemani	Alexey Koshkin	ENS de Lyon
Johannes Geiselmann	Alain Lombard	Univ Grenoble Alpes
Johannes Geiselmann	Shiny Martis	INSA de Lyon
Hidde de Jong	Charlotte Coton	INRA Moulon
Hidde de Jong	Kapil Newar	Univ Grenoble Alpes
Hidde de Jong	Agustín Yabo	Univ Nice–Sophia-Antipolis
Stéphan Lacour	Julien Trouillon	CEA Grenoble
Stéphan Lacour	Camille Brunet	CHU Grenoble
Stéphan Lacour	Raphael Fourquet	INSA de Lyon/Univ de Lyon
Stéphan Lacour	Kevin Pick	Univ de Lyon
Delphine Ropers	Manon Barthe	Univ Paul Sabatier, Toulouse
Delphine Ropers	Charlotte Roux	Univ Paul Sabatier, 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. Interventions

Delphine Ropers participated in the colloquium FIGAS 2019 - Un rêve pour les filles et les garçons : la Science (Nov 2019). She was moderator of the session "Stéréotypes dans le numérique", and co-presented with Thierry Vieville "Mais comment éduquer les garçons à l'équité des genres au niveau informatique et numérique ?".

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^{o} 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^{o} 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^{o} 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] A. MARGUET, M. LAVIELLE, E. CINQUEMANI.*Inheritance and variability of kinetic gene expression parameters in microbial cells: modeling and inference from lineage tree data*, in "Bioinformatics", 2019, vol. 35, n^{o} 14, p. i586–i595 [*DOI :* 10.1093/BIOINFORMATICS/BTZ378], https://hal.archives-ouvertes.fr/hal-02317115

[9] 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

[10] 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

[11] 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 $\sigma(S)$ 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

[12] 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

[13] 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

Doctoral Dissertations and Habilitation Theses

[14] E. CINQUEMANI.*Identification, estimation and control of gene expression and metabolic network dynamics*, Université Grenoble-Alpes, ED MSTII, November 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02424024

Articles in International Peer-Reviewed Journal

[15] E. CINQUEMANI.*Stochastic reaction networks with input processes: Analysis and application to gene expression inference*, in "Automatica", 2019, vol. 101, p. 150-156 [*DOI :* 10.1016/J.AUTOMATICA.2018.11.047], https://hal.inria.fr/hal-01925923

[16] M. DI SALVO, S. PUCCIO, C. PEANO, S. LACOUR, P. ALIFANO.*RhoTermPredict: an algorithm for predicting Rho-dependent transcription terminators based on Escherichia coli, Bacillus subtilis and Salmonella enterica databases*, in "BMC Bioinformatics", December 2019, vol. 20, n^O^ 1, p. 1-11 [*DOI :* 10.1186/S12859-019-2704-X], https://hal.inria.fr/hal-02423208

[17] M. HOFFMANN, A. MARGUET.*Statistical estimation in a randomly structured branching population*, in "Stochastic Processes and their Applications", 2019, vol. 129, n^O^ 12, p. 5236-5277 [*DOI :* 10.1016/J.SPA.2019.02.015], https://hal.archives-ouvertes.fr/hal-01662203

[18] S. LEONARD, S. MEYER, S. LACOUR, W. NASSER, F. HOMMAIS, S. REVERCHON.*APERO: a genome-wide approach for identifying bacterial small RNAs from RNA-Seq data*, in "Nucleic Acids Research", May 2019, vol. 47, n^O^ 15, p. 1-12 [*DOI :* 10.1093/NAR/GKZ485], https://hal.archives-ouvertes.fr/hal-02152104

[19] A. MARGUET, M. LAVIELLE, E. CINQUEMANI.*Inheritance and variability of kinetic gene expression parameters in microbial cells: modeling and inference from lineage tree data*, in "Bioinformatics", 2019, vol. 35, n^O^ 14, p. i586-i595 [*DOI :* 10.1093/BIOINFORMATICS/BTZ378], https://hal.archives-ouvertes.fr/hal-02317115

[20] A. MARGUET.*A law of large numbers for branching Markov processes by the ergodicity of ancestral lineages*, in "ESAIM: Probability and Statistics", 2019, p. 1-23, https://arxiv.org/abs/1707.07993 , forthcoming [*DOI :* 10.1051/PS/2018029], https://hal.archives-ouvertes.fr/hal-01567317

[21] A. MARGUET.*Uniform sampling in a structured branching population*, in "Bernoulli", 2019, vol. 25, n^O^ 4A, p. 2649-2695, https://arxiv.org/abs/1609.05678 [*DOI :* 10.3150/18-BEJ1066], https://hal.archives-ouvertes.fr/hal-01362366

[22] Y. MARTIN, M. PAGE, C. BLANCHET, H. DE JONG.*WellInverter: a web application for the analysis of fluorescent reporter gene data*, in "BMC Bioinformatics", December 2019, vol. 20, n^O^ 1, 309 [*DOI :* 10.1186/S12859-019-2920-4], https://hal.inria.fr/hal-02195461

[23] S. PINHAL, D. ROPERS, J. GEISELMANN, H. DE JONG.*Acetate metabolism and the inhibition of bacterial growth by acetate*, in "Journal of Bacteriology", July 2019, vol. 201, n^{o} 13, p. 147 - 166 [*DOI :* 10.1128/JB.00147-19], https://hal.inria.fr/hal-02195459

[24] 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", 2019, vol. 78, n^{o} 4, p. 985–1032 [*DOI :* 10.1007/s00285-018-1299-6], https://hal.inria.fr/hal-01929475

International Conferences with Proceedings

[25] E. CINQUEMANI, F. MAIRET, I. YEGOROV, H. DE JONG, J.-L. GOUZÉ.*Optimal control of bacterial growth for metabolite production: The role of timing and costs of control*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, IEEE, June 2019, p. 2657-2662 [*DOI :* 10.23919/ECC.2019.8796079], https://hal.inria.fr/hal-02384034

Other Publications

[26] V. BANSAYE, B. CLOEZ, P. GABRIEL, A. MARGUET.*A Non-Conservative Harris' Ergodic Theorem*, March 2019, https://arxiv.org/abs/1903.03946 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02062882

[27] A. MARGUET, C. SMADI.*Long Time Behaviour of a General Class of Branching Markov Processes*, August 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02274642

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

Remi Ronfard [Team leader, Inria, Senior Researcher, HDR]
Melina Skouras [Inria, Researcher]

Faculty Members

Stefanie Hahmann [Institut polytechnique de Grenoble, Professor, HDR]
Jean-Claude Léon [Institut polytechnique de Grenoble, Professor, until Apr 2019, HDR]
Olivier Palombi [Univ Grenoble Alpes, Associate Professor, HDR]

Technical Staff

Remi Colin de Verdiere [Inria, Engineer]
Valerian Daunis [Inria, Engineer, from Mar 2019 until Apr 2019]
Julien Daval [Inria, Engineer, until Apr 2019]

PhD Students

Amelie Fondevilla [Univ. Grenoble Alpes, PhD student, until Nov 2019]
Thomas Buffet [Inria, PhD Student]

Pierre Casati [Inria, PhD Student]
Qianqian Fu [Univ Grenoble Alpes, PhD Student]
Geoffrey Guingo [Univ de Strasbourg, PhD Student, until Aug 2019]
Youna Le Vaou [Cifre Groupe PSA, PhD Student]
Vaishnavi Ameya Murukutla [Univ Grenoble Alpes, PhD Student]

Post-Doctoral Fellow

Musaab Khalid Osman Mohammed [Institut polytechnique de Grenoble, Post-Doctoral Fellow, until Apr 2019]

Administrative Assistant

Marion Ponsot [Inria, Administrative Assistant]

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. Target applications

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 7.

4.2. Visual arts

- Sculpture.
- Modeling and animation for 3D films and games.
- Virtual cinematography and tools for theater directors.

4.3. Engineering

- Industrial design.
- Mechanical & civil engineering.

4.4. Natural sciences

- Geology.
- Virtual functional anatomy.

4.5. Education and creative tools

- Sketch-based teaching.
- Creative environments for novice users.
- Museography.

5. Highlights of the Year

5.1. Highlights of the Year

Maxime Garcia, Amélie Fondevilla and Geoffrey Guingo defended their PhD theses.

We published two papers [16], [20] at ACM Transaction on Graphics (Proceedings of SIGGRAPH) which is the most prestigious and selective conference in computer graphics.

5.1.1. Awards

Mélina Skouras was elected Eurographics Junior Fellow in May 2019.
Stefanie Hahmann was elected SMA Fellow (Solid Modeling Association) in June 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. 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. 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.

6.3.2. 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.

7. New Results

7.1. Star-Shaped Metrics for Mechanical Metamaterial Design

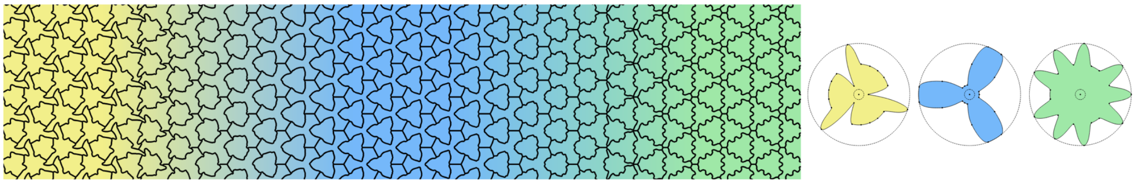


Figure 1. Our method generates a smoothly-graded pattern (left) when interpolating between three star-shaped distance functions (right) on a regular honeycomb lattice. Each distance function is compactly parameterized with polar coordinates, allowing for simple interpolation in metric space as indicated by color-coding.

We present a method for designing mechanical metamaterials based on the novel concept of Voronoi diagrams induced by star-shaped metrics. As one of its central advantages, our approach supports interpolation between arbitrary metrics, as depicted in Figure 1. This capability opens up a rich space of structures with interesting aesthetics and a wide range of mechanical properties, including isotropic, tetragonal, orthotropic, as well as smoothly graded materials. We evaluate our method by creating large sets of example structures, provided as accompanying material. We validate the mechanical properties predicted by simulation through tensile tests on a set of physical prototypes.

7.2. Computational Design of Fabric Formwork

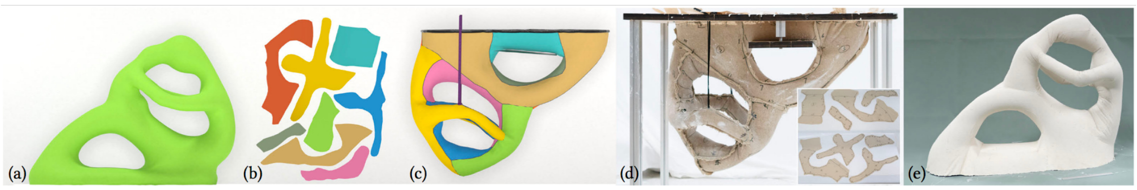


Figure 2. A fertility model designed and fabricated using our computational approach. For a target 3D model (a), our system can automatically compute a set of flat panels (b) that can be sewn together to serve as fabric containers to form a target shape by pressure of liquid plaster poured in – see (c) for the simulation under force equilibrium of membrane tension, liquid pressure and external supports. The generated flat panels are used to conduct the physical fabrication of fabric formwork (d). After drying and unwrapping the fabric container, a sculpture with the designed target shape has been fabricated (e).

This work (illustrated in Figure 2) presents an inverse design tool for fabric formwork - a process where flat panels are sewn together to form a fabric container for casting a plaster sculpture. Compared to 3D printing techniques, the benefit of fabric formwork is its properties of low-cost and easy transport. The process of fabric formwork is akin to molding and casting but having a soft boundary. Deformation of the fabric container is governed by force equilibrium between the pressure forces from liquid fill and tension in the stretched fabric. The final result of fabrication depends on the shapes of the flat panels, the fabrication orientation and the placement of external supports. Our computational framework generates optimized flat panels and fabrication orientation with reference to a target shape, and determines effective locations for external supports. We demonstrate the function of this design tool on a variety of models with different shapes and topology. Physical fabrication is also demonstrated to validate our approach.

7.3. Spatial Motion Doodles

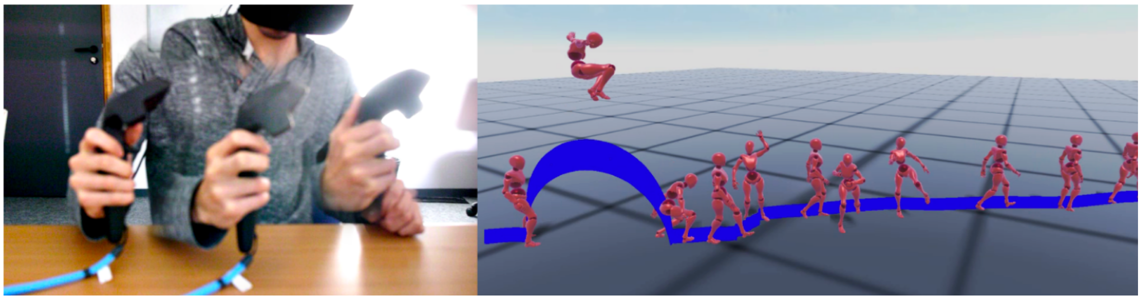


Figure 3. Left: A user drawing a spatial motion doodle (SMD) which is the six-dimensional trajectory of a moving frame (position and orientation), here attached to the HTC Vive controller. Right: The SMD is parsed into a string of motion tokens,allowing to recognize actions and extract the associated motion qualities. This information is transferred to an articulated character to generate an expressive 3D animation sequence.

We present a method for easily drafting expressive character animation by playing with instrumented rigid objects (see Figure 3). We parse the input 6D trajectories (position and orientation over time) – called spatial motion doodles – into sequences of actions and convert them into detailed character animations using a dataset of parameterized motion clips which are automatically fitted to the doodles in terms of global trajectory and timing. Moreover, we capture the expressiveness of user-manipulation by analyzing Laban effort qualities in the input spatial motion doodles and transferring them to the synthetic motions we generate. We validate the ease of use of our system and the expressiveness of the resulting animations through a series of user studies, showing the interest of our approach for interactive digital storytelling applications dedicated to children and non-expert users, as well as for providing fast drafting tools for animators.

7.4. Text-to-Movie Authoring of Anatomy Lessons

With popular use of multimedia and 3D content in anatomy teaching there is a need for a simple yet comprehensive tool to create and edit pedagogical anatomy video lessons. This work introduces an automated video authoring tool (shown in Figure 4) created for teachers. It takes text written in a novel domain specific language (DSL) called the Anatomy Storyboard Language (ASL) as input and translates it to real time 3D animation. Preliminary results demonstrates the ease of use and effectiveness of the tool for quickly drafting video lessons in realistic medical anatomy teaching scenarios.

7.5. Approximate Reconstruction of 3D Scenes From Bas-Reliefs

For thousands of years, bas-reliefs such as the one depicted in Figure 5 have been used to depict scenes of everyday life, mythology and historic events. Yet, the precise geometry of those scenes remains difficult to interpret and reconstruct. Over the past decade, methods have been developed for generating bas-reliefs from 3D scenes. With this work, we investigate the inverse problem of interpreting and reconstructing 3D scenes from their bas-relief depictions. Even approximate reconstructions can be useful for art historians and museum exhibit designers, as a first entry to the complete interpretation of the narratives told in stone or marble. To create such approximate reconstructions, we present methods for extracting 3D base mesh models of all characters depicted in a bas-relief. We take advantages of the bas-relief geometry and high-level knowledge of human body proportions to recover body parts and their three-dimensional structure, even in severe cases of contact and occlusion. We present experimental results for 6 bas-relief depictions of Greek mythological and historical scenes involving 18 characters and draw conclusions for future work.

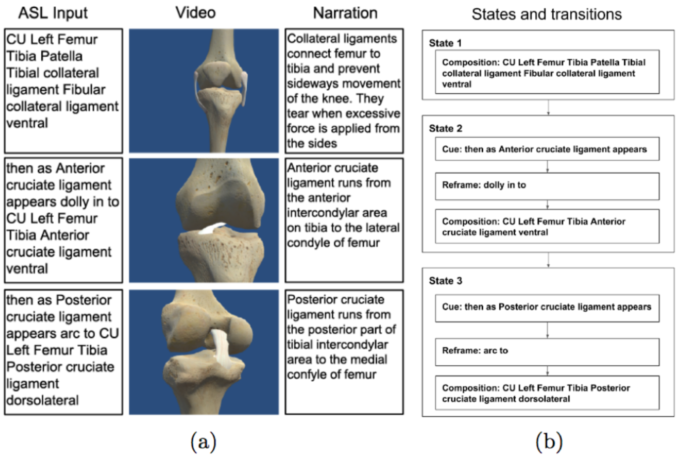


Figure 4. Text-to-movie generation example with hierarchical finite state machines representation.

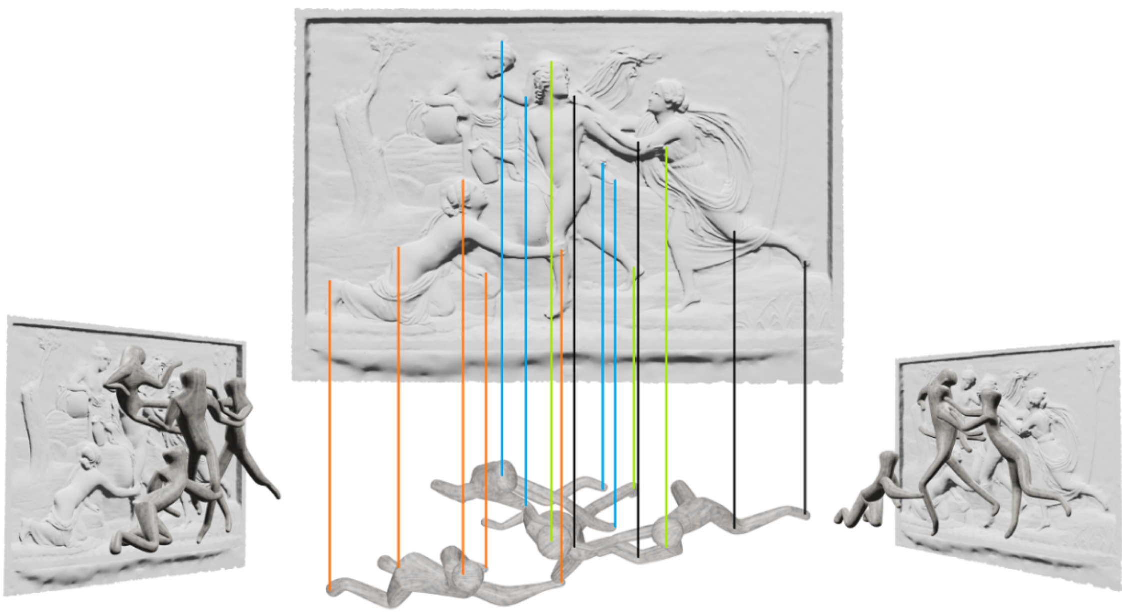


Figure 5. 3D interpretation of the mythological story of Hylas and the Water Nymphs, after a bas-relief marble by Bertel Thorvaldsen (1833). Hylas was sent to fetch water for the camp. Finding a pool in a clearing, he was encircled by water nymphs reaching up to kiss him and there disappeared with them forever. Using hand-drawn silhouette shapes and 2D skeletons of the four characters, we compute a plausible 3D reconstruction of the scene with rigged and skinned models suitable for 3D animation.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts 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, IMAGINE 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 entirely re-designed for supporting ultra high definition video. 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 was also designed, tested and evaluated with the Litt&Arts team at Univ. Grenoble Alpes, as part of CDP project Performance Lab.

9.2.2. 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), was a follow-up and a generalization of Dynam’it and Collodi 1. The goal was 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 was 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. Three short animations were created for this purpose by a professional animator from film examples of dancers (Gene Kelly in "Singing in the rain", Fred Astaire and Cyd Charisse in "The band wagon"). Those examples demonstrate that sketch-based animation can be used to create complex character animation even in very challenging situations. Those results were presented during the two final reviews of the COLLODI2 project in Valence and Paris in December 2019 and published as a research report.

9.2.3. 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.

9.2.4. 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.

9.2.5. 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.

9.2.6. 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

Rémi Ronfard is a member of the selection committee for Inria-MCC (Ministry of Culture and Communication) activities; and a member of the steering committee for the Eurographics workshop on intelligent cinematography and editing (WICED).

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

Remi Ronfard and Julie Valéro (Litt & Arts, Univ. Grenoble Alpes) are chairing the first national meeting on computer theater (Journées d'Informatique Théâtrale), which will take place in Grenoble in February 2020.

Stefanie Hahmann was conference general co-chair for the ACM Symposium on Solid and Physical Modeling (SPM), Vancouver 2019.

10.1.1.2. Member of the Organizing Committees

Stefanie Hahmann was member of the organization committee of the SIAM Conference on Geometric Design, Vancouver 2019.

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

Mélina Skouras was a member of the Program Committees for ACM Siggraph Asia 2019 and Eurographics 2019.

Stefanie Hahmann was a member of the Program Committees for Symposium on Geometry Processing (SGP) 2019.

Rémi Ronfard was a member of the Program Committees for Motion, Interaction and Games (MIG) 2019, Intelligent Cinematography and Editing (WICED) 2019, and EXPRESSIVE 2019.

10.1.2.2. Reviewer

Rémi Ronfard was a reviewer for Siggraph 2019 and Siggraph Asia 2019.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Stefanie Hahmann is an Associate Editor of CAG (Computers & Graphics, Elsevier) and CAD (Computer Aided Design, Elsevier).

10.1.3.2. Reviewer - Reviewing Activities

Mélina Skouras was a reviewer for ACM Siggraph 2019, ACM Transactions on Graphics (TOG) and the International Journal for Numerical Methods in Engineering.

Stefanie Hahmann was a reviewer for Graphical Models (GMOD), Computers & Graphics (CAG) and Computer Aided Design (CAD).

Rémi Ronfard was a reviewer for ACM transactions on Graphics (TOG).

10.1.4. Invited Talks

Mélina Skouras gave an invited talk at the ACM/Eurographics Symposium on Computer Animation in July 2019.

Stefanie Hahmann gave an invited talk at the ARCADES workshop (Algebraic Representations in Computer-Aided Design for complEx Shapes) in Vienna, Austria, in November 2019.

10.1.5. Leadership within the Scientific Community

Remi Ronfard and Julie Valéro (Litt & Arts, Univ. Grenoble Alpes) are animating 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.

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.

10.1.6. Scientific Expertise

Stefanie Hahmann was a member af the Advisory Board of the European ITN-ETN Marie-Curie Training Network ARCADES on Algebraic Representations for Computer-Aided Design of Complex Shapes from 2015 until 2019.

Stefanie Hahmann was a member of the selection committees for the SMA Young Investigator Award, Solid Modeling Association.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Bachelor : Stefanie Hahmann, Numerical Methods, 42 HETD, 240 students, L3, Ensimag-Grenoble INP.

Master : Stefanie Hahmann, Geometric Modeling, 47 HETD, 60 students, M1, Ensimag-Grenoble INP.

Master : Stefanie Hahmann, Surface Modeling, 37 HETD, 30 students, M2, Ensimag-Grenoble INP.

Master: Mélina Skouras, Surface modeling, 14.5 HETD, M2, Ensimag, Grenoble, France

Master: Mélina Skouras, Numerical Mechanics, 8 HETD, M2, ENS de Lyon, France

Master: Rémi Ronfard teaches computer animation to MOSIG M2 students, 18 HETD, Grenoble INP, Univ. Grenoble Alpes.

PhD: Rémi Ronfard is an associate researcher in the Spatial Media team at ENSADLAB, where he teaches computer graphics to doctoral art students in the SACRE doctoral school, 60 HETD, Univ. Paris Sciences et Lettres (PSL).

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 20 Masters (PFE) thesis defences in 2019.

10.2.2. Supervision

PhD: Geoffrey Guingo, Synthesis of animated textures, supervised by Marie-Paule Cani, Jean-Michel Dischler and Basile Sauvage, was defended on December 3, 2019.

PhD: Amélie Fondevilla, Modélisation et animation de surfaces développables, supervised by Stefanie Hahmann and Damien Rohmer, was defended on Decembre 18, 2019.

PhD: Maxime Garcia, Animation transfer: character animation by playing and acting, since October 2016, supervised by Rémi Ronfard, was defended on December 19, 2019.

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 : 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 : Nachwa Aboubakr, Observation and modeling of human activities, since October 2016, supervised by James Crowley and Rémi Ronfard.

PhD in progress: David Jourdan, Design of free-from surfaces using self-actuated materials, 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: Mélanie Carrière, 3D Design of Ancient Garments, June 2019, supervised by Mélina Skouras and Stefanie Hahmann.

Master's thesis: Shalu Dwivedi, Simulation of Laser-Cut Metamaterials, August 2019, supervised by Mélina Skouras.

Master thesis: Paul-Elian Tabarant, student at Telecom ParisTech, 3D sketching in virtual reality, July 2019, subervised by Stefanie Hahmann and Georges-Pierre Bonneau.

M1 Internship: Marion Taconné, Extensions to the aeroMorph Project, September 2019, supervised by Mélina Skouras.

M1 IRL: Manon Valle, Study of different norms for modeling inverse problems, May 2019, supervised by Mélina Skouras.

M1 IRL: Nathan Shourick, Existence of solutions to design inverse problems, May 2019, supervised by Mélina Skouras, Florence Bertails-Descoubes and Mickaël Ly.

10.2.3. Juries

Stefanie Hahmann was part of the jury CRCN 2019 at Inria Rennes, April 2019.

Stefanie Hahmann was member of the HDR committee a reviewer of the HDR Thesis of Alexandra Bac at Univ. Aix-Marseille.

Stefanie Hahmann also was a reviewer of 2 PhD theses, Université de Strasbourg (Cedric Bobenrieth) and Université de Bourgogne (Lucas Morlet) in 2019.

Stefanie Hahmann was an external reviewer of the PhD thesis of Alessandro Marro at University of Kaiserslautern, Germany.

11. Bibliography

Major publications by the team in recent years

[1] G. CORDONNIER, E. GALIN, J. GAIN, B. BENES, E. GUÉRIN, A. PEYTAVIE, M.-P. CANI.*Authoring Landscapes by Combining Ecosystem and Terrain Erosion Simulation*, in "ACM Transactions on Graphics", July 2017, vol. 36, n^{o} 4, 134, The paper was presented at Siggraph 2017 [*DOI :* 10.1145/3072959.3073667], https://hal.archives-ouvertes.fr/hal-01518967

[2] B. DALSTEIN, R. RONFARD, M. VAN DE PANNE. *Vector Graphics Complexes*, in "ACM Transactions on Graphics", July 2014, vol. 33, n^o 4, Article No. 133 [*DOI :* 10.1145/2601097.2601169], https://hal.inria.fr/hal-00983262

[3] B. DALSTEIN, R. RONFARD, M. VAN DE PANNE. *Vector Graphics Animation with Time-Varying Topology*, in "ACM Transactions on Graphics", August 2015, 12, https://hal.inria.fr/hal-01155246

[4] A. H. DICKO, T. LIU, B. GILLES, L. KAVAN, F. FAURE, O. PALOMBI, M.-P. CANI. *Anatomy Transfer*, in "ACM Transactions on Graphics", November 2013, vol. 32, n^o 6, Article No. 188 [*DOI :* 10.1145/2508363.2508415], https://hal.inria.fr/hal-00862502

[5] A. EMILIEN, U. VIMONT, M.-P. CANI, P. POULIN, B. BENES. *WorldBrush: Interactive Example-based Synthesis of Procedural Virtual Worlds*, in "ACM Transactions on Graphics", August 2015, vol. 34, n^o 4, 11 [*DOI :* 10.1145/2766975], https://hal.inria.fr/hal-01147913

[6] M. GUAY, M.-P. CANI, R. RONFARD. *The Line of Action: an Intuitive Interface for Expressive Character Posing*, in "ACM Transactions on Graphics", November 2013, vol. 32, n^o 6, Article No. 205 [*DOI :* 10.1145/2508363.2508397], https://hal.inria.fr/hal-00861503

[7] M. GUAY, R. RONFARD, M. GLEICHER, M.-P. CANI. *Space-time sketching of character animation*, in "ACM Transactions on Graphics", August 2015, vol. 34, n^o 4, Article No. 118 [*DOI :* 10.1145/2766893], https://hal.archives-ouvertes.fr/hal-01153763

[8] A. JUNG, S. HAHMANN, D. ROHMER, A. BEGAULT, L. BOISSIEUX, M.-P. CANI. *Sketching Folds: Developable Surfaces from Non-Planar Silhouettes*, in "ACM Transactions on Graphics (TOG)", 2015, vol. 34, n^o 5, p. 155:1–155:12, https://hal.inria.fr/hal-01152904

[9] 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

[10] C. SCHRECK, D. ROHMER, S. HAHMANN, M.-P. CANI, S. JIN, C. C. WANG, J.-F. BLOCH. *Non-smooth developable geometry for interactively animating paper crumpling*, in "ACM Transactions on Graphics", December 2015, vol. 35, n^o 1, p. 10:1-10:18 [*DOI :* 10.1145/2829948], https://hal.inria.fr/hal-01202571

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] A. FONDEVILLA. *Reverse-Engineering Fashion Products: From a single-view Sketch to a 3D Model*, UGA Grenoble France, December 2019

[12] M. GARCIA. *Performance transfer: animating virtual charaters by playing and acting*, UGA Grenoble France, December 2019

Articles in International Peer-Reviewed Journal

[13] T. BUFFET, D. ROHMER, L. BARTHE, L. BOISSIEUX, M.-P. CANI. *Implicit Untangling: A Robust Solution for Modeling Layered Clothing*, in "ACM Transactions on Graphics", July 2019, vol. 38, n^o 4, Article No. 120:1-12 [*DOI :* 10.1145/3306346.3323010], https://hal.archives-ouvertes.fr/hal-02129156

[14] G. CORDONNIER, B. BOVY, J. BRAUN.*A Versatile, Linear Complexity Algorithm for Flow Routing in Topographies with Depressions*, in "Earth Surface Dynamics", June 2019, vol. 7, n^o 2, p. 549-562 [*DOI :* 10.5194/ESURF-2018-81], https://hal.archives-ouvertes.fr/hal-02136750

[15] Y. LE VAOU, J.-C. LÉON, S. HAHMANN, S. MASFRAND, M. MIKA.*As-Stiff-As-Needed Surface Deformation Combining ARAP Energy with an Anisotropic Material*, in "Computer-Aided Design", April 2020, vol. 121, p. 1-15 [*DOI :* 10.1016/J.CAD.2019.102803], https://hal.inria.fr/hal-02428505

[16] J. MARTÍNEZ, M. SKOURAS, C. SCHUMACHER, S. HORNUS, S. LEFEBVRE, B. THOMASZEWSKI.*Star-Shaped Metrics for Mechanical Metamaterial Design*, in "ACM Transactions on Graphics", July 2019, vol. 38, n^o 4, Article No. 82 :1-13, Special issue, SIGGRAPH 2019 [*DOI :* 10.1145/3306346.3322989], https://hal.inria.fr/hal-02118846

[17] G. NIETO, F. DEVERNAY, J. L. CROWLEY.*Rendu basé image avec contraintes sur les gradients*, in "Traitement du Signal", 2019, p. 1-26, forthcoming [*DOI :* 10.3166/HSP.X.1-26], https://hal.archives-ouvertes.fr/hal-01900200

[18] O. PALOMBI, F. JOUANOT, N. NZIENGAM, B. OMIDVAR-TEHRANI, M.-C. ROUSSET, A. SANCHEZ.*OntoSIDES: Ontology-based student progress monitoring on the national evaluation system of French Medical Schools*, in "Artificial Intelligence in Medicine", May 2019, vol. 96, p. 59-67 [*DOI :* 10.1016/J.ARTMED.2019.03.006], https://hal.archives-ouvertes.fr/hal-02089318

[19] S. SUNDARAM, M. SKOURAS, D. S. KIM, L. VAN DEN HEUVEL, W. MATUSIK.*Topology optimization and 3D-printing of multimaterial magnetic actuators and displays*, in "Science Advances ", July 2019, vol. 5, n^o 7, p. 1-43 [*DOI :* 10.1126/SCIADV.AAW1160], https://hal.inria.fr/hal-02397188

[20] X. ZHANG, G. FANG, M. SKOURAS, G. GIESELER, C. C. WANG, E. WHITING.*Computational Design of Fabric Formwork*, in "ACM Transactions on Graphics", July 2019, vol. 38, n^o 4, p. 1-13 [*DOI :* 10.1145/3306346.3322988], https://hal.inria.fr/hal-02397265

International Conferences with Proceedings

[21] M. CARRIÈRE, M. SKOURAS, S. HAHMANN.*3D Design Of Ancient Garments*, in "GCH 2019 - EUROGRAPHICS Workshop on Graphics and Cultural Heritage", Sarajevo, Bosnia and Herzegovina, November 2019, p. 1-5, https://hal.inria.fr/hal-02312274

[22] P. CASATI, R. RONFARD, S. HAHMANN.*Approximate Reconstruction of 3D Scenes From Bas-Reliefs*, in "GCH 2019 - EUROGRAPHICS Workshop on Graphics and Cultural Heritage", Sarajevo, Bosnia and Herzegovina, November 2019, p. 1-10, https://hal.inria.fr/hal-02313660

[23] M. GARCIA, R. RONFARD, M.-P. CANI.*Spatial Motion Doodles: Sketching Animation in VR Using Hand Gestures and Laban Motion Analysis*, in "MIG 2019 - ACM SIGGRAPH Conference on Motion, Interaction and Games", Newcastle upon Tyne, United Kingdom, October 2019 [*DOI :* 10.1145/3359566.3360061], https://hal.archives-ouvertes.fr/hal-02303803

[24] 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 "HFR - 10th International Workshop on Human-Friendy Robotics", Napoli, Italy, Springer Proceedings in Advanced Robotics, 2019, vol. 7, p.

207-221, https://arxiv.org/abs/1801.01281 - In Human-Friendly Robotics 2017 [*DOI :* 10.1007/978-3-319-89327-3_16], https://hal.archives-ouvertes.fr/hal-01674511

[25] V. A. MURUKUTLA, E. CATTAN, O. PALOMBI, R. RONFARD.*Text-to-Movie Authoring of Anatomy Lessons*, in "AIME 2019 - 17th conference on artificial intelligence in medicine", Poznan, Poland, Lecture Notes in Computer Science, Springer, June 2019, vol. 11526, p. 304-308 [*DOI :* 10.1007/978-3-030-21642-9_39], https://hal.inria.fr/hal-02143822

Conferences without Proceedings

[26] N. ABOUBAKR, J. L. CROWLEY, R. RONFARD.*Recognizing Manipulation Actions from State-Transformations*, in "EPIC@CVPR 2019 - Fourth International Workshop on Egocentric Perception Interaction and Computing", Long Beach, United States, June 2019, p. 1-4, https://arxiv.org/abs/1906.05147 - Accepted for presentation at EPIC@CVPR2019 workshop, https://hal.archives-ouvertes.fr/hal-02197549

Research Reports

[27] N. ABOUBAKR, J. L. CROWLEY, R. RONFARD.*Recognizing Manipulation Actions from State-Transformations (Technical Report)*, Univ. Grenoble Alps, CNRS, Inria, Grenoble INP, LIG, 38000 Grenoble, France, June 2019, https://hal.archives-ouvertes.fr/hal-02197595

[28] J. DAVAL, M. BEAUD DE BRIVE, V. DAUNIS, R. RONFARD.*Sketch-based animation in the RUMBA software - Final report of COLLODI2 project.*, Inria, September 2019, n^o RT-0505, https://hal.inria.fr/hal-02374171

Team MARACAS

Models and Algorithms for Reliable Communication Systems

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
Networks and Telecommunications

Table of contents

Team MARACAS

Creation of the Project-Team: 2020 January 01

Keywords:

Computer Science and Digital Science:

A1.2.5. - Internet of things
A1.2.6. - Sensor networks
A1.2.7. - Cyber-physical systems
A1.5.2. - Communicating systems
A3.4.1. - Supervised learning
A3.4.3. - Reinforcement learning
A3.4.8. - Deep learning
A5.9. - Signal processing
A5.9.2. - Estimation, modeling
A5.9.6. - Optimization tools
A8.6. - Information theory
A8.7. - Graph theory
A8.8. - Network science
A8.11. - Game Theory
A9.2. - Machine learning
A9.3. - Signal analysis
A9.9. - Distributed AI, Multi-agent

Other Research Topics and Application Domains:

B1.1.10. - Systems and synthetic biology
B4.5.1. - Green computing
B6.2.2. - Radio technology
B6.4. - Internet of things
B6.6. - Embedded systems
B8.1. - Smart building/home
B8.2. - Connected city

1. Team, Visitors, External Collaborators

Research Scientists

Malcolm Egan [Inria, Researcher, from Sep 2019 - INSA Lyon, Starting Research position until Sep 2019]
Samir Medina Perlaza [Inria, Researcher]

Faculty Members

Jean Marie Gorce [Team leader, INSA Lyon, Professor, HDR]
Claire Goursaud [INSA Lyon, Associate Professor, HDR]
Leonardo Sampaio Cardoso [INSA Lyon, Associate Professor]

External Collaborators

Nikolai Lebedev [École supérieure de chimie physique électronique de Lyon]
Chantal Muller [INSA Lyon, HDR]

Technical Staff

Pascal Girard [INSA Lyon, Engineer]
Matthieu Imbert [Inria, Engineer]
Olivier Nicolas [INSA Lyon, Engineer, until Nov 2019]
Othmane Oubejja [Inria, Engineer]

PhD Students

Dadja Anade Akpo [INSA Lyon, PhD Student]
Lélio Chetot [Inria, PhD Student]
Diane Duchemin [Inria, PhD Student]
Mathieu Goutay [Nokia, PhD Student, from Mar 2019, granted by CIFRE]
Hassan Kallam [Insavalor, PhD Student]
Nizar Khalfet [Inria, PhD Student, until Nov 2019]
David Kibloff [Inria, PhD Student, until Sep 2019]
Cyrille Morin [Inria, PhD Student]

Post-Doctoral Fellow

Bayram Akdeniz [Inria, Post-Doctoral Fellow]

Visiting Scientists

Michalis Eliodorou [MTN Cyprus, Sep 2019]
Matei Moldoveanu [U. of Sheffield, UK, from Jul 2019 until Sep 2019]

Administrative Assistants

Sophie Karsai [Inria, Administrative Assistant, until Jul 2019]
Anouchka Ronceray [Inria, Administrative Assistant, from Jul 2019]

2. Overall Objectives

2.1. Motivation

During the last century, the industry of communications was devoted to improving human connectivity, leading to a seamless worldwide coverage to cope with increasing data rate demands and mobility requirements. The Internet revolution drew on a robust and efficient multi-layer architecture ensuring end-to-end services. In a classical network architecture, the different protocol layers are compartmentalized and cannot easily interact. For instance, source coding is performed at the application layer while channel coding is performed at the physical (PHY) layer. This multi-layer architecture blocked any attempt to exploit low level cooperation mechanisms such as relaying, phy-layer network coding or joint estimation. During the last decade, a major shift, often referred to as *the Internet of Things (IoT)*, was initiated toward a machine-to-machine (M2M) communication paradigm, which is in sharp contrast with classical centralized network architectures. The IoT enables machine-based services exploiting a massive quantity of data virtually spread over a complex, redundant and distributed architecture.

This new paradigm makes the aforementioned classical network architecture based on a centralized approach out-of-date.

The era of *Internet of Everything* deeply modifies the paradigm of communication systems. They have to transmute into reactive and adaptive intelligent systems, under stringent QoS constraints (latency, reliability) where the networking service is intertwined in an information-centric network. The associated challenges are linked to the intimate connections between communication, computation, control and storage. Actors, nodes or agents in a network can be viewed as forming a distributed system of computations—a *computing network*.

2.2. Scientific methodology

It is worth noting that working on these new architectures can be tackled from different perspectives, e.g. data management, protocol design, middleware, algorithmic design... Our main objective in Maracas is to address this problem from a communication theory perspective. Our background in communication theory includes information theory, estimation theory, learning and signal processing. Our strategy relies on three fundamental and complementary research axes:

- Mathematical modeling: information theory is a powerful framework suitable to evaluate the limits of complex systems and relies on probability theory. We will explore new bounds for complex networks (multi-objective optimization, large scale, complex channels,...) in association with other tools (stochastic geometry, queuing theory, learning,...)
- Algorithmic design: a number of theoretical results obtained in communication theory, despite their high potential are still far from a practical use. We will thus work on exploiting new algorithmic techniques. Back and forth efforts between theory and practice is necessary to identify the most promising opportunities. The key elements are related to the exploitation of feedbacks, signaling and decentralized decisions. Machine learning algorithms will be explored.
- Experimentation and cross-layer approach: theoretical results and simulation are not enough to provide proofs of concept. We will continue to put efforts on experimental works either on our own (e.g. FIT/CorteXlab and SILECS) or in collaboration with industries (Nokia, Orange, Thalès,...) and other research groups.

While our expertise is mostly related to the optimization of wireless networks from a communication perspective, the project of Maracas is to broaden our scope in the context of *Computing Networks*, where a challenging issue is to optimize jointly architectures and applications, and to break the classical network/data processing separation. This will drive us to change our initial positioning and to really think in terms of information-centric networks following, e.g. [60], [58], [66].

To summarize, *Computing Networks* can be described as highly distributed and dynamic systems, where information streams consist in a huge number of transient data flows from a huge number of nodes (sensors, routers, actuators, etc...) with computing capabilities at the nodes. These *Computing Networks* are nothing but the invisible nonetheless necessary skeleton of cloud and fog-computing based services.

Our research strategy is to describe these *Computing Networks* as complex large scale systems in an information theory framework, but in association with other tools, such as stochastic geometry, stochastic network calculus, game theory [19] or machine learning.

The multi-user communication capability is a central feature, to be tackled in association with other concepts and to assess a large variety of constraints related to the data (storage, secrecy,...) or related to the network (energy, self-healing,...).

The information theory literature or more generally the communication theory literature is rich of appealing techniques dedicated to efficient multi-user communications: e.g. physical layer network coding, amplify-and-forward, full-duplexing, coded caching at the edge, superposition coding. But despite their promising performance, none of these technologies play a central role in current protocols. The reasons are two-fold : i) these techniques are usually studied in an oversimplified theoretical framework which neglect many practical aspects (feedback, quantization,...), and that is not able to tackle large scale networks and ii) the proposed algorithms are of a high complexity and are not compatible with the classical multi-layer network architecture.

Maracas addresses these questions, leveraging on its past outstanding experience from wireless network design.

The aim of Maracas is to push from theory to practice a fully cross-layer design of *Computing Networks* , based on multi-user communication principles relying mostly on information theory, signal processing, estimation theory, game theory and optimization. We refer to all these tools under the umbrella of *communication theory* .

As such, Maracas project goes much beyond wireless networks. The *Computing Networks* paradigm applies to a wide variety or architectures including wired networks, smart grids, nanotechnology based networks. One Maracas research axis will be devoted to the identification of new research topics or scenarios where our algorithms and mathematical models could be useful.

3. Research Program

3.1. General description

As presented in the first section, *Computing Networks* is a concept generalizing the study of multi-user systems under the communication perspective. This problematic is partly addressed in the aforementioned references. Optimizing *Computing Networks* relies on exploiting simultaneously multi-user communication capabilities, in the one hand, and storage and computing resources in the other hand. Such optimization needs to cope with various constraints such as energy efficiency or energy harvesting, delays, reliability or network load.

The notion of reliability (used in MARACAS acronym) is central when considered in the most general sense : ultimately, the reliability of a *Computing Network* measures its capability to perform its intended role under some confidence interval. Figure 1 represents the most important performance criteria to be considered to achieve reliable communications. These metrics fit with those considered in 5G and beyond technologies [63].

On the theoretical side, multi-user information theory is a keystone element. It is worth noting that classical information theory focuses on the power-bandwith tradeoff usually referred as Energy Efficiency-Spectral Efficiency (EE-SE) tradeoff (green arrow on 1). However, the other constraints can be efficiently introduced by using a non-asymptotic formulation of the fundamental limits [62], [64] and in association with other tools devoted to the analysis of random processes (queuing theory, ...).

Maracas aims at studying *Computing Networks* from a communication point of view, using the foundations of information theory in association with other theoretical tools related to estimation theory and probability theory.

In particular, Maracas combines techniques from communication and information theory with statistical signal processing, control theory, and game theory. Wireless networks is the emblematic application for Maracas, but other scenarios are appealing for us, such as molecular communications, smart grids or smart buildings.

Several teams at Inria are addressing computing networks, but working on this problem with an emphasis on communication aspects is unique within Inria.

The complexity of *Computing Networks* comes first from the high dimensionality of the problem: i) thousands of nodes, each with up to tens setting parameters and ii) tens variable objective functions to be minimized/maximized.

In addition, the necessary decentralization of the decision process, the non stationary behavior of the network itself (mobility, ON/OFF Switching) and of the data flows, and the necessary reduction of costly feedback and signaling (channel estimation, topology discovering, medium access policies...) are additional features that increase the problem complexity.

The original positioning of Maracas holds in his capability to address three complementary challenges :

1. **to develop a sound mathematical framework inspired by information theory.**
2. **to design algorithms, achieving performance close to these limits.**
3. **to test and validate these algorithms on experimental testbeds.**

3.2. Research program

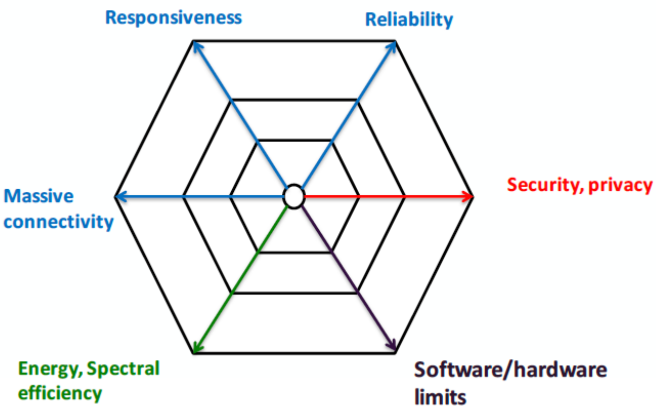


Figure 1. Main metrics for future networks (5G and beyond)

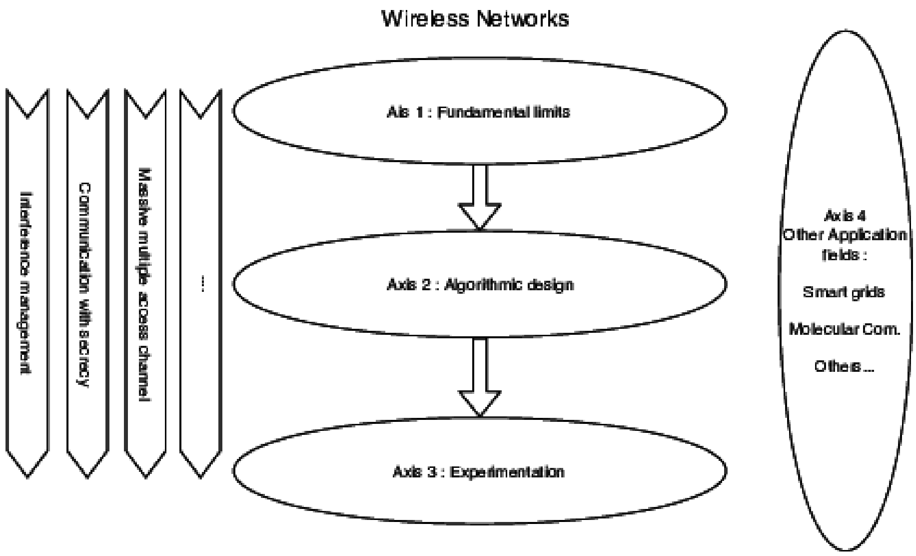


Figure 2. Maracas organization

Our research is organized in 4 research axes:

- **Axis 1 - Fundamental Limits of Reliable Communication Systems**: Information theory is revisited to integrate reliability in the wide sense. The non-asymptotic theory which made progress recently and attracted a lot of interest in the information theory community is a good starting point. But for addressing computing network in a wide sense, it is necessary to go back to the foundation of communication theory and to derive new results, e.g. for non Gaussian channels [8] of for multi-constrained systems [17].

  This also means revisiting the fundamental estimation-detection problem [65] in a general multi-criteria, multi-user framework to derive tractable and meaningful bounds.

  As mentioned in the introduction, *Computing Networks* also relies on a data-centric vision, where transmission, storage and processing are jointly optimized. The strategy of *caching at the edge* [57] proposed for cellular networks shows the high potential of considering simultaneously data and network properties. Maracas is willing to extend his skills on source coding aspects to tackle with a data-oriented modeling of *Computing Networks*.

- **Axis 2 - Algorithms and protocols**: Our second objective is to elaborate new algorithms and protocols able to achieve or at least to approach the aforementioned fundamental limits. While the exploration of fundamental limits is helpful to determine the most promising strategies (e.g. relaying, cooperation, interference alignment) to increase system performance, the transformation of these degrees of freedom into real protocols is a non trivial issue. One reason is the exponentially growing complexity of multi-user communication strategies, with the number of users, due to the necessity of some coordination, feedback and signaling. The general problem is a decentralized and dynamic multi-agents multi-criteria optimization problem and the general formulation is a non-linear and non-convex large scale problem.

  The conventional research direction aims at reducing the complexity by relaxing some constraints or by reducing the number of degrees of freedom. For instance, topology interference management is a seducing model used to reduce feedback needs in decentralized wireless networks leading to original and efficient algorithms [67], [59].

  Another emerging research direction relies on using machine learning techniques [54] as a natural evolution of cognitive radio based approaches. Machine learning in the wide sense is not new in radio networks, but the most important works in the past were devoted to reinforcement learning approaches. The use of deep learning (DL) is much more recent, with two important issues : i) identifying the right problems that really need DL algorithms and ii) providing extensive data sets from simulation and real experiments. Our group started to work on this topic in association with Nokia in the joint research lab. As we are not currently expert in deep learning, our primary objective is to identify the strategic problems and to collaborate in the future with Inria experts in DL, and in the long term to contribute not only to the application of these techniques, but also to improve their design according to the constraints of computing networks.

- **Axis 3 - Experimental validation** : With the rapid evolution of network technologies, and their increasing complexity, experimental validation is necessary for two reasons: to get data, and to validate new algorithms on real systems.

  Maracas activity leverages on the FIT/CorteXlab platform (http://www.cortexlab.fr/), and our strong partnerships with leading industry including Nokia Bell Labs, Orange labs, Sigfox or Sequans. Beyond the platform itself which offers a worldwide unique and remotely accessible testbed , Maracas also develops original experimentations exploiting the reproducibility, the remote accessibility, and the deployment facilities to produce original results at the interface of academic and industrial research [1], [10]. FIT/CorteXlab uses the GNU Radio environment to evaluate new multi-user communication systems.

  Our experimental work is developed in collaboration with other Inria teams especially in the Rhone-Alpes centre but also in the context of the future SILECS project https://www.silecs.net/ which will implement the convergence between FIT and Grid'5000 infrastructures in France, in cooperation with European partners and infrastructures. SILECS is a unique framework which will allow us to test our algorithms, to generate data, as required to develop a data-centric approach for computing networks.

Last but not least, software radio technologies are leaving the confidentiality of research laboratories and are made available to a wide public market with cheap (few euros) programmable equipment, allowing to setup non standard radio systems. The existence of home-made and non official radio systems with legacy ones could prejudice the deployment of Internet of things. Developing efficient algorithms able to detect, analyse and control the spectrum usage is an important issue. Our research on FIT/CorteXlab will contribute to this know-how.

- **Axis 4 - Other application fields** : Even if the wireless network context is still challenging and provides interesting problems, Maracas targets to broaden its exploratory playground from an application perspective. We are looking for new communication systems, or simply other multi-user decentralized systems, for which the theory developed in the context of wireless networks can be useful. Basically, Maracas might address any problem where multi-agents are trying to optimize their common behavior and where the communication performance is critical (e.g. vehicular communications, multi-robots systems, cyberphysical systems). Following this objective, we already studied the problem of missing data recovery in smart grids [11] and the original paradigm of molecular communications [6].

  Of course, the objective of this axis is not to address random topics but to exploit our scientific background on new problems, in collaboration with other academic teams or industry. This is a winning strategy to develop new partnerships, in collaboration with other Inria teams.

4. Application Domains

4.1. 5G, 6G, and beyond

The fifth generation (5G) broadens the usage of cellular networks but requires new features, typically very high rates, high reliability, ultra low latency, for immersive applications, tactile internet, M2M communications.

From the technical side, new elements such as millimeter waves, massive MIMO, massive access are under evaluation. The initial 5G standard validated in 2019, is finally not really disruptive with respect to the 4G and the clear breakthrough is not there yet. The ideal network architecture for billions of devices in the general context of Internet of Things, is not well established and the debate still exists between several proposals such as NB-IoT, Sigfox, Lora. We are developing a deep understanding of these techniques, in collaboration with major actors (Orange Labs, Nokia Bell Labs, Sequans, Sigfox) and we want to be able to evaluate, to compare and to propose evolutions of these standards with an independent point of view.

This is why we are interested in developing partnerships with major industries, access providers but also with service providers to position our research in a joint optimization of the network infrastructure and the data services, from a theoretical perspective as well as from experimentation.

4.2. Energy sustainability

The energy footprint and from a more general perspective, the sustainability of wireless cellular networks and wireless connectivity is somehow questionable.

We develop our models and analysis with a careful consideration of the energy footprint : sleeping modes, power adaptation, interference reduction, energy gathering, ... many techniques can be optimized to reduce the energetic impact of wireless connectivity. In a *computing networks* approach, considering simultaneously transmission, storage and computation constraints may help to reduce drastically the overall energy footprint.

4.3. Smart building, smart cities, smart environments

Smart environments rely on the deployment of many sensors and actuators allowing to create interactions between the twinned virtual and real worlds. These smart environments (e.g. smart building) are for us an ideal playground to develop new models based on information theory and estimation theory to optimize the network architecture including storage, transmission, computation at the right place.

Our work can be seen as the dark side of cloud/edge computing. In collaboration with other teams expert in distributed computing or middleware (typically at CITIlab, with the team Dynamid of Frédéric Le Mouel) and in the framework of the chaire SPIE/ICS-INSA Lyon, we want to optimize the mechanisms associated to these technologies : in a multi-constrained approach, we want to design new distributed algorithms appropriate for large scale smart environments.

4.4. Machine learning based radio

During the first 6G wireless meeting which was held in Lapland, Finland in March 2019, machine learning (ML) was clearly identified as one of the most promising breakthroughs for future 6G wireless systems expected to be in use around 2030 (https://www.6gsummit.com/). The research community is entirely leveraging the international ML tsunami. We strongly believe that the paradigm of wireless networks is moving toward to a new era. Our view is supported by the fact that artificial Intelligence (AI) in wireless communications is not new at all. The telecommunications industry has been seeking for 20 years to reduce the operational complexity of communication networks in order to simplify constraints and to reduce costs on deployments. This obviously relies on data-driven techniques allowing the network to self-tune its own parameters. Over the successive 3GPP standard releases, more and more sophisticated network control has been introduced. This has supported increasing flexibility and further self-optimization capabilities for radio resource management (RRM) as well as for network parameters optimization.

We target the following key elements :

- Obtaining data from experimental scenarios, at the lowest level (baseband I/Q signals) in multi-user scenarios (based upon FIT/CorteXlab).
- Developing a framework and algorithms for deep learning based radio.
- Developing new reinforcement learning techniques in high dimensional state-action spaces.
- Embedding NN structures on radio devices (FPGA or m-controllers) and in FIT/CorteXlab.
- Evaluating the gap between these algorithms and fundamental limits from information theory.
- Building an application scenario in a smart environment to experiment a fully cross-layer design (e.g. within a smart-building context, how could a set of object could learn their protocols efficiently ?)

4.5. Molecular communications

Many communication mechanisms are based on acoustic or electromagnetic propagation; however, the general theory of communication is much more widely applicable. One recent proposal is molecular communication, where information is encoded in the type, quantity, or time or release of molecules. This perspective has interesting implications for the understanding of biochemical processes and also chemical-based communication where other signaling schemes are not easy to use (e.g., in mines). Our work in this area focuses on two aspects: (i) the fundamental limits of communication (i.e., how much data can be transmitted within a given period of time); and (ii) signal processing strategies which can be implemented by circuits built from chemical reaction-diffusion systems.

A novel perspective introduced within our work is the incorporation of coexistence constraints. That is, we consider molecular communication in a crowded biochemical environment where communication should not impact pre-existing behavior of the environment. This has lead to new connections with communication subject to security constraints as well as the stability theory of stochastic chemical reaction-diffusion systems and systems of partial differential equations which provide deterministic approximations.

5. Highlights of the Year

5.1. Highlights of the Year

Over the last year, the MARACAS team has made a number of significant contributions in the form of journal publications and international conference proceedings, invited lectures in international conferences and schools, as well as contributions in the form of organization of international conferences and editorial roles in international IEEE journals. These include 9 high quality journal publications and over 10 international conference proceedings, spanning many areas of communication and information theory as well as signal processing. In particular, these results contribute to the ongoing development of 5G wireless communication systems and also to emerging areas of communications in the form of the smart grid and molecular communications.

5.1.1. Awards

- Samir Perlaza: Visiting Research Collaborator (Honorific Position), Term 2019 - 2020, at the Department of Electrical Engineering, Princeton University. Annual Renew under evaluation of a Departmental Committee.
- Samir Perlaza: Fellowship of The Finnish Society of Sciences and Letters for visiting the School of Energy Systems at Lappeenranta University of Technology, Finland. April, 2019.
- Bayram Akdeniz: 2nd place in the Molecular MIMO Competition at the IEEE Communication Theory Workshop.
- Cyrille Morin: 1st place in the Machine learning challenge at the 6th Training School on Machine and Deep Learning Techniques for (Beyond) 5G Wireless Communication Systems.

6. New Software and Platforms

6.1. cortexlab-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

6.2. cortexlab-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

6.3. cortexlab-webapp

KEYWORDS: Experimentation - SDR (Software Defined Radio)

FUNCTIONAL DESCRIPTION: User management module, which aims at easing platform usage and improving the metadata that we can associate with each experimenter and experiment. This metadata aims at improving the metrics we can gather about the platform's usage

- Partner: Insa de Lyon
- Contact: Pascal Girard

6.4. CorteXlab-IoT Framework

Framework for PHY-MAC layers Prototyping in Dense IoT Networks using CorteXlab Testbed

KEYWORDS: SDR (Software Defined Radio) - Iot - CorteXlab - GNU Radio

FUNCTIONAL DESCRIPTION: This framework was developed in the project ”Enhanced Physical Layer for Cellular IoT” (EPHYL). It provides a customizable and open source design for IoT networks prototyping in a massive multi-user, synchronized and reproducible environment thanks to the hardware and software capabilities of the testbed.

- Author: Othmane Oubejja
- Contact: Othmane Oubejja
- Publication: Framework for PHY-MAC layers Prototyping in Dense IoT Networks using FIT/CorteXlab Testbed
- URL: https://github.com/CorteXlab/gr-ephyl

6.5. Platforms

6.5.1. FIT/CorteXlab

FIT (Future Internet of Things) is a french Equipex (Équipement d'excellence) which aims to develop an experimental facility, a federated and competitive infrastructure with international visibility and a broad panel of customers. FIT is be composed of four main parts: a Network Operations Center (FIT NOC), a set of IoT test-beds (FIT IoT-Lab), a set of wireless test-beds (FIT-Wireless) which includes the FIT/CorteXlab platform deployed previoulsy by the Socrate team and now managed by Maracas team in the Citi lab, and finally a set of Cloud test-beds (FIT-Cloud). In 2014 the construction of the room was done and SDR nodes have been installed in the room: 42 industrial PCs (Aplus Nuvo-3000E/P), 22 NI radio boards (usrp) and 18 Nutaq boards (PicoSDR, 2x2 and 4X4) can be programmed remotely, from internet now.

New features are under developments, and the platform is used or has been used by several research groups : i) Nokia Bell Labs - Inria common labs for research in machine learning, ii) ANR projects EPHYL and ARBURST for research on massive access for IoT, iii) ERC of prof. Michele Wigger, IMT Paris, for validating distributed coding strategies, iv) Nokia Bell Labs New Jersey for a proof of concept of wireless caching, Greentouch international consortium to develop a prototype of interference alignment.

7. New Results

7.1. Results of axis 1: fundamental limits

We worked in 2019 on the following main research directions:

1. Fundamental limits of IoT networks

Figure 3. FIT/CorteXlab facility

Table 1.

Principal Investigators:	Malcolm Egan, Samir Perlaza, Jean-Marie Gorce
Students:	Dadja Toussaint Anade-Akbo, Lélio Chetot
Funding:	Orange Labs, ANR Arburst
Partners:	Philippe Mary (IETR, Rennes), Laurent Clavier (IRCICA, Lille) JM Kélif (Orange Labs) H. Vincent Poor (Princeton University, NJ, USA)
Publications:	[34], [48], [36], [49], [35]

One of the main figures of merit in an IoT cell is the capability to support a massive access from distributed nodes, but with very small information quantity [12]. This perspective raises fundamental questions relative to the theoretical limits and performance of this kind of very large scale deployments. Fundamental limits are neither well known nor even well formulated. What is the maximal number of IoT nodes we may deploy in a given environment? At which energetic cost? With which transmission reliability or latency? These multiple questions highlight that the problem is not unique and the capacity is not the only (and even not the main) challenge to be addressed. We aim at establishing the fundamental limits of a decentralized system in a bursty regime which includes short packets of information and impulsive interference regime. We are targeting the fundamental limits and their mathematical expression, according to the usual information theory framework capturing the capacity region by establishing converse and achievability theorems.

2. Stability and sensitivity of fundamental limits

Table 2.

Principal Investigator:	Malcolm Egan, Samir Perlaza
Students:	-
Funding:	
Partners:	H. Vincent Poor, Alex Disto, Princeton University Vyacheslav Kungurtsev, Czech Technical University
Publications:	[8],[33]

The analysis of the fundamental limits on communications systems is performed under some assumptions including Gaussian noise, channel input symbols with average power, among others. Nonetheless, despite that these constraints were well suited for describing communications systems in the early 90's, the evolution of these systems make these assumptions vacuous today. Often, noise is better described by α–stable stochastic processes in IoT networks and channel inputs are subject to constraints in the amplitude, energy harvesting etc. From this perspective, our contributions are based on the notion of capacity sensitivity to study the capacity of continuous memoryless point-to-point channels. The capacity sensitivity reflects how the capacity changes with small perturbations in any of the parameters describing the channel, e.g., cost constraints on the input distribution as well as on the noise distribution.

3. Energy self-sustained wireless networks

Table 3.

Principal Investigator:	Samir Perlaza
Students:	Nizar Khalfet
Funding:	H2020 ComMed
Partners:	I. Kikridis (U. of Cyprus)
Publications:	[29], [42], [43], [50]

The main scientific challenge is to set up a theoretical framework for designing and developing fully decentralized energy-self-sustained communications systems. The main motivation stems from the fact that wireless networks deployed in hard-to-reach places, e.g., remote geographical areas, concrete structures, human body or war zones are often limited by the lifetime of their batteries. This contrasts with the fact that hardware is built to last for very long periods. One of the solutions being considered today for solving the energy limitation problem is the use of energy harvesting (EH) techniques. Within this context, our work focuses on the study of wireless communications systems based on EH sources. EH is expected to be the enabler of energy self-sustainability by eliminating the critical dependence on manual battery recharging.

However, a solid answer on whether or not EH is a viable solution can be given only if the corresponding fundamental limits of data transmission based on EH are known. This is mainly because these limits are based on the laws of Physics and thus, determine the barrier between feasible and unfeasible systems. We study the fundamental limits of three strongly correlated problems regarding the energy supply of future wireless networks: (i) Data transmission over centralized and decentralized EH multi-user channels; (ii) Simultaneous energy and information transmission in multi-user channels; and (iii) Energy cooperation. In a near future, we expect to exploit these results to design algorithms and protocols and later to perform a proof of concept on FIT/CorteXlab. We believe that a solid theoretical framework may help to drive the future design and performance evaluation of applications involving EH based wireless communications systems within smart buildings, smart cities.

4. Security and Privacy

Table 4.

Principal Investigator:	Samir Perlaza
Students:	David Kibloff
Funding:	Inria-DGA PhD
Partners:	Guilaume Villemaud (Socrate), Ligong Wang (ETIS, Cergy) Raphael Shaeffer (TU Berlin)
Publications:	[44], [51]

Information theory is also well adapted to study the fundamental limits of privacy and secrecy. Indeed, the wiretap channel and the covert communication [53] models have been shown to be appropriate for privacy preserving communications in wireless communications. With the PhD of David Kibloff defended in October 2019, we explored the following problem. Given a code used to send a message to two receivers through a degraded discrete memoryless broadcast channel (DM-BC), the sender wishes to alter the codewords to achieve the following goals: (i) the original broadcast communication continues to take place, possibly at the expense of a tolerable increase of the decoding error probability; and (ii) an additional covert message can be transmitted to the stronger receiver such that the weaker receiver cannot detect the existence of this message. The main results are: (a) feasibility of covert communications is proven by using a random coding argument for general DM-BCs; and (b) necessary conditions for establishing covert communications are described and an impossibility (converse) result is presented for a particular class of DM-BCs. Together, these results characterize the asymptotic fundamental limits of covert communications for this particular class of DM-BCs within an arbitrarily small gap. Future extensions will concern the Gaussian and other continuous channels, or more complex scenarios where some subsets of nodes are willing to communicate while some external observers cannot even detect the existence of these messages. Covert communication allows to introduce a side constraint that prevent a network to be attacked.

5. Structured Codes for Quantization and Channel Estimation

Table 5.

Principal Investigator:	Malcolm Egan
Publications:	[25]

Finite frames are sequences of vectors in finite dimensional Hilbert spaces that play a key role in signal processing and coding theory. In this work, we study the class of tight unit-norm frames for $\mathbb{C}^d$ that also form regular schemes, which we call tight regular schemes (TRS). Many common frames that arising in vector quantization and channel state estimation, 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 cubature rules for integrals over spheres in $\mathbb{C}^d$—with weights dependent on the Voronoi regions of each frame element. Aided by additional numerical evidence, we conjecture that all TRSs in fact satisfy this property.

7.2. Results of axis 2: algorithms

1. Massive random access in LPWAN

Table 6.

Principal Investigator:	Jean-Marie Gorce, Claire Goursaud
Students:	Diane Duchemin, Lélio Chetot
Funding:	ANR Ephyl, Inria-Nokia common lab
Partners:	Sequans, Supelec Rennes, ISEP, CEA Leti, Nokia
Publications:	[30], [31], [37], [47]

The optimization of IoT access techniques was the objective of the ANR Ephyl collaborative project, where we studied different solutions at the PHY and MAC layers as presented in [47].

The main question Maracas group addressed in this research is the detection of simultaneous random transmissions from distributed nodes. The underlying mechanism is a coded slotted Aloha allowing to avoid hand-skake mechanisms. Each node can transmit randomly and the receiver tries to detect several packets simultaneously. Our objective is to identify a good code family, and to determine the fundamental trade-off in terms of nodes density versus reliability. During this year, we focused on the detection of a small subset of simultaneous active nodes, exploiting optimal detection. We developed a MAP based iterative detector at a multi-antennas receiver in [30]. We also proposed a low complexity detector in [37].

This joint coding-decoding optimization problem will be also investigated from extensive simulations and experimental data (see section 3.4), and represents an interesting problem to evaluate deep learning based approaches.

2. Interference management

Table 7.

Principal Investigator:	Léonardo Cardoso, Jean-Marie Gorce
Students:	Hassan Khalam
Funding:	Fed4PMR (PIA)
Partners:	Thales
Publications:	[41]

Interference management and resource management is a very complex problem in wireless environment (e.g. [55]). The capacity region is known for some specific scenarios and some specific channel conditions. But the optimal performance relies on perfect feedback mechanisms, to get channel state information at the transmitters and to coordinate them. As proposed by Jafar et al, topological interference management (TIM) [56] is a seducing framework to balance performance with feedback complexity. In the context of the Fed4PMR project, we develop new algorithms to allow partial coordination between interfering transmitters [41], relying only on some partial interference information. This approach suits particularly well with the requirements of PMR networks, since their deployments is not optimized. The algorithm relies on an association of degrees of freedom evaluation, graph theory and interference alignment.

Based on this first study, we will explore the suitability of TIM in other application scenarios (especially for the standard IEE802.11ax under preparation). For short, TIM allows to build optimal graph representations of a wireless networks, with reduced coordination needs. TIM can be seen as an approach to optimally quantize a complex interfering graph and to distribute its knowledge in an optimal fashion.

3. Learning in radio systems

Table 8.

Principal Investigator:	Léonardo Cardoso, Malcolm Egan, Jean-Marie Gorce
Student:	Cyrille Morin, Mathieu Goutay
Funding:	ADR Analytics, Inria-Nokia common lab AI chair ANR program (applied)
Partners:	Jakob Hoydis, Nokia Bell Labs
Publications:	[45]

Following the artificial intelligence tsunami, the research community in wireless systems (both industry and academia) is engaged in a strong competition to determine how this revolution could change the paradigm of wireless networks. Following the preliminary studies made by Jakob Hoydis [54], we investigate in this research action, the potential of deep learning in radio communications. The central question is to identify which processing could take advantage from neural networks against classical approaches.

Our joint strategy with Nokia follows: we target the production of a huge set of experimental data with FIT/CorteXlab to facilitate the comparison of different solutions and to train neural networks on real data. We currently investigate three original problems : transmitter identification from its RF signature (Cyrille Morin PhD) [45], self-synchronization procedures based on neural networks (Cyrille Morin PhD) and dirty RF compensation (Mathieu Goutay PhD, patents submitted). Last but not least, we believe that an intelligent radio should be able to learn from its environment and to adapt its behavior. Therefore, in the future, we will explore reinforcement principles associated to neural networks and applied to learning based radio.

This topic is very hot, and most top ranked conference have special sessions on this topic. We believe that our partnership with Nokia, our data sets from FIT/CorteXlab and our experience in estimation theory let us be highly competitive.

7.3. Results of axis 3: experimental assessment

During 2019-2020, our experimental work was mostly devoted to the development of new functions of FIT/CorteXlab, and to the development of experimental evaluations with external partners.

1. Development of a user and administrative graphical interface

Table 9.

Principal Investigator:	Pascal Girard, Matthieu Imbert, Léonardo Cardoso
Funding:	FIT/CorteXlab
Partners:	FIT consortium

The objective is to develop a web-based user-friendly interface for using CorteXlab. Several modules are planned and the first module is the user management module, which aims at easing platform usage and improving the metadata that we can associate with each experimenter and experiment. This metadata aims at improving the metrics we can gather about the platform's usage.

2. Development of a docker-based experiment conducting middleware.

Table 10.

Principal Investigator:	Matthieu Imbert, Léonardo Cardoso
Funding:	FIT/CorteXlab
Partners:	FIT consortium

CorteXlab relies on Minus, an experiment conducting middleware which allows users to submit experimental tasks to the platform, handles the automatic execution of these experiments, and gathers their results. The initial design for Minus relies on a fixed toolchain (mainly composed of GNURadio, hardware drivers, and additional external or in-house software or GNURadio blocks, FPGA tools, etc.). Experimenters are supposed to use this fixed toolchain in a batch-like workflow. It is hard for experimenters to extend the limits of the fixed toolchain (e.g. to use a custom library or software, or a different version of GNURadio), and the development phase of an experiment can be painful due to the batch-like interface. To improve this, we have developed a new experimental workflow based on docker [61] images and containers which allows experimenters to use our in-house provided docker images [52], adapt them if needed, or even create completely custom ones. These images have the benefit that they can be used identically on the experimenters' workstations, on the CorteXlab platform, or another platform, and they can be used interactively if needed, even on CorteXlab. This increases greatly the ease of use of the platform, the reproducibility and share-ability of experiments, and the breadth of its usage.

3. Reference scenario for massive IoT access

Table 11.

Principal Investigator:	Othmane Oubejja, Jean-Marie Gorce Matthieu Imbert, Léonardo Cardoso
Funding:	ANR Ephyl, ANR ARburst
Partners:	CEA Leti, Supelec Rennes, Sequans
Publications:	[46]

In this work we developed an experimental setup for dense IoT access evaluation, as part of the project "Enhanced Physical Layer for Cellular IoT" (EPHYL), using FIT/CorteXlab radio testbed. The aim of this work is to provide a customizable and open source design for IoT networks prototyping in a massive multi-user, synchronized and reproducible environment thanks to the hardware and software capabilities of the testbed. The massive access feature is managed by emulating a base station and several sensors per radio nodes. As shown in Fig.4, two categories of modular network components are used in our design: a base station unit and a multi-sensor emulator unit. These components are separately hosted in dedicated and remotely accessible radio nodes.

The features of this design can be accessed through customizable demos as documentation and resources are available online. As a result, it is possible for any interested user to plug custom algorithms, evaluate diverse communication scenarios and perform necessary physical measurements.

7.4. Results of axis 4: other application fields

1. Smart Grid

Table 12.

Principal Investigators:	Samir Perlaza
Student:	Matei Moldoveanu (visitor)
Partners:	Inaki Esnaola
Publications:	[40]

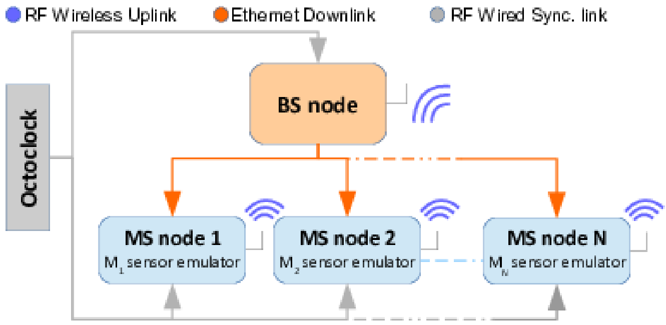


Figure 4. EPHYL IoT network representation

We study the recovery of missing data from multiple smart grid datasets within a matrix completion framework. The datasets contain the electrical magnitudes required for monitoring and control of the electricity distribution system. Each dataset is described by a low rank matrix. Different datasets are correlated as a result of containing measurements of different physical magnitudes generated by the same distribution system. To assess the validity of matrix completion techniques in the recovery of missing data, we characterize the fundamental limits when two correlated datasets are jointly recovered. We then proceed to evaluate the performance of Singular Value Thresholding (SVT) and Bayesian SVT (BSVT) in this setting. We show that BSVT outperforms SVT by simulating the recovery for different correlated datasets. The performance of BSVT displays the tradeoff behaviour described by the fundamental limit, which suggests that BSVT exploits the correlation between the datasets in an efficient manner.

2. Molecular Communications

Table 13.

Principal Investigators:	Malcolm Egan
Postdoc:	Bayram Akdeniz
Funding:	Inria Projet Recherche Exploratoire (PRE)
Partners:	Valeria Loscri (FUN Team, Inria) Marco Di Renzo (CNRS), Bao Tang (University of Graz, Austria) Trung Duong (Queen's University Belfast) Ido Nevat (TUMCREATE, Singapore)
Publications:	[38], [39], [24], [26]

Some of the most ambitious applications of molecular communications are expected to lie in nanomedicine and advanced manufacturing. In these domains, the molecular communication system is surrounded by a range of biochemical processes, some of which may be sensitive to chemical species used for communication. Under these conditions, the biological system and the molecular communication system impact each other. As such, the problem of coexistence arises, where both the reliability of the molecular communication system and the function of the biological system must be ensured. In this paper, we study this problem with a focus on interactions with biological systems equipped with chemosensing mechanisms, which arises in a large class of biological systems. We motivate the problem by considering chemosensing mechanisms arising in bacteria chemo-taxis, a ubiquitous and well-understood class of biological systems. We then propose strategies for a molecular communication system to minimize disruption of biological system equipped with a chemosensing mechanism. This is achieved by exploiting tools from the theory of chemical reaction networks. To investigate the capabilities of our strategies, we obtain fundamental information theoretic limits by establishing a new connection with the problem of covert communications.

3. Intelligent Transportation

Table 14.

Principal Investigators:	Malcolm Egan
Partners:	Michel Jakob (Czech Technical University in Prague), Nir Oren (University of Aberdeen)
Publications:	[27]

Market mechanisms are now playing a key role in allocating and pricing on-demand transportion services. In practice, most such services use posted-price mechanisms, where both passengers and drivers are offered a journey price which they can accept or reject. However, providers such as Liftago and GrabTaxi have begun to adopt a mechanism whereby auctions are used to price drivers. These latter mechanisms are neither posted-price nor classical double auctions, and can instead be considered a hybrid mechanism. In this work, we develop and study the properties of a novel hybrid on-demand transport mechanism. Due to the need for incorporating statistical knowledge and communication of system state information, communication-theoretic methods can play a useful role.

In particular, as these mechanisms require knowledge of passenger demand, we analyze the data-profit tradeoff as well as how passenger and driver preferences influence mechanism performance. We show that the revenue loss for the provider scales with $\sqrt{n \log n}$ for n passenger requests under a multi-armed bandit learning algorithm with beta distributed preferences. We also investigate the effect of subsidies on both profit and the number of successful journeys allocated by the mechanism, comparing these with a posted-price mechanism, showing improvements in profit with a comparable number of successful requests.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

We have currently the following partnerships

1. Inria-Nokia Bell Labs common lab (600k€) : we are involved in two research actions (Analytics, and Network Information Theory), with the funding of two PhDs and 1 postdoc (to be hired) for Maracas.
2. SPIE-ICS (1Meuros, 2017-2021) : 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. The SPIE-ICS / Insa Lyon chaire on IoT has been setup in 2017 by JM Gorce for the benefit of the CITIlab. JM Gorce was the head of this chair from 2016 to 2019 and is now vice-head (Frédéric Le Mouel is heading the chair since sept 2019). The remaining budget for Maracas corresponds to one postdoc to be hired nad overhead costs.
3. Sigfox : we are collaborating with Sigfox for several years. Maracas explored the performance of UNB networks with an emphasis on robust signal processing techniques (PhD defended on Dec 2018) and a new contract is in preparation for a PhD grant to be started in September, 2020.
4. Orange Labs : our research contract ended in 2018 and we are preparing a new contract.

8.2. Bilateral Grants with Industry

1. PhD grant of Mathieu Goutay (with Nokia Bell Labs, 2019-2022).

9. Partnerships and Cooperations

9.1. Regional Initiatives

- QAMUT *Quantum Algorithms for Multi Users wireless Transmissions* (2019-2021, leader : MARACAS, partners LIP and Institut Camille Jourdan). This project aims to propose new multi-user detection algorithms for wireless transmission systems, based on a quantum architecture.
- *Statistical Hypothesis Testing with Persistent Homology* 2019-2021, leader: MARACAS, partners CRAL. This project aims to develop statistical signal processing methods exploiting persistent homology.

9.2. National Initiatives

9.2.1. ANR

- ANR EPHYL *Enhanced PHY for Cellular Low Power Communication IoT* (2016-2019, 183 keuros, leader : Sequans). This project aims 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.
- ANR ARBURST *Acheivable region of bursty wireless networks* (2016-2020, 195 KEuros, leader : MARACAS). In this project, we propose an original approach complementary to other existing projects, devoted to the study of IoT networks fundamental limits. Instead of proposing one specific technical solution, our objective is to define a unified theoretical framework. 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.
- ANR EquipEx FIT/CorteXlab (2009-2020, 1M€, leader : UPMC). The FIT projet is a national equipex headed by the Lip6 laboratory. As a member of Inria, Maracas is in charge of the development of the Experimental Cognitive Radio platform (CorteXlab) that is used as a testbed 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).

9.2.2. Autres sections...

1. SILECS is a research infrastructure being built to gather the efforts of several testbeds, relying on the success of Grid’5000 and FIT https://www.silecs.net/.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

1. COM-MED, *COMMunication systems with renewable Energy micro-griD*
   - Programm: H2020
   - Duration: October 2016 - October 2019
   - Coordinator: Inria
   - Inria contact: Samir M. Perlaza
   - Summary : 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.
2. WindMill, *Machine Learning for Wireless Communications*
   - Programm: H2020; European Training Network (ETN).
   - Duration: January 2019 - December 2022.
   - Coordinator: Aalborg University, DK
   - Inria contact: Jean-Marie Gorce
   - Summary : With their evolution towards 5G and beyond, wireless communication networks are entering an era of massive connectivity, massive data, and extreme service demands. A promising approach to successfully handle such a magnitude of complexity and data volume is to develop new network management and optimization tools based on machine learning. This is a major shift in the way wireless networks are designed and operated, posing demands for a new type of expertise that requires the combination of engineering, mathematics and computer science disciplines. The ITN project WindMill addresses this need by providing Early Stage Researchers (ESRs) with an expertise integrating wireless communications and machine learning. The project will train 15 ESRs within a consortium of leading international research institutes and companies comprising experts in wireless communications and machine learning. This a very timely project, providing relevant inter-disciplinary training in an area where machine learning represents a meaningful extension of the current methodology used in wireless communication systems. Accordingly, the

project will produce a new generation of experts, extremely competitive on the job market, considering the scale by which machine learning will impact the future and empower the individuals that are versed in it. The project will also nurture the sense of responsibility of the ESRs and the other participants through personal engagement in the training program and by promoting teamwork through collaborative joint projects.

9.3.2. Collaborations in European Programs, Except FP7 & H2020

- Program: PHC Amadeus 2020
- Title: Towards Rigorous Design of Molecular Communication Systems
- Duration: 1/2020 - 12/2021
- Coordinator: Malcolm Egan (MARACAS)
- Other Partners: Institute of Mathematics and Scientific Computing, University of Graz, Austria; CNRS.
- Abstract: The main aim of this project is to bring together experts in molecular communication (Univ. Lyon, Inria, CNRS) and in chemical reaction-diffusion systems (Univ. Graz) to (i) develop novel design of molecular communication systems using up-to-date mathematical results in chemical reaction-diffusion systems, and (ii) strengthen the mathematical theory about chemical reaction networks arising from designation of communication systems.
- Program: COST
- Title: COST Action CA15104, IRACON Inclusive Radio Communications
- Duration: 3/2016 - 3/2020
- Coordinator: Prof. Claude Oestges, University Catholique de Louvain, Belgium.
- Other Partners: many, see website.
- Abstract: This COST Action aims at scientific breakthroughs by introducing novel design and analysis methods for the 5th-generation (5G) and beyond-5G radio communication networks. Challenges include i) modelling the variety of radio channels that can be envisioned for future inclusive radio, ii) capacity, energy, mobility, latency, scalability at the physical layer and iii) network automation, moving nodes, cloud and virtualisation architectures at the network layer, as well as iv) experimental research addressing Over-the-Air testing, Internet of Things, localization and tracking and new radio access technologies. The group of experts supporting this proposal comes from both academia and industry, from a wide spread of countries all over Europe, with the support of some non-COST institutions and R&D associations and standardisation bodies worldwide. The proposers have also long experience on COST Actions in the Radiocommunications field.

9.4. International Initiatives

9.4.1. Inria International Partners

9.4.1.1. Informal 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.
- 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.
- 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.
- UMNG (Universidad Militar de Nueva Granada), Telecommunications Department, Bogota, Colombia. Ongoing collaboration on security for GSM networks using deep learning. Scientific-in-charge at Inria: Leonardo S. Cardoso.
- Department of Power, Electronic and Communication Engineering, University of Novi Sad, Serbia. This collaboration is on GNU radio and signal processing around FIT/CorteXlab with Prof. Dejan Vukobratovic. Scientific-in-charge at Inria: Jean-Marie Gorce.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

9.5.1.1. Internships

1. Huy Duy Do, February-July 2019, Master Thesis MONABIPHOT, ENS Cachan, "Biological Circuits for Detection in Molecular Communication".
2. Antoine Dejonghe, September 2018-July 2020, Telecommunication Department's Research Track, INSA-Lyon, "Techniques for Massive Access in Dense IoT Networks" (Provisional Title)
3. Nuria Vinyes, September 2019-January 2020, Master Thesis, UPC Barcelona, "Simultaneous Information and Energy Transmission: Towards Feasible Systems"
4. Charlotte Hoefler-Hoerle – Undergraduate Student at INSA de Lyon (programme "parcours recherche" de l'INSA de Lyon), Leonardo S. Cardoso and Samir M. Perlaza.
5. INSA de Lyon, D´epartement des T´el´ecommunications. I have advised the following students during their final projects for obtaining the title of Engineer of INSA of Lyon: Samia Bouchareb (2015) and Naslaty Ali Kari (2016), L´elio Chetot (2016), Matias Dwek (2016), and Mamy Niang (2016), Charlotte Hoefler-Hoerle (2019), Adam Ben-Ltaifa (2019), Carl Hatoum (2019).
6. ENS de Lyon, D´epartement d'Informatique. I have advised the following students during their M2-level projects: Lucas Venturini (2019) and Tran Xuan Thang (2019).
7. Matei Catalin Moldoveanu – Master Student at University of Sheffield (Research Intern, Summer 2019).

9.5.2. Visits to International Teams

9.5.2.1. Research Stays Abroad

- Léonardo Cardoso visited Carles Anton, CTTC (Barcelona, Spain), June 2019.

10. Dissemination

10.1. Promoting Scientific Activities

- Interview to Samir M. Perlaza "A view of the Internet of Things", Cognitive Networks Technical Committee Newsletter, IEEE Communications Society, vol. 5, No. 1, May 2019

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- Samir Perlaza: Chair of two special sessions: "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. "Data Analytics for Power Systems" hosted at the 2019 IEEE Data Science Workshop, June 2-5, 2019, Minneapolis, MN, USA.
- Samir Perlaza: Chair of three international Workshops: "Mathematical Tools for IoT Networks Modeling (MOTION)" hosted at the IEEE Wireless Communications and Networking Conference (WCNC). April, 15-18 2019, Marrakech, Morocco. "Resource Allocation, Cooperation and Competition in Wireless Networks (RAWNET)", June 19, 2020, Volvos, Greece. Workshop on Information and Decision Making. A satellite event of the IEEE International Symposium on Information Theory (ISIT), Institut Henri Poincaré, Paris, France, July 10, 2019.
- Samir Perlaza: Publication Chair of the International Symposium on Information Theory (ISIT), July, 2018, Paris, France. (Main conference of the IEEE Information Theory Society)

10.1.1.2. Member of the Organizing Committees

- Leonardo S. Cardoso: Organization and Scientific Commitee of the European GNU Radio Days 2019, June, 2019, Besançon, France.

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

1. Leonardo S. Cardoso
   - IEEE International Conference on Communications (ICC). 7-11 June, 2020, Dublin, Ireland. Member of the Technical Program Committee of the Wireless Communications Symposium.
   - IEEE WCNC 2019, 2020
   - EuCNC 2019, 2020
2. Malcolm Egan
   - IEEE Global Communications Conference (GLOBECOM) 2019, 2020
   - IEEE International Conference on Communications (ICC). 7-11 June, 2020, Dublin, Ireland. Member of the Technical Program Committee of the Wireless Communications Symposium.
   - ACM NanoCommunications Conference (NanCom) 2019, 2020
   - International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS) 2019
3. Jean-Marie Gorce
   - APCC 2019, 2020
   - WF-5G 2019, 2020
   - PIMRC 2019
4. Claire GOURSAUD
   - IEEE ICASSP 2019

5. Samir Perlaza
   - IEEE International Conference on Communications (ICC). 7-11 June, 2020, Dublin, Ireland. Member of the Technical Program Committee of the Wireless Communications Symposium.
   - IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2020) 31 August – 3 September, 2020, London, UK. Member of the Technical Program Committee, Track 3: Mobile and Wireless Networks.
   - International Conference on Computing, Networking and Communications (ICNC). 18-21 February, 2019, Honolulu, Hawaii, USA. Member of the Technical Program Committee.
   - IEEE International Conference on Communications (ICC). 20-24 May 2019, Shanghai, China. Member of the Technical Program Committee of the Wireless Communications Symposium.
   - IEEE Global Communications Conference (Globecom). 9-13 December 2018, Abu Dhabi, UAE. Member of the Technical Program Committee of Workshop on Green and Sustainable 5G Wireless Networks.
   - IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2018) 9-12 September 2018, Bologna, Italy. Member of the Technical Program Committee, Track 3: Mobile and Wireless Networks.

10.1.2.2. Reviewer

1. Malcolm Egan
   - IEEE ISIT 2019
2. Samir Perlaza
   - ISIT, ITW, Eusipco, VTC, Rawnet, WiOpt, GameComm, Globecom, Colcom, PIMRC, Crowncom, WCMC, W-GREEN, ICT, IWCMC, ICC, Infocom, among others.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

1. Malcolm Egan
   - Associate Editor IEEE Communications Letters
   - Guest Editor IEEE Access Special Section on Molecular Communication Networks
2. Jean-Marie Gorce
   - Associate Editor Springer Journal of Wireless Communications and Networking (JWCN)
3. Claire GOURSAUD
   - Associate Editor of European Transactions on Telecommunications (ETT)
   - Associate Editor of Internet Technologies Letters (ITL)
4. Samir Perlaza
   - 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. (One of the main journals of the IEEE Communication Society)
   - Associate Editor of the IET Smart Grid for the term 2018-2021.
   - Guest Editor of the IEEE Internet of Things Journal, Special Issue on Artificial Intelligence Powered Edge Computing for Internet of Things. Publication date: 2020.

10.1.3.2. Reviewer - Reviewing Activities

1. Malcolm Egan

- IEEE Transactions on Information Theory
- IEEE Transactions on Signal Processing
- IEEE Transactions on Wireless Communications
- IEEE Transactions on Communications
- IEEE Transactions on NanoBioscience
- IEEE Transactions on Molecular, Biological, and Multi-Scale Communications

2. Jean-Marie Gorce
   - IEEE Transactions on Wireless Communications
3. Claire Goursaud
   - IEEE Sensors
   - IEEE Transactions on Communications
   - IEEE Wireless Communication Letters
   - IEEE Internet of Things Journal
4. Samir Perlaza
   - IEEE Journal on Selected Areas in Communications, IEEE Journal on Selected Topics in Signal Processing, IEEE Trans. on Wireless Communications, IEEE Trans. on Communications, IEEE Trans on Vehicular Technologies, IEEE Communications Letters, IEEE Trans. on Mobile Computing, EURASIP Journal on Advances in Signal Processing, EURASIP Journal on Wireless Communications and Networking.

10.1.4. Invited Talks

1. Leonardo S. Cardoso
   - “Transmitter Classification with Supervised Deep Learning”. Invited talk at CTTC, June 29, 2019, Barcelona, Spain.
   - “Exploring the Radio Spectrum with GNU Radio”. Invited tutorial at European GNU Radio Days , June 19, 2019, Besançon, France..
2. Malcolm Egan
   - Invited Lecture, European School on Information Theory, April 15-19, Nice, 2019.
   - Invited Lecture, International Workshop on Mathematical Tools and Technologies for IoT and mMTC Networks Modeling, 15-18 April, Morocco, 2019.
   - Invited Paper, Third International Balkan Conference on Communications and Networking 2019, 10-12 June, North Macedonia, 2019.
3. Jean-Marie Gorce
   - Invited Paper, International Workshop on Mathematical Tools and Technologies for IoT and mMTC Networks Modeling, 15-18 April, Morocco, 2019.
   - Invited Paper, Special Session on Ultra-Reliable Low-Latency Communicationsgo to top at the 20th international workshop on signal processing advances in wireless communications, 2-5 July, France, 2019.
   - Invited talk entitled *Etat de l’art, focus sur la mobilité et prospectives des communications pour l’IoT*; lors de l’atelier IdéO *Quelles infrastructures numériques et quel IoT pour une mobilité intelligente ?* , organisé par les pôles de compétitivité Systematic and Moveo (regional poles), Paris, Jeudi 26 Sept 2019.
4. Samir Perlaza
   - “Transforming Broadcast Codes to Perform Covert Communications”. Invited talk at Department of Automatic Control and Systems Engineering, University of Sheffield, November 26, 2019, Sheffield, UK.

- “Transforming Broadcast Codes to Perform Covert Communications”. Invited talk at Centre d’Enseignement et de Recherche en Informatique (CERI) Université d’Avignon, November 14, 2019, Avignon, France.
- “Simultaneous Information and Energy Transmission in Decentralized Networks”. Invited talk at Inria, Centre de Recherche Sophia Antipolis Méditerranée, September 24, 2019, Sophia Antipolis, France.
- “Simultaneous Information and Energy Transmission Systems”. Invited talk at Eurecom, September 26 2019, Sophia Antipolis, France.
- “On Ultra-Reliable and Low Latency Simultaneous Information and Energy Transmission Systems”. Invited talk at Lappeenranta University of Technology, Finland. School of Energy Systems. April 08 2019. Lappeenranta, Finland
- “Information-Theoretic Security in the Smart Grid”. Invited talk at Lappeenranta University of Technology, Finland. School of Energy Systems. April 07, 2019, Lappeenranta, Finland.
- “Simultaneous Wireless Information and Energy Transmission”. Invited talk at Université de Lille. Institut d’électronique de microélectronique et de nanotechnologie (IEMN), October 4, 2018, Lille, France.

10.1.5. Scientific Expertise

- Jean-Marie Gorce
  - Chair of the evaluation committee for ETIS lab (Cergy) for HCERES.
  - Evaluator of research projects for Region Picardie.
  - Jury member of the senior researchers recruitment competition, Inria.

10.1.6. Research Administration

- Jean-Marie Gorce was:
  - Vice-head for research for the Grenoble Rhone-Alpes research centre.
  - Member (representative for Inria) of the Groupe Academique for University of Lyon.
  - Member of the national evaluation committee, Inria.
- Claire Goursaud was:
  - elected as a CNU member in the 61th section for 2019-2024
- Samir Perlaza was:
  - Member of the On-Line Committee of the IEEE Information Theory Society.
  - Membre du Conseil du Laboratoire CITI à l’INSA de Lyon. Term 2015 - 2019.
  - Committé du Prix Paul CASEAU for the 2017 edition of the thesis award.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master : JM Gorce, M Egan, L Chetot Advanced Digital Communications, 40h eq TD, M1, Telecommunications dept, INSA Lyon, France.

Master : JM Gorce, M Egan, D Duchemin Radio Access Networs, 32h eqTD, M1, Telecommunications dept, INSA Lyon, France.

Master : S Perlaza, JM Gorce, Selected Topics in Information Theory, 32h eqTD, M2, Compute sciences dept, ENS Lyon, France.

Master : L Cardoso, C Morin, Software Radio, 32h eqTD, M2, Telecommunications dept, INSA Lyon, France.

Licence : L Cardoso, C Goursaud, Digital Communications, 80h eqTD, L3, Telecommunications dept, INSA Lyon, France.

Licence : L Cardoso, C Goursaud, Research projetcs, 32h eqTD, L3, Telecommunications dept, INSA Lyon, France.

Master : C Goursaud, Communications Systems, 32h eqTD, M1, Telecommunications dept, INSA Lyon, France.

License : L Cardoso, Electromagnetism and Wave Physics, 104 eqTD, L2, First Cycle Dept, INSA Lyon, France.

License : L Cardoso, Mathematics for Engineering, 60h eqTD, L1, First Cycle Dept, INSA Lyon, France.

10.2.2. Supervision

PhD : David Kibloff, Contributions Théoriques sur les Communications Furtives, INSA Lyon, 17 sept 2019, Samir Perlaza and Guillaume Villemaud.

PhD: Trang Mai, Advanced Technologies in Molecular and Wireless Communication Networks: Analysis, Design and Optimization, Queen's University Belfast, 30 Dec 2019, Trung Duong, Malcolm Egan, and Hien Quoc Ngo.

PhD in progress : Dadja Toussaint Anade-Akpo, Non-asymptotic fundamental limits of impulsive radio communications, 01 oct 2017, Philippe Mary and Jean-Marie Gorce.

PhD in progress : Lélio Chetot, From finite blocklength information theory to multi-user M2M communication protocols, 01 oct 2018, Malcolm Egan and Jean-Marie Gorce.

PhD in progress : Diane Duchemin, Distributed coding for dense IoT networks, October 2016, Claire Goursaud and Jean-Marie Gorce

PhD in progress : Mathieu Goutay, Prédistortion digitale profonde, 01 Feb 2019, Jakob Hoydis and Jean-Marie Gorce.

PhD in progress : Hassan Kallam, Topology aided Multi-User Interference Management in Wireless Networks, 01 Feb 2017, Loonardo S. Cardoso and Jean-Marie Gorce.

PhD in progress : José Rugelles, Deep Learning for Security in GSM Based IoT Systems, 01 Jan 2017, Loonardo S. Cardoso and Edward Guillén.

PhD in progress : Nizar Khalfet, Study of stochastique energy sources to power communication system, 07 nov 2016, Samir Perlaza and Jean-Marie Gorce.

PhD in progress : Cyrille Morin, Deep learning for next generation communication systems, 20 Feb 2018, Leonardo S. Cardoso and Jean-Marie Gorce.

PhD in progress : Ce Zheng, Statistical models for IoT interference, Commenced Oct 2018, Malcolm Egan and Laurent Clavier.

Xiuzhen Ye - Academic Period: 2019 - 2022 – University of Sheffield, Department of Automatic Control and Systems Engineering, Sheffield, UK. Taux d'encadrement : Iñaki Esnaola (University of Sheffield, UK) 50%; and Rob Harrisson (University of Sheffield, UK) 25% and Samir Perlaza (external advisor) 25% Publications: Several under preparation. Thesis Title: Data Injection Attacks in Power Systems

Michalis Eliodorous – PhD Student at University of Cyprus (on-going collaboration during during his PhD)

10.2.3. Juries

1. Leonardo Cardoso
   - Examiner; Thesis Jury for Sumit Kumar supervised by Florian Kaltenberger in Institut Eurecom (Telecom Paristech), Sophia Antipolis, 12 March 2019.

- Examiner; Thesis Jury for Nicolas Araújo supervised by Laurent Clavier in the University de Lille, Lille, 11 July 2019.

2. Malcolm Egan
   - Examiner; Thesis Jury for Pierre Escamilla supervised by Michele Wigger and Abdellatif Zaidi in Telecom ParisTech
   - Invited; Thesis Jury for Yasser Mestrah supervised by Laurent Clavier in the University of Reims Champagne-Ardenne and IMT Lille Douai
3. Jean-Marie Gorce
   - Reviewer; Thesis jury for Gourab Ghatak supervised by Antonio De Domenico (CEA) and Marceau Coupechoux, Telecoms Paris, IMT. 24 janv 2019.
   - Chair; Thesis jury for Alexandre Marcastel, supervised by Veronica Belmega ,Panayotis Mertikopoulos and Inbar Fijalkow, ETIS, Cergy. 21 Feb 2019.
   - Reviewer; HdR jury for Veronica Belmega , ETIS, Cergy. 29 March 2019.
   - Reviewer; HdR jury for Thomas Watteyne, Inria, Sorbonne University. 7 May 2019.
   - Examiner; Thesis jury for Abir Ben Hadj Fredj, supervised by Jean-Claude Belfiore and Ghaya Rekaya-Ben Othman, Telecoms Paris, IMT. 28 June 2019.
   - Examiner; Thesis jury for David Kibloff, supervised by Samir Perlaza and Guillaume Villemaud, INSA Lyon. 17 Sept 2019.
   - Reviewer; HdR jury for Marios Kountouris, Telecoms Paris, IMT. 17 Oct 2019.
   - Reviewer; Thesis jury for Selma Zamoum, supervised by Marie-Laure Boucheret and Jérôme Lacan, Université de Toulouse. 28 Nov 2019.
   - Reviewer; Thesis jury for Kaoutar Abdelalim supervised by Karine Amis and Getachew Redieteab, IMT Atlantique. 9 Dec 2019.
   - Reviewer; Thesis jury for Ibrahim Fawaz, supervised by Philippe Ciblat and Mireille Sarkiss, Telecoms Paris, IMT. 9 Dec 2019.
   - Reviewer; Thesis jury for Marc Kacou, supervised by Valery Guillet, Gheorghe Zaharia, and Ghais El Zein, INSA Rennes. 12 Dec 2019.
   - Reviewer; Thesis jury for Xiaojun Xi supervised by Marco Di Renzo, Supelec, Université Paris Saclay. 19 Dec 2019.
   - Examiner; Thesis jury for Jian Song supervised by Marco Di Renzo, Supelec, Université Paris Saclay. 19 Dec 2019.
4. Claire Goursaud
   - Reviewer for Alex The Phuong Nguyen supervised by Frederic Guilloud and Raphaêl Le Bidan, IMT Atlantique
5. Samir Perlaza
   - Miguel Arrieta, Phd Thesis “Universal Privacy Guarantees for Smart Meters” at the Department of Automatic Control and Systems Engineering, University of Sheffield, November 29 2019, Sheffield, UK.
   - Ms. Nihan Cicek, Mphil Thesis "Probabilistic Energy Management Systems in PV-Rich Communities” Electrical & Electronic Engineering, The University of Melbourne, Australia

11. Bibliography

Major publications by the team in recent years

[1] G. C. Alexandropoulos, P. Ferrand, J.-M. Gorce, C. B. Papadias.*Advanced coordinated beamforming for the downlink of future LTE cellular networks*, in "IEEE Communications Magazine", July 2016, vol. 54, n^o^ 7, p. 54 - 60, Arxiv: 16 pages, 6 figures, accepted to IEEE Communications Magazine [*DOI :* 10.1109/MCOM.2016.7509379], https://hal.inria.fr/hal-01395615

[2] S. BELHADJ AMOR, S. PERLAZA, I. KRIKIDIS, H. V. POOR.*Feedback Enhances Simultaneous Wireless Information and Energy Transmission in Multiple Access Channels*, in "IEEE Transactions on Information Theory", August 2017, vol. 63, n^{o} 8, p. 5244 - 5265 [*DOI :* 10.1109/TIT.2017.2682166], https://hal.inria.fr/hal-01857373

[3] M. DE FREITAS, M. EGAN, L. CLAVIER, A. GOUPIL, G. W. PETERS, N. AZZAOUI.*Capacity Bounds for Additive Symmetric α -Stable Noise Channels*, in "IEEE Transactions on Information Theory", August 2017, vol. 63, n^{o} 8, p. 5115-5123 [*DOI :* 10.1109/TIT.2017.2676104], https://hal.univ-reims.fr/hal-02088563

[4] 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

[5] 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

[6] M. EGAN, T. Q. DUONG, M. DI RENZO, J.-M. GORCE, I. NEVAT, V. LOSCRI.*Cognitive Molecular Communication*, in "3rd Workshop on Molecular Communications", 2018

[7] 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

[8] M. EGAN, S. PERLAZA, V. KUNGURTSEV.*Capacity sensitivity in additive non-gaussian noise channels*, in "2017 IEEE International Symposium on Information Theory (ISIT)", IEEE, 2017, p. 416–420

[9] I. ESNAOLA, S. PERLAZA, H. V. POOR, O. KOSUT.*Maximum Distortion Attacks in Electricity Grids*, in "IEEE Transactions on Smart Grid", 2016, vol. 7, n^{o} 4, p. 2007-2015 [*DOI :* 10.1109/TSG.2016.2550420], https://hal.archives-ouvertes.fr/hal-01343248

[10] Y. FADLALLAH, A. M. TULINO, D. BARONE, G. VETTIGLI, J. LLORCA, J.-M. GORCE.*Coding for Caching in 5G Networks*, in "IEEE Communications Magazine", February 2017, vol. 55, n^{o} 2, p. 106 - 113 [*DOI :* 10.1109/MCOM.2017.1600449CM], https://hal.inria.fr/hal-01492353

[11] C. GENES, I. ESNAOLA, S. PERLAZA, L. F. OCHOA, D. COCA.*Robust Recovery of Missing Data in Electricity Distribution Systems*, in "IEEE Transactions on Smart Grid", 2018

[12] J.-M. GORCE, Y. FADLALLAH, J.-M. KELIF, H. V. POOR, A. GATI.*Fundamental limits of a dense iot cell in the uplink*, in "Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2017 15th International Symposium on", IEEE, 2017, p. 1–6

[13] C. GOURSAUD, J.-M. GORCE.*Dedicated networks for IoT : PHY / MAC state of the art and challenges*, in "EAI endorsed transactions on Internet of Things", October 2015 [*DOI :* 10.4108/EAI.26-10-2015.150597], https://hal.archives-ouvertes.fr/hal-01231221

[14] A. GUIZAR, C. GOURSAUD, J.-M. GORCE.*Performance of IR-UWB cross-layer ranging protocols under on-body channel models with body area networks*, in "Annals of Telecommunications - annales des

télécommunications", March 2016, p. 453–46, http://link.springer.com/article/10.1007/s12243-016-0500-4 [*DOI :* 10.1007/s12243-016-0500-4], https://hal.archives-ouvertes.fr/hal-01290211

[15] T. C. MAI, M. EGAN, T. Q. DUONG, M. DI RENZO.*Event Detection in Molecular Communication Networks with Anomalous Diffusion*, in "IEEE Communications Letters", February 2017, vol. 21, n^{o} 6, p. 1249 - 1252 [*DOI :* 10.1109/LCOMM.2017.2669315], https://hal.archives-ouvertes.fr/hal-01671181

[16] Y. MO, M.-T. DO, C. GOURSAUD, J.-M. GORCE.*Up-Link Capacity Derivation for Ultra-Narrow-Band IoT Wireless Networks*, in "International Journal of Wireless Information Networks", June 2017, vol. 24, n^{o} 3, p. 300-316 [*DOI :* 10.1007/s10776-017-0361-4], https://hal.inria.fr/hal-01610466

[17] S. PERLAZA, A. TAJER, H. V. POOR.*Simultaneous Energy and Information Transmission: A Finite Block-Length Analysis*, in "IEEE International Workshop on Signal Processing Advances in Wireless Communications", 2018

[18] V. QUINTERO, S. PERLAZA, I. ESNAOLA, J.-M. GORCE.*When Does Output Feedback Enlarge the Capacity of the Interference Channel?*, in "IEEE Transactions on Communications", September 2017, vol. 66, n^{o} 2, p. 615-628, Part of this work was presented at the 11th EAI International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM), Grenoble, France, May 30-Jun 1 2016 [*DOI :* 10.1109/TCOMM.2017.2753252], https://hal.archives-ouvertes.fr/hal-01432525

[19] V. QUINTERO, S. PERLAZA, I. ESNAOLA, J.-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

[20] D. TSILIMANTOS, J.-M. GORCE, K. JAFFRÈS-RUNSER, H. V. POOR.*Spectral and Energy Efficiency Trade-Offs in Cellular Networks*, in "IEEE Transactions on Wireless Communications", January 2016, vol. 15, n^{o} 1, p. 54-66 [*DOI :* 10.1109/TWC.2015.2466541], https://hal.inria.fr/hal-01231819

Publications of the year

Articles in International Peer-Reviewed Journal

[21] F. DINC, B. CEVDET AKDENIZ, E. EROL, D. GOKAY, E. TEKGUL, A. EMRE PUSANE, T. TUGCU.*Analytical derivation of the impulse response for the bounded 2-D diffusion channel*, in "Modern Physics Letters A", 2019, vol. 383, p. 1589 - 1600 [*DOI :* 10.1016/J.PHYSLETA.2019.02.025], https://hal.inria.fr/hal-02416000

[22] F. DINC, B. CEVDET AKDENIZ, A. E. PUSANE, T. TUGCU.*A General Analytical Approximation to Impulse Response of 3-D Microfluidic Channels in Molecular Communication*, in "IEEE Transactions on NanoBioscience", July 2019, vol. 18, n^{o} 3, p. 396-403 [*DOI :* 10.1109/TNB.2019.2905417], https://hal.inria.fr/hal-02415993

[23] F. DINC, L. THIELE, B. CEVDET AKDENIZ.*The effective geometry Monte Carlo algorithm: Applications to molecular communication*, in "Modern Physics Letters A", August 2019, vol. 383, p. 2594 - 2603 [*DOI :* 10.1016/J.PHYSLETA.2019.05.029], https://hal.inria.fr/hal-02415980

[24] M. EGAN, T. Q. DUONG, M. D. RENZO.*Biological Circuits for Detection in MoSK-Based Molecular Communication*, in "IEEE Access", 2019, p. 21094 - 21102, forthcoming [*DOI :* 10.1109/ACCESS.2019.2897173], https://hal.archives-ouvertes.fr/hal-02001903

[25] M. EGAN.*Properties of Tight Frames that are Regular Schemes*, in "Cryptography and Communications - Discrete Structures, Boolean Functions and Sequences", 2019, p. 1-15, forthcoming [*DOI :* 10.1007/S12095-019-00378-2], https://hal.inria.fr/hal-02148974

[26] M. EGAN, V. LOSCRI, T. Q. DUONG, M. D. RENZO.*Strategies for Coexistence in Molecular Communication*, in "IEEE Transactions on NanoBioscience", January 2019, vol. 18, n^{o} 1, p. 51-60 [*DOI :* 10.1109/TNB.2018.2884999], https://hal.archives-ouvertes.fr/hal-01928205

[27] M. EGAN, N. OREN, M. JAKOB.*Hybrid Mechanisms for On-Demand Transport*, in "IEEE Transactions on Intelligent Transportation Systems", January 2019, p. 1-13 [*DOI :* 10.1109/TITS.2018.2886579], https://hal.inria.fr/hal-01950684

[28] F. GOKBUDAK, B. CEVDET AKDENIZ, T. TUGCU, A. E. PUSANE.*Spatial Receptor Allocation for a Multiple Access Hub in Nanonetworks*, in "IEEE Transactions on Molecular, Biological and Multi-Scale Communications", October 2019 [*DOI :* 10.1109/TMBMC.2019.2946819], https://hal.inria.fr/hal-02415965

[29] N. KHALFET, S. PERLAZA.*Simultaneous Information and Energy Transmission in the Two-User Gaussian Interference Channel*, in "IEEE Journal on Selected Areas in Communications", January 2019, vol. 37, n^{o} 1, p. 156-170 [*DOI :* 10.1109/JSAC.2018.2872365], https://hal.archives-ouvertes.fr/hal-01874019

Invited Conferences

[30] D. DUCHEMIN, L. CHETOT, J.-M. GORCE, C. GOURSAUD.*Coded random access for massive MTC under statistical channel knowledge*, in "SPAWC 2019 - 20th IEEE International Workshop on Signal Processing Advances in Wireless Communications", Cannes, France, IEEE, July 2019, p. 1-5 [*DOI :* 10.1109/SPAWC.2019.8815491], https://hal.inria.fr/hal-02153735

[31] D. DUCHEMIN, J.-M. GORCE, C. GOURSAUD.*Low complexity Detector for massive uplink random access with NOMA in IoT LPWA networks*, in "WCNC 2019 - IEEE Wireless Communications and Networking Conference", Marrakech, Morocco, IEEE, April 2019, p. 1-6, https://hal.inria.fr/hal-02146649

[32] M. EGAN, B. CEVDET AKDENIZ.*On the Limits of Molecular Communication with Coexisting Biological Systems Towards the Internet of Bio-Nano Things*, in "BalkanCom 2019 - Third International Balkan Conference on Communications and Networking", Skopje, Macedonia, June 2019, p. 1-5, https://hal.inria.fr/hal-02143434

[33] M. EGAN.*On Capacity Sensitivity in Additive Vector Symmetric α-Stable Noise Channels*, in "WCNC 2019 - IEEE Wireless Communications and Networking Conference", Marrakech, Morocco, IEEE, April 2019, p. 1-6, https://hal.inria.fr/hal-02126849

International Conferences with Proceedings

[34] D. T. ANADE AKPO, J.-M. GORCE, P. MARY.*Nouvelle borne atteignable de la probabilité d'erreur pour des transmissions en paquets courts*, in "GRETSI 2019 - XXVIIème Colloque francophonede traitement du signal et des images", Lille, France, August 2019, p. 1-4, https://hal.inria.fr/hal-02428615

[35] L. CHETOT, J.-M. GORCE, J.-M. KELIF.*Fundamental Limits in Cellular Networks with Point Process Partial Area Statistics*, in "WiOpt 2019 - 17th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks", Avignon, France, June 2019, p. 1-8, https://hal.inria.fr/hal-02146103

[36] M. DE FREITAS, M. EGAN, L. CLAVIER, A. SAVARD, J.-M. GORCE.*Power Control in Parallel Symmetric α-Stable Noise Channels*, in "SPAWC 2019 - 20th IEEE International Workshop on Signal Processing Advances in Wireless Communications", Cannes, France, IEEE, July 2019, p. 1-5 [*DOI :* 10.1109/SPAWC.2019.8815443], https://hal.inria.fr/hal-02130226

[37] D. DUCHEMIN, L. CHETOT, J.-M. GORCE, C. GOURSAUD.*Détecteur pour l'accès aléatoire massif entre machines avec connaissance statistique du canal en lien ascendant*, in "GRETSI 2019", Lille, France, August 2019, p. 1-4, https://hal.inria.fr/hal-02297508

[38] M. EGAN, T. Q. BAO, B. CEVDET AKDENIZ.*On the Input-Output Relationship for Molecular Communications in General First-Order Chemical Reaction-Diffusion Systems*, in "NanoCom 2019 - Sixth Annual ACM International Conference on Nanoscale Computing and Communication", Dublin, Ireland, ACM, September 2019, p. 1-6 [*DOI :* 10.1145/3345312.3345477], https://hal.inria.fr/hal-02304088

[39] M. EGAN, V. LOSCRI, I. NEVAT, T. Q. DUONG, M. DI RENZO.*Estimation and Optimization for Molecular Communications with a Coexistence Constraint*, in "NanoCom 2019 - Sixth Annual ACM International Conference on Nanoscale Computing and Communication", Dublin, Ireland, ACM, September 2019, vol. NanoCom 2019 - Sixth Annual ACM International Conference on Nanoscale Computing and Communication, p. 1-6 [*DOI :* 10.1145/3345312.3345472], https://hal.archives-ouvertes.fr/hal-02178421

[40] C. GENES, I. ESNAOLA, S. PERLAZA, D. COCA.*Recovery of missing data in correlated smart grid datasets*, in "DSW 2019 - Second IEEE Data Science Workshop", Minneapolis, United States, IEEE, June 2019, vol. DSW 2019 - Second IEEE Data Science Workshop, p. 263-269 [*DOI :* 10.1109/DSW.2019.8755592], https://hal.archives-ouvertes.fr/hal-02165164

[41] H. KALLAM, L. CARDOSO, J.-M. GORCE.*Topological Interference Management: Trade-off Between DoF and SIR for Cellular Systems*, in "ICT 2019 - 26th International Conference on Telecommunications", Hanoi, Vietnam, April 2019, p. 96-101 [*DOI :* 10.1109/ICT.2019.8798819], https://hal.inria.fr/hal-02364450

[42] N. KHALFET, S. PERLAZA.*On the Maximum Energy Transmission Rate in Ultra-Reliable and Low Latency SIET*, in "BalkanCom 2019 - Third International Balkan Conference on Communications and Networking", Skopje, Macedonia, BalkanCom 2019 - Third International Balkan Conference on Communications and Networking, June 2019, p. 1-4, https://hal.archives-ouvertes.fr/hal-02138312

[43] N. KHALFET, S. PERLAZA, A. TAJER, H. VINCENT POOR.*On Ultra-Reliable and Low Latency Simultaneous Information and Energy Transmission Systems*, in "SPAWC 2019 - 20th IEEE International Workshop on Signal Processing Advances in Wireless Communications", Cannes, France, SPAWC 2019 - 20th IEEE International Workshop on Signal Processing Advances in Wireless Communications, IEEE, July 2019, p. 1-5 [*DOI :* 10.1109/SPAWC.2019.8815395], https://hal.archives-ouvertes.fr/hal-02136606

[44] D. KIBLOFF, S. PERLAZA, L. WANG.*Embedding Covert Information on a Given Broadcast Code*, in "ISIT 2019 - IEEE International Symposium on Information Theory", Paris, France, July 2019, vol. ISIT 2019 - IEEE International Symposium on Information Theory, p. 1-5, https://hal.archives-ouvertes.fr/hal-02135854

[45] C. MORIN, L. CARDOSO, J. HOYDIS, J.-M. GORCE, T. VIAL.*Transmitter Classification With Supervised Deep Learning*, in "CROWNCOM 2019 - 14th EAI International conference on Cognitive Radio Oriented Wireless Networks", Poznan, Poland, June 2019, p. 1-14, https://arxiv.org/abs/1905.07923 , https://hal.inria.fr/hal-02132970

[46] O. OUBEJJA, D. DUCHEMIN, M. IMBERT, L. CARDOSO, J.-M. GORCE.*Framework for PHY-MAC layers Prototyping in Dense IoT Networks using FIT/CorteXlab Testbed*, in "CNERT 2019 - International Workshop on Computer and Networking Experimental Research using Testbeds", Paris, France, IEEE, April 2019, p. 1-2, https://hal.inria.fr/hal-02150687

[47] G. VIVIER, V. MANNONI, F. DEHMAS, V. BERG, L. MROUEH, Y. YU, D. DUCHEMIN, J.-M. GORCE, J. MANCO-VASQUEZ, F. BADER.*Beyond LoRa and NB-IoT: Proposals for Future LPWA Systems*, in "EuCNC 2019 - European Conference on Networks and Communications", Valencia, Spain, IEEE (editor), EuCNC 2019 - European Conference on Networks and Communications, June 2019, p. 523-527 [*DOI :* 10.1109/EUCNC.2019.8802007], https://hal.archives-ouvertes.fr/hal-02106227

[48] C. C. ZHENG, M. EGAN, L. CLAVIER, G. W. PETERS, J.-M. GORCE.*Copula-Based Interference Models for IoT Wireless Networks*, in "ICC 2019 - 53rd IEEE International Conference on Communications", Shanghai, China, Proceedings of the 53rd IEEE International Conference on Communications, IEEE, May 2019, p. 1-6 [*DOI :* 10.1109/ICC.2019.8761783], https://hal.archives-ouvertes.fr/hal-02065780

[49] C. C. ZHENG, M. EGAN, L. CLAVIER, G. W. PETERS, J.-M. GORCE.*On the Validity of Isotropic Complex α-Stable Interference Models for Interference in the IoT*, in "GRETSI 2019 - Colloque francophone de traitement du signal et des images", Lille, France, August 2019, p. 1-4, https://hal.archives-ouvertes.fr/hal-02263787

Research Reports

[50] N. KHALFET, S. PERLAZA, A. TAJER, H. VINCENT POOR.*On Ultra-Reliable and Low Latency Simultaneous Information and Energy Transmission Systems*, Inria Grenoble Rhône-Alpes, June 2019, n^{o} RR-9261, p. 1-30, https://hal.archives-ouvertes.fr/hal-02161130

[51] D. KIBLOFF, S. PERLAZA, L. WANG.*Broadcast Codes Can Be Enhanced to Perform Covert Communications*, Inria Grenoble - Rhône-Alpes, January 2019, n^{o} RR-9249, p. 1-70, https://hal.archives-ouvertes.fr/hal-01991847

References in notes

[52] *CorteXlab provided docker images*, 2019, https://wiki.cortexlab.fr/doku.php?id=docker_images

[53] M. R. BLOCH.*Covert communication over noisy channels: A resolvability perspective*, in "IEEE Transactions on Information Theory", 2016, vol. 62, n^{o} 5, p. 2334–2354

[54] S. DÖRNER, S. CAMMERER, J. HOYDIS, S. TEN BRINK.*Deep learning based communication over the air*, in "IEEE Journal of Selected Topics in Signal Processing", 2018, vol. 12, n^{o} 1, p. 132–143

[55] M. C. FILIPPOU, P. DE KERRET, D. GESBERT, T. RATNARAJAH, A. PASTORE, G. A. ROPOKIS.*Coordinated shared spectrum precoding with distributed CSIT*, in "IEEE Transactions on Wireless Communications", 2016, vol. 15, n^{o} 8, p. 5182–5192

[56] S. A. JAFAR.*Topological interference management through index coding*, in "IEEE Transactions on Information Theory", 2014, vol. 60, n^o 1, p. 529–568

[57] M. G. KHOSHKHOLGH, K. NAVAIE, K. G. SHIN, V. LEUNG, H. YANIKOMEROGLU.*Caching or No Caching in Dense HetNets?*, in "arXiv preprint arXiv:1901.11068", 2019

[58] S. LI, M. A. MADDAH-ALI, Q. YU, A. S. AVESTIMEHR.*A fundamental tradeoff between computation and communication in distributed computing*, in "IEEE Transactions on Information Theory", 2018, vol. 64, n^o 1, p. 109–128

[59] W. LIU, S. XUE, J. LI, L. HANZO.*Topological Interference Management for Wireless Networks*, in "IEEE Access", 2018, vol. 6, p. 76942–76955

[60] Y. MAO, C. YOU, J. ZHANG, K. HUANG, K. B. LETAIEF.*A survey on mobile edge computing: The communication perspective*, in "IEEE Communications Surveys & Tutorials", 2017, vol. 19, n^o 4, p. 2322–2358

[61] D. MERKEL.*Docker: Lightweight Linux Containers for Consistent Development and Deployment*, in "Linux J.", March 2014, vol. 2014, n^o 239

[62] Y. POLYANSKIY, H. V. POOR, S. VERDÚ.*Channel coding rate in the finite blocklength regime*, in "IEEE Transactions on Information Theory", 2010, vol. 56, n^o 5, 2307

[63] J. SACHS, L. A. A. ANDERSSON, J. ARAÚJO, C. CURESCU, J. LUNDSJÖ, G. RUNE, E. STEINBACH, G. WIKSTRÖM.*Adaptive 5G Low-Latency Communication for Tactile InternEt Services*, in "Proceedings of the IEEE", Feb 2019, vol. 107, n^o 2, p. 325-349

[64] V. Y. TAN.*Asymptotic estimates in information theory with non-vanishing error probabilities*, in "Foundations and Trends® in Communications and Information Theory", 2014, vol. 11, p. 1–184

[65] G. VAZQUEZ-VILAR, A. G. I FABREGAS, T. KOCH, A. LANCHO.*Saddlepoint approximation of the error probability of binary hypothesis testing*, in "2018 IEEE International Symposium on Information Theory (ISIT)", IEEE, 2018, p. 2306–2310

[66] Q. YAN, S. YANG, M. WIGGER.*Storage, computation, and communication: A fundamental tradeoff in distributed computing*, in "2018 IEEE Information Theory Workshop (ITW)", IEEE, 2018, p. 1–5

[67] X. YI, G. CAIRE.*Topological interference management with decoded message passing*, in "IEEE Transactions on Information Theory", 2018, vol. 64, n^o 5, p. 3842–3864

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, HDR]
Cyril Soler [Inria, Researcher, HDR]

Faculty Members

Georges-Pierre Bonneau [Univ Grenoble Alpes, Professor, HDR]
Joëlle Thollot [Institut polytechnique de Grenoble, Professor, HDR]
Romain Vergne [Univ Grenoble Alpes, Associate Professor]

PhD Students

Alban Fichet [Inria, PhD Student]
Morgane Gerardin [Univ Grenoble Alpes, PhD Student]
Nolan Mestres [Univ Grenoble Alpes, PhD Student, from Oct 2019]
Ronak Molazem [Inria, PhD Student]
Vincent Tavernier [Univ Grenoble Alpes, PhD Student]
Sunrise Wang [Inria, PhD Student]

Visiting Scientist

Hugo Rens [Univ Paul Sabatier, from May 2019 until Jul 2019]

Administrative Assistant

Diane Courtiol [Inria, Administrative Assistant]

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.

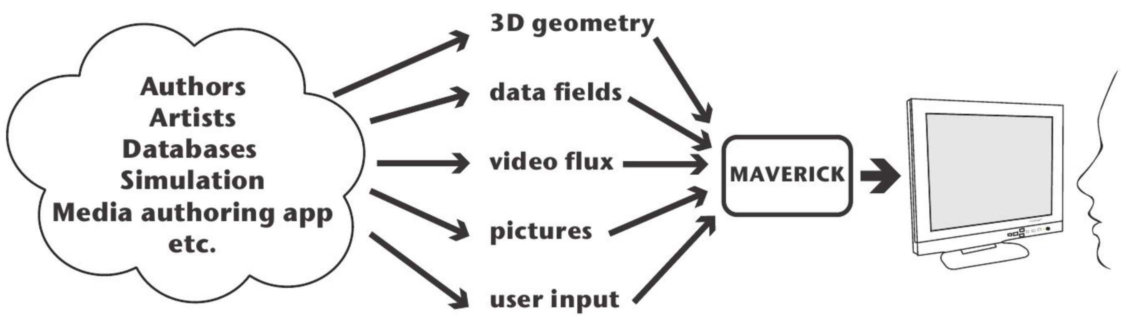


Figure 1. Position of the Maverick research team inside the graphics pipeline.

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 experience of collaboration with industrial partners. These collaborations bring us new problems to solve, with short-term or medium-term transfer 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. New Software and Platforms

5.1. 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/

5.2. 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

5.3. 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

5.4. 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

5.5. Vrender

KEYWORDS: 3D - Vector graphics

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://maverick.inria.fr/Software/VRender/

5.6. X3D TOOLKIT

X3D Development pateform

KEYWORDS: X3D - Geometric modeling

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://maverick.inria.fr/Software/X3D/

5.7. 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. New Results

6.1. Texture synthesis

6.1.1. Procedural Phasor Noise

Participants: Thibault Tricard, Semyon Efremov, Cédric Zanni, Fabrice Neyret, Jonàs Martínez, Sylvain Lefebvre.

Procedural pattern synthesis is a fundamental tool of Computer Graphics, ubiquitous in games and special effects. By calling a single procedure in every pixel – or voxel – large quantities of details are generated at low cost, enhancing textures, producing complex structures within and along surfaces. Such procedures are typically implemented as pixel shaders. We propose a novel procedural pattern synthesis technique that exhibits desirable properties for modeling highly contrasted patterns, that are especially well suited to produce surface and microstructure details. In particular, our synthesizer affords for a precise control over the profile, orientation and distribution of the produced stochastic patterns, while allowing to grade all these parameters spatially. Our technique defines a stochastic smooth phase field – a phasor noise – that is then fed into a periodic function (e.g. a sine wave), producing an oscillating field with prescribed main frequencies and preserved contrast oscillations. In addition, the profile of each oscillation is directly controllable as shown Figure 2. Our technique builds upon a reformulation of Gabor noise in terms of a phasor field that affords for a clear separation between local intensity and phase. Applications range from texturing to modeling surface displacements, as well as multi-material microstructures in the context of additive manufacturing.

This paper was published in ACM TOG [6] and presented at Siggraph 2019.

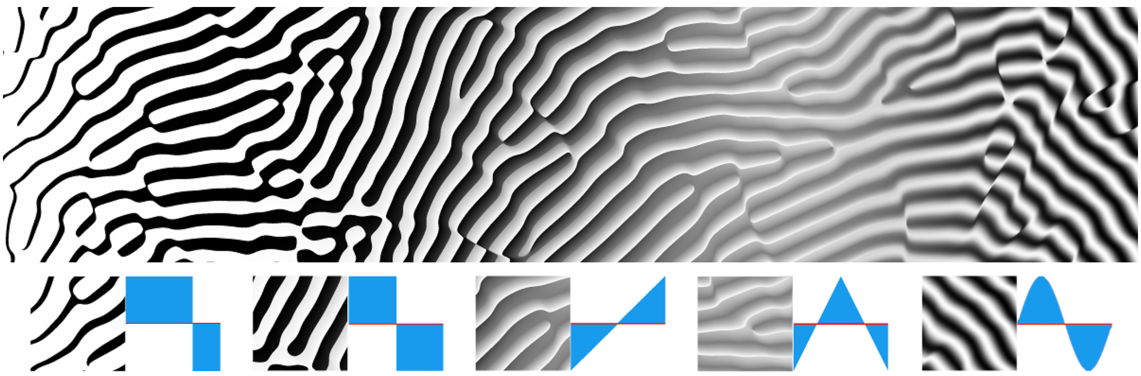


Figure 2. High-contrast patterns produced by our approach. Note how the profile of the oscillations smoothly transition from a rectangular wave (20% black), to a square wave, to a triangular profile and finally a sine wave. At the same time, the orientation of the waves changes from left to right. The field visualized here is purely procedural. It is obtained by feeding our phasor noise into periodic profile functions (shown in blue), that are interpolated from left to right.

6.1.2. Making Gabor Noise Fast and Normalized

Participants: Vincent Tavernier, Fabrice Neyret, Romain Vergne, Joëlle Thollot.

Gabor Noise is a powerful procedural texture synthesis technique, but it 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.

This paper has been presented at Eurographics-short 2019 [11].

6.2. Illumination simulation and materials

6.2.1. Harmonic Analysis of the Light Transport Operator

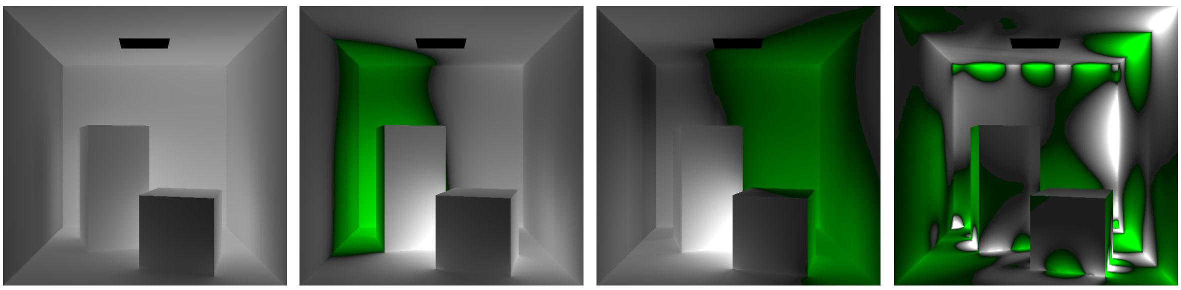
Participants: Ronak Molazem, Cyril Soler.

In this work we study the eigenvalues and eigenfunctions of the light transport operator. While computing the spectrum of the light transport operator is a simple task in Lambertian scenes by applying a traditional eigensolver to the linear system obtained from discretized geometry, it becomes a real challenge in general environments where discretizing the geometry is not possible anymore. “Diagonalizing” light transport however can be a very effective way to perform re-lighting and rapidely compute light transport solutions.

In this work we propose an analysis of the properties of the spectrum of the light transport operator, connecting the calculation of eigenvalues to resolvent theory. We show that the eigenfunctions are generally not orthogonal nor positive, but they can still be used to efficiently represent light distributions.

We analyse the performance of different methods to compute eigenvalues and images of their eigenfunctions using path tracing. We prove in particular that it is possible to compute the eigenfunctions of the light transport operator by integrating "circular" light paths of various lengths accross the scene.

This work is part of the PhD of Ronak Molazem and is funded by the ANR project "CaLiTrOp". At the time of writing this (Dec. 2019), we’re about to submit a paper to ACM Transactions on Graphics.



Λ_0 Λ_1 Λ_2 Λ_{144}

Figure 3. Path-traced images of four eigenfunctions of the light transport operator in the Cornell Box. A green scale is used to represent negative values.

6.2.2. Low Dimension Approximations of Light Transport

Participants: Ronak Molazem, Cyril Soler.

Light transport is known to be a low rank linear operator: the vector space formed by solutions of a light transport problem for different initial conditions is of low dimension. Approximating this space using appropriate bases is therefore of primordial help to efficiently compute solutions to light transport problems.

In this work, we’re interested into generating such approximations using *ad-hoc* methods that rely on deep learning. The goal is to be able to efficiently generate a sensible basis for light transport solutions on which we can efficiently project a noisy image. Other applications of this work include relighting pictures, in which an approximate geometry is used to project the illumination in the image, that can further be manipulated while staying in the space of expected light transport solutions.

This work is an ongoing collaboration with Unity Research Grenoble, and part of the PhD of Ronak Molazem, currently in her second year of PhD, and is funded by the ANR project "CaLiTrOp".

6.2.3. Precomputed Multiple Scattering for Rapid Light Simulation in Participating Media

Participants: Nicolas Holzschuch, Liangsheng Ge, Beibei Wang.

Rendering translucent materials is costly: light transport algorithms need to simulate a large number of scattering events inside the material before reaching convergence. The cost is especially high for materials with a large albedo or a small mean-free-path, where higher-order scattering effects dominate. In [7], we present a new method for fast computation of global illumination with participating media. Our method uses precomputed multiple scattering effects, stored in two compact tables. These precomputed multiple scattering tables are easy to integrate with any illumination simulation algorithm. We give examples for virtual ray lights (VRL), photon mapping with beams and paths (UPBP), Metropolis Light Transport with Manifold Exploration (MEMLT). The original algorithms are in charge of low-order scattering, combined with multiple scattering computed using our table. Our results show significant improvements in convergence speed and memory costs, with negligible impact on accuracy.

6.2.4. Fast Computation of Single Scattering in Participating Media with Refractive Boundaries using Frequency Analysis

Participants: Nicolas Holzschuch, Yulin Liang, Lu Wang, Beibei Wang.

Many materials combine a refractive boundary and a participating media on the interior. If the material has a low opacity, single scattering effects dominate in its appearance. Refraction at the boundary concentrates the incoming light, resulting in an important phenomenon called volume caustics. This phenomenon is hard to simulate. Previous methods used point-based light transport, but attributed point samples inefficiently, resulting in long computation time. In [3], we use frequency analysis of light transport to allocate point samples efficiently. Our method works in two steps: in the first step, we compute volume samples along with their covariance matrices, encoding the illumination frequency content in a compact way. In the rendering step, we use the covariance matrices to compute the kernel size for each volume sample: small kernel for high-frequency single scattering, large kernel for lower frequencies. Our algorithm computes volume caustics with fewer volume samples, with no loss of quality. Our method is both faster and uses less memory than the original method. It is roughly twice as fast and uses one fifth of the memory. The extra cost of computing covariance matrices for frequency information is negligible.

6.2.5. Reparameterizing discontinuous integrands for differentiable rendering

Participants: Nicolas Holzschuch, Wenzel Jakob, Guillaume Loubet.

Differentiable rendering has recently opened the door to a number of challenging inverse problems involving photorealistic images, such as computational material design and scattering-aware reconstruction of geometry and materials from photographs. Differentiable rendering algorithms strive to estimate partial derivatives of pixels in a rendered image with respect to scene parameters, which is difficult because visibility changes are inherently non-differentiable.

We propose [5] a new technique for differentiating path-traced images with respect to scene parameters that affect visibility, including the position of cameras, light sources, and vertices in triangle meshes. Our algorithm computes the gradients of illumination integrals by applying changes of variables that remove or strongly reduce the dependence of the position of discontinuities on differentiable scene parameters. The underlying parameterization is created on the fly for each integral and enables accurate gradient estimates using standard Monte Carlo sampling in conjunction with automatic differentiation. Importantly, our approach does not rely on sampling silhouette edges, which has been a bottleneck in previous work and tends to produce high-variance gradients when important edges are found with insufficient probability in scenes with complex visibility and high-resolution geometry. We show that our method only requires a few samples to produce gradients with low bias and variance for challenging cases such as glossy reflections and shadows. Finally, we use our differentiable path tracer to reconstruct the 3D geometry and materials of several real-world objects from a set of reference photographs.

6.3. Expressive rendering

6.3.1. Procedural Stylization

Participants: Maxime Isnel, Mohamed Amine Farhat, Romain Vergne, Joëlle Thollot.

Stylizing 3D scenes is a long term goal for the expressive rendering community. During the master internship of Maxime Isnel we have worked on a procedurel approach based on a procedural solid noise used in image space to generate brush strokes or 2.5D visual primitives, such as fur. The overview of the approach is shown Figure 4. This project is still in progress and will continue with a post-doc in 2020.

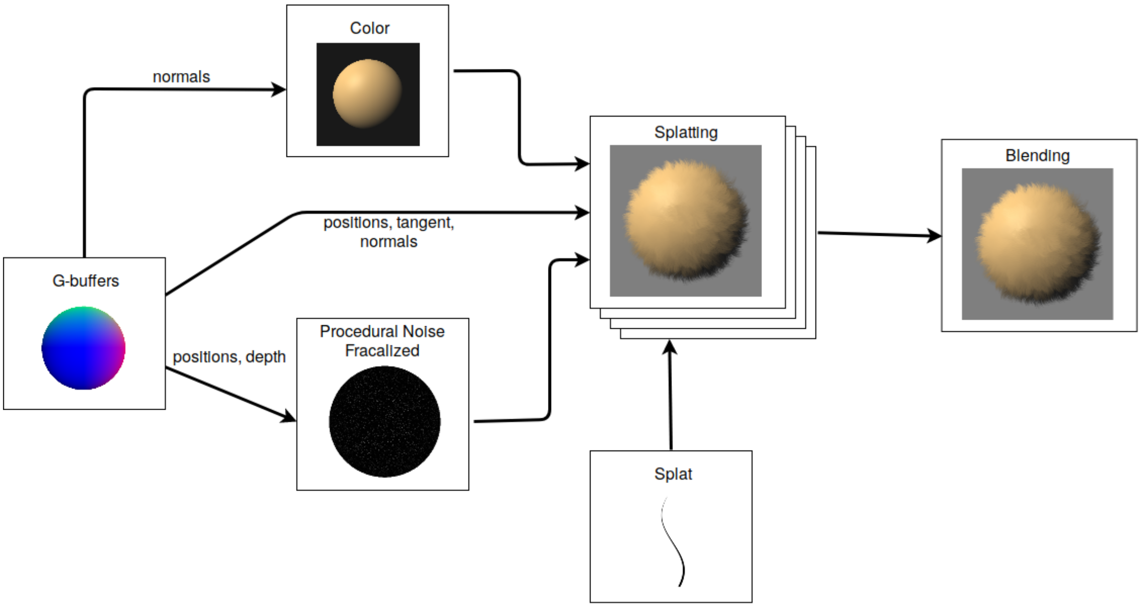


Figure 4. Based on a procedural solid noise and the use of geometry buffers, we propose an image-space approach to stylize a 3D object on the GPU.

7. Partnerships and Cooperations

7.1. 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.

7.2. National Initiatives

7.2.1. 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.

7.2.2. 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.

7.2.3. 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.

7.3. European Initiatives

Together with Stefanie Hahmann and Melina Skouras from project-team IMAGINE, Georges-Pierre Bonneau is part of the H2020 FET-Open Challenging Current Thinking project $ADAM^2$, grant ID 862025, accepted in June 2019 and starting officially January 1st 2020. The Imagine and Maverick teams at Inria are in charge of modelling of micro-structured geometries and design of meta-metarials. More information is available at www.adam2.eu.

7.4. International Initiatives

7.4.1. ASICIAO: Erasmus+ capacity building project

Joëlle Thollot is an active member of the ASICIAO Erasmus+ project. In this project four European higher education institutions support six schools from Senegal and Togo in their pursuit of autonomy by helping them to develop their own method of improving quality in order to obtain the CTI accreditation and the EUR-ACE label and, by doing so, to reach international standards.

7.5. International Research Visitors

7.5.1. Visits of International Scientists

7.5.1.1. Internships

Anmol Hanagodimath spent 6 months of internship in our team as part of his master thesis of Delft university. He was supervised by Romain Vergne and Joëlle Thollot in Grenoble and Elman Eisemann in Delft.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events: Selection

8.1.1.1. Member of the Conference Program Committees

- Romain Vergne: AFIG 2019
- Romain Vergne: SIBGRAPI 2019
- Joëlle Thollot: *Expressive* 2019.
- Nicolas Holzschuch: Eurographics 2020, Sibgrapi 2019, Eurographics Symposium on Rendering Steering Committee

8.1.2. Journal

All members of the Maverick team work as reviewers for the most prestigious journals, including ACM TOG, IEEE TVCG, CGF, etc.

8.1.3. Invited Talks

- Fabrice Neyret presented a 1 hour invited talk "Managing Ultra-high Complexity in Real-time Graphics: Some Hints and Ingredients" at HPG'2019 (ACM-SIGGRAPH-Eurographics Symposium on High-Performance Graphics) [9].
- Fabrice Neyret presented a 1 hour opening invited talk "Graphics vs Physics: the light and dark sides of the Force" [8] in pair with Emmanuel Delangre (Ladhyx - Ecole Polytechnique) at GraPhyz'2019, a new international workshop bridging Physics and Computer Graphics communities.
- Romain Vergne presented a 1 hour invited talk "ensuring congruency between shape and light" at IRIT: Institut de Recherche en Informatique de Toulouse.

8.1.4. 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 Rendering Working Group of the French Association ofr Computer Graphics (AFIG).
- Romain Vergne is co-responsible of the PhD students of the Laboratoire Jean Kuntzmann.
- Nicolas Holzschuch is an elected member of CNESER (National Council for Higher Education and Research), 2019 — 2023.
- 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.
- Nicolas Holzschuch is co-head of the Inria International Laboratory "Inria - EPFL".
- Nicolas Holzschuch is an elected member of Conseil Académique of the COMUE Université Grenoble-Alpes

8.2. Teaching - Supervision - Juries

8.2.1. 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. Joëlle Thollot is in charge of the MMIS ENSIMAG cursus (master level) with Stefanie Hahmann. Nicolas Holzschuch teaches advanced courses in computer graphics at the master level.

Licence: Joëlle Thollot, Théorie des langages, 45h, L3, ENSIMAG, France

Licence: Joëlle Thollot, Séminaire d'innovation, 10h, L3, ENSE3, France

Licence: Joëlle Thollot, MAP, 10h, L3, ENSIMAG, France

Master : Joelle Thollot, TD de créativité, 7h, M1, ENSIMAG, France.

Master : Joelle Thollot, English courses using theater, 18h, M1, ENSIMAG, France.

Licence : Romain Vergne, Introduction to algorithms, 64h, L1, UGA, France.

Licence : Romain Vergne, WebGL, 29h, L3, IUT2 Grenoble, France.

Licence : Romain Vergne, Programmation, 68h. L1, UGA, France.

Master : Romain Vergne, Image synthesis, 27h, M1, UGA, 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.

8.2.2. Supervision

PhD: Alban Fichet, Efficient representation for measured reflectance, UGA, 13/12/2019, 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

PhD in progress: Nolan Mestre, Rendering of panorama maps, 1/10/2019, Joëlle Thollot, Romain Vergne.

8.2.3. Juries

Romain Vergne, member of the jury, PhD of Julien Fayer, University of Toulouse, 19/04/2019.

Joëlle Thollot, president of the jury, PhD of Geoffrey Guingo, Université Grenoble-Alpes, 3/12/2019

Joëlle Thollot, president of the jury, PhD of Maxime Garcia, Université Grenoble-Alpes, 19/12/2019

Georges-Pierre Bonneau, reviewer, PhD of Yohann Bearzi, Université Claude Bernard Lyon 1, 8/11/2019

Georges-Pierre Bonneau, reviewer, PhD of Maxime Soler, Sorbonne Universités, 20/06/19

8.3. Popularization

8.3.1. Articles and contents

- Fabrice Neyret maintains the blog shadertoy-Unofficial and various shaders examples on Shadertoy site to popularize GPU technologies as well as disseminates academic models within computer graphics, computer science, applied math and physics fields. About 26k pages viewed and 12k unique visitors (87% out of France) in 2019.
- in 2019, Fabrice Neyret launched the blog desmosGraph-Unofficial to popularize the use of interactive grapher DesmosGraph for research, communication and pedagogy. For this year, about 1k pages viewed and 800 unique visitors (95% out of France).

8.3.2. Interventions

In the scope of MathC2+ (partnership with French ministry of Education and "Fondation Sciences mathématiques de Paris") Inria Rhône-Alpes hosted in June, 2019 a selection of about 30 motivated high-school students (level "second") for a day of presentations, demos and discussions. Fabrice Neyret did a 1h presentation on "Math and Physics in special effects and video games" plus discussions about his curriculum and questions.

9. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] H. DENG, B. WANG, R. WANG, N. HOLZSCHUCH.*A Practical Path Guiding Method for Participating Media*, in "Computational Visual Media", February 2020, https://hal.inria.fr/hal-02486014

[2] L. GE, B. WANG, L. WANG, X. MENG, N. HOLZSCHUCH.*Interactive Simulation of Scattering Effects in Participating Media Using a Neural Network Model*, in "IEEE Transactions on Visualization and Computer Graphics", January 2020 [*DOI :* 10.1109/TVCG.2019.2963015], https://hal.archives-ouvertes.fr/hal-02486039

[3] Y. LIANG, B. WANG, L. WANG, N. HOLZSCHUCH.*Fast Computation of Single Scattering in Participating Media with Refractive Boundaries using Frequency Analysis*, in "IEEE Transactions on Visualization and Computer Graphics", April 2019, p. 1-9, forthcoming [*DOI :* 10.1109/TVCG.2019.2909875], https://hal.inria.fr/hal-02089988

[4] W. LIN, B. WANG, L. WANG, N. HOLZSCHUCH.*A Detail Preserving Neural Network Model for Monte Carlo Denoising*, in "Computational Visual Media", February 2020, https://hal.inria.fr/hal-02488602

[5] G. LOUBET, N. HOLZSCHUCH, W. JAKOB.*Reparameterizing Discontinuous Integrands for Differentiable Rendering*, in "ACM Transactions on Graphics", November 2019, vol. 38, n^{o} 6, p. 228:1 - 228:14 [*DOI :* 10.1145/3355089.3356510], https://hal.inria.fr/hal-02497191

[6] T. TRICARD, S. EFREMOV, C. ZANNI, F. NEYRET, J. MARTÍNEZ, S. LEFEBVRE.*Procedural Phasor Noise*, in "ACM Transactions on Graphics", July 2019, vol. 38, n^{o} 4, Article No. 57:1-13 [*DOI :* 10.1145/3306346.3322990], https://hal.archives-ouvertes.fr/hal-02118508

[7] B. WANG, L. GE, N. HOLZSCHUCH.*Precomputed Multiple Scattering for Rapid Light Simulation in Participating Media*, in "IEEE Transactions on Visualization and Computer Graphics", 2019, p. 1-14, forthcoming [*DOI :* 10.1109/TVCG.2018.2890466], https://hal.inria.fr/hal-01970399

Invited Conferences

[8] F. NEYRET, E. DELANGRE.*(keynote) Graphics vs Physics : the light and dark sides of the Force*, in "GraPhyz 2019 - Workshop physique/mécanique/informatique graphique", Grenoble, France, October 2019, https://hal.inria.fr/hal-02333788

[9] F. NEYRET.*(keynote) Managing Ultra-high Complexity in Real-time Graphics: Some Hints and Ingredients*, in "HPG 2019 - High-Performance Graphics", Strasbourg, France, July 2019, https://hal.inria.fr/hal-02333768

International Conferences with Proceedings

[10] M. GERARDIN, N. HOLZSCHUCH, P. MARTINETTO.*Rendering transparent materials with a complex refractive index: semi-conductor and conductor thin layers*, in "MAM 2019 - 7th annual Workshop on Material Appearance Modeling", Strasbourg, France, July 2019, p. 1-4, https://hal.inria.fr/hal-02191534

[11] V. TAVERNIER, F. NEYRET, R. VERGNE, J. THOLLOT.*Making Gabor Noise Fast and Normalized*, in "Eurographics 2019 - 40th Annual Conference of the European Association for Computer Graphics", Gênes, Italy, T. E. ASSOCIATION (editor), Eurographics 2019 - Short Papers, 2019, p. 37-40 [*DOI :* 10.2312/EGS.20191009], https://hal.inria.fr/hal-02104389

[12] A. VODA, O. MARTIN, P. R. NAVES NETO, J.-D. GASCUEL, S. SCHMERBER.*Visual-vestibular compensation in balance recovery: A transfer function model-based analysis*, in "FOSBE 2019 - 8th IFAC Conference on Foundations of Systems Biology in Engineering", Valencia, Spain, October 2019, p. 1-6, https://hal.archives-ouvertes.fr/hal-02371059

Other Publications

[13] L. CIAVALDINI RIVIERE, E. CHALMIN, M. DUVAL, P. MARTINETTO, P. BORDET, K. FROMENT, C. DEJOIE, N. BLANC, N. HOLZSCHUCH, D. ZIÉBELIN, P.-A. DAVOINE.*Patrimalp, Science of Heritage : perspectives on cultural artworks*, February 2019, World meeting « Heritage Sciences and Technologies, 13-16 Février 2019, Paris, Poster, https://hal.archives-ouvertes.fr/hal-02416129

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

Florence Forbes [Team leader, Inria, Senior Researcher, HDR]
Sophie Achard [CNRS, Senior Researcher, from Sep 2019, HDR]
Julyan Arbel [Inria, Researcher, HDR]

Stephane Girard [Inria, Senior Researcher, HDR]

Faculty Member

Jean-Baptiste Durand [Institut polytechnique de Grenoble, Associate Professor]

External Collaborators

Sophie Achard [CNRS, from May 2019 until Aug 2019, HDR]
Thibaud Rahier [Criteo, from Feb 2019 until Apr 2019]

Technical Staff

Fatima Fofana [Inria, Engineer, until Aug 2019]

PhD Students

Karina Ashurbekova [Univ Grenoble Alpes, PhD Student]
Meryem Bousebata [Univ Grenoble Alpes, PhD Student]
Fabien Boux [Univ Grenoble Alpes, PhD Student]
Daria Bystrova [Univ Grenoble Alpes, PhD Student, from Oct 2019]
Alexandre Constantin [Univ Grenoble Alpes, PhD Student]
Benoit Kugler [Univ Grenoble Alpes, PhD Student]
Veronica Munoz Ramirez [Univ Grenoble Alpes, PhD Student]
Brice Olivier [Univ Grenoble Alpes, PhD Student, until May 2019]
Giovanni Poggiato [Univ Grenoble Alpes, PhD Student, from Nov 2019]
Mariia Vladimirova [Inria, PhD Student]

Post-Doctoral Fellows

Alexis Arnaud [Inria, Post-Doctoral Fellow, until Jul 2019]
Marta Crispino [Inria, Post-Doctoral Fellow, until Mar 2019]
Pascal Dkengne Sielenou [Inria, Post-Doctoral Fellow, until Oct 2019]
Hongliang Lu [Inria, Post-Doctoral Fellow, until Jan 2019]
Antoine Usseglio Carleve [Inria, Post-Doctoral Fellow]
Fei Zheng [Inria, Post-Doctoral Fellow]

Visiting Scientists

Aboubacrène Ag Ahmad [Univ. Gaston Berger, Senegal, from Nov 2019]
Hien Nguyen [Inria, from Nov 2019]
Darren Wraith [QUT, Brisbane, Australia, from Dec 2019]

Administrative Assistant

Marion Ponsot [Inria, Administrative Assistant]

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, Stephane Girard, Julyan Arbel, Daria Bystrova, Giovanni Poggiato, 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, Jean-Baptiste Durand, Florence Forbes, Karina Ashurbekova, Hongliang Lu, 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: Julyan Arbel, Daria Bystrova, Giovanni Poggiato, Stephane Girard, Florence Forbes, Antoine Usseglio Carleve, Pascal Dkengne Sielenou, 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 [82]. 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 [84] which combines non-parametric regression techniques with parametric dimension reduction aspects. This is also the case in *extreme value analysis* [81], 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 $X_{n,n}$.

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 [10] 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 [83]. 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 [80]. 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 [84].

4. Application Domains

4.1. Image Analysis

Participants: Alexis Arnaud, Veronica Munoz Ramirez, Florence Forbes, Stephane Girard, 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). 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.

4.2. Biology, Environment and Medicine

Participants: Alexis Arnaud, Florence Forbes, Stephane Girard, Jean-Baptiste Durand, Julyan Arbel, 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

New appointments:

- Florence Forbes has been appointed as a member of the advisory committee of the Helmholtz AI Cooperation Unit https://helmholtz.ai/.

Data Challenges

- Pixyl winner of the Société Francaise de Radiologie Data Challenge 2019

  Pixyl, a Grenoble-based start-up originating in the team and Inserm, accompanied by a team of neuroradiologists and academics, distinguished itself in the AI challenge held during the 2019 edition of the Journées Francophone de Radiologie, which took place from 11 to 14 October in Paris. The Challenge was about prediction of multiple sclerosis patient disability from a single MRI image

5.1.1. Awards

- Meryem Bousebata received the second best presentation award at the “10th conference of the international society for Integrated Disaster Risk Management (IDRiM)" organized by CNRS-University of Nice and AFPCN and held from 16 to 18 October 2019 in Nice.
- Mariia Vladimirova received the best poster award for her work [45] at the “12th Conference on Bayesian Nonparametrics”, Oxford University, UK, June 24-28, 2019.

BEST PAPERS AWARDS :

[52]

M. BOUSEBATA, G. ENJOLRAS, S. GIRARD.*Bayesian estimation of natural extreme risk measures. Application to agricultural insurance*, in "IDRiM 2019 - 10th conference of the international society for Integrated Disaster Risk Management", Nice, France, October 2019, https://hal.archives-ouvertes.fr/hal-02276292

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: 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. Mini-batch learning of exponential family finite mixture models

Participant: Florence Forbes.

Joint work with: Hien Nguyen, La Trobe University Melbourne Australia and Geoffrey J. McLachlan, University of Queensland, Brisbane, Australia.

Mini-batch algorithms have become increasingly popular due to the requirement for solving optimization problems, based on large-scale data sets. Using an existing online expectation-maximization (EM) algorithm framework, we demonstrate [28] how mini-batch (MB) algorithms may be constructed, and propose a scheme for the stochastic stabilization of the constructed mini-batch algorithms. Theoretical results regarding the convergence of the mini-batch EM algorithms are presented. We then demonstrate how the mini-batch framework may be applied to conduct maximum likelihood (ML) estimation of mixtures of exponential family distributions, with emphasis on ML estimation for mixtures of normal distributions. Via a simulation study, we demonstrate that the mini-batch algorithm for mixtures of normal distributions can outperform the standard EM algorithm. Further evidence of the performance of the mini-batch framework is provided via an application to the famous MNIST data set.

7.1.2. Component elimination strategies to fit mixtures of multiple scale distributions

Participants: Florence Forbes, Alexis Arnaud.

We address the issue of selecting automatically the number of components in mixture models with non-Gaussian components. As amore efficient alternative to the traditional comparison of several model scores in a range, we consider procedures based on a single run of the inference scheme. Starting from an over-fitting mixture in a Bayesian set-ting, we investigate two strategies to eliminate superfluous components.We implement these strategies for mixtures of multiple scale distributions which exhibit a variety of shapes not necessarily elliptical while remaining analytical and tractable in multiple dimensions. A Bayesian formulation and a tractable inference procedure based on variational approximation are proposed. Preliminary results on simulated and real data show promising performance in terms of model selection and computational time. This work has been presented at RSSDS 2019 - Research School on Statistics and Data Science in Melbourne, Australia [33].

7.1.3. Approximate Bayesian Inversion for high dimensional problems

Participants: Florence Forbes, Benoit Kugler.

Joint work with: Sylvain Douté from Institut de Planétologie et d'Astrophysique de Grenoble (IPAG).

The overall objective is to develop a statistical learning technique capable of solving complex inverse problems in setting with specific constraints. More specifically, the challenges are 1) the large number of observations to be inverted, 2) their large dimension, 3) the need to provide predictions for correlated parameters and 4) the need to provide a quality index (eg. uncertainty).

In the context of Bayesian inversion, one can use a regression approach, such as in the so-called Gaussian Locally Linear Mapping (GLLiM) [7], to obtain an approximation of the posterior distribution. In some cases, exploiting this approximate distribution remains challenging, for example because of its multi-modality. In this work, we investigate the possible use of Importance Sampling to build on the standard GLLiM approach by improving the approximation induced by the method and to better handle the potential existence of multiple solutions. We may also consider our approach as a way to provide an informed proposal distribution as requested by Importance Sampling techniques. We experiment our approach on simulated and real data in the context of a photometric model inversion in planetology. Preliminary results have been presented at StatLearn 2019 [76]

7.1.4. 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. . A dictionary-based learning (DBL) method was investigated to bypass inherent MRF limitations in high dimension: reconstruction time and memory requirement. The DBL method is a 3-step procedure: (1) a quasi-random sampling strategy to produce the dictionary, (2) a statistical inverse regression model to learn from the dictionary a probabilistic mapping between MR fingerprints and parameters, and (3) this mapping to provide both parameter estimates and their confidence levels. On synthetic data, experiments show that the quasi-random sampling outperforms the grid when designing the dictionary for inverse regression. Dictionaries up to 100 times smaller than usually employed in MRF yield more accurate parameter estimates with a 500 time gain. Estimates are supplied with a confidence index, well correlated with the estimation bias. On microvascular MRI data, results showed that dictionary-based methods (MRF and DBL) yield more accurate estimates than the conventional, closed-form equation, method. On MRI signals from tumor bearing rats, the DBL method shows very little sensitivity to the dictionary size in contrast to the

MRF method. The proposed method efficiently reduces the number of required simulations to produce the dictionary, speeds up parameter estimation, and improve estimates accuracy. The DBL method also introduces a confidence index for each parameter estimate. Preliminary results have been presented at the third *Congrés National d'Imagerie du Vivant* (CNIV 2019) [53] and at the fourth *Congrés de la Société Française de Résonance Magnétique en Biologie et Médecine* (SFRMBM 2019) [54].

7.1.5. Characterization of daily glycemic variability in subjects with type 1 diabetes using a mixture of metrics

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 is an important component of glycemic control for patients with type 1 diabetes. 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. In a first study [24], our goal 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.

In a second study, we address the issue of choosing an appropriate measure of GV. Many metrics have been proposed to account for this variability but none is unanimous among physicians. The inadequacy of existing measurements lies in the fact that they view the variability from different aspects, so that no consensus has been reached among physicians as to which metrics to use in practice. Moreover, although glycemic variability, from one day to another, can show very different patterns, few metrics have been dedicated to daily evaluations. In this work [50], [30], a reference (stable-glycemia) statistical model is built based on a combination of daily computed canonical glycemic control metrics including variability. The metrics are computed for subjects from the TRIMECO islet transplantation trial , selected when their β-score (composite score for grading success) is greater than 6 after a transplantation. Then, for any new daily glycemia recording, its likelihood with respect to this reference model provides a multi-metric score of daily glycemic variability severity. In addition, determining the likelihood value that best separates the daily glycemia with a zero β-score from that greater than 6, we propose an objective decision rule to classify daily glycemia into "stable" or "unstable". The proposed characterization framework integrates multiple standard metrics and provides a comprehensive daily glycemic variability index, based on which, long term variability evaluations and investigations on the implicit link between variability and β-score can be carried out. Evaluation, in a daily glycemic variability classification task, shows that the proposed method is highly concordant to the experience of diabetologists. A multivariate statistical model is therefore proposed to characterize the daily glycemic variability of subjects with type 1 diabetes. The model has the advantage to provide a single variability score that gathers the information power of a number of canonical scores, too partial to be used individually. A reliable decision rule to classify daily variability measurements into stable or unstable is also provided.

7.1.6. Dirichlet process mixtures under affine transformations of the data

Participant: Julyan Arbel.

Joint work with: Riccardo Corradin and Bernardo Nipoti from Milano Bicocca, Italy.

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 [63], 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.

7.1.7. Approximate Bayesian computation via the energy statistic

Participants: Julyan Arbel, Florence Forbes, Hongliang Lu.

Joint work with: Hien Nguyen, La Trobe University Melbourne Australia.

Approximate Bayesian computation (ABC) has become an essential part of the Bayesian toolbox for addressing problems in which the likelihood is prohibitively expensive or entirely unknown, making it intractable. ABC defines a quasi-posterior by comparing observed data with simulated data, traditionally based on some summary statistics, the elicitation of which is regarded as a key difficulty. In recent years, a number of data discrepancy measures bypassing the construction of summary statistics have been proposed, including the Kullback-Leibler divergence, the Wasserstein distance and maximum mean discrepancies. In this work [79], we propose a novel importance-sampling (IS) ABC algorithm relying on the so-called two-sample energy statistic. We establish a new asymptotic result for the case where both the observed sample size and the simulated data sample size increase to infinity, which highlights to what extent the data discrepancy measure impacts the asymptotic pseudo-posterior. The result holds in the broad setting of IS-ABC methodologies, thus generalizing previous results that have been established only for rejection ABC algorithms. Furthermore, we propose a consistent V-statistic estimator of the energy statistic, under which we show that the large sample result holds. Our proposed energy statistic based ABC algorithm is demonstrated on a variety of models, including a Gaussian mixture, a moving-average model of order two, a bivariate beta and a multivariate g-and-k distribution. We find that our proposed method compares well with alternative discrepancy measures.

7.1.8. Industrial applications of mixture modeling

Participant: Julyan Arbel.

Joint work with: Kerrie Mengersen and Earl Duncan from QUT, School of Mathematical Sciences, Brisbane, Australia, and Clair Alston-Knox, Griffith University Brisbane, Australia, and Nicole White, Institute for Health and Biomedical Innovation, Brisbane, Australia.

In [61], 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. The first of these focuses on the iconic and pervasive need for process monitoring, and reviews a range of mixture approaches that have been proposed to tackle complex multimodal and dynamic or online processes. The second study reports on mixture approaches to resource allocation, applied here in a spatial health context but which are applicable more generally. The next study provides a more detailed description of a multivariate Gaussian mixture approach to a biosecurity risk assessment problem, using big data in the form of satellite imagery. This is followed by a final study that again provides a detailed description of a mixture model, this time using a nonparametric formulation, for assessing an industrial impact, notably the influence of a toxic spill on soil biodiversity.

7.2. Semi and non-parametric methods

7.2.1. Deep learning models to study the early stages of Parkinson's Disease

Participants: Florence Forbes, Veronica Munoz Ramirez, Virgilio Kmetzsch Rosa E Silva.

Joint work with: Michel Dojat from Grenoble Institute of Neuroscience.

Current physio-pathological data suggest that Parkinson's Disease (PD) symptoms are related to important alterations in subcortical brain structures. However, structural changes in these small structures remain difficult to detect for neuro-radiologists, in particular, at the early stages of the disease (*de novo* PD patients) [58], [43], [59]. The absence of a reliable ground truth at the voxel level prevents the application of traditional supervised deep learning techniques. In this work, we consider instead an anomaly detection approach and show that auto-encoders (AE) could provide an efficient anomaly scoring to discriminate *de novo* PD patients using quantitative Magnetic Resonance Imaging (MRI) data.

7.2.2. Estimation of extreme risk measures

Participants: Stephane Girard, Antoine Usseglio Carleve.

Joint work with: A. Daouia (Univ. Toulouse), L. Gardes (Univ. Strasbourg) 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.

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 [23], 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 [6]. 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. have recently received a lot of attention, especially in actuarial and financial risk man-agement. Their estimation, however, typically requires to consider non-explicit asymmetric leasts-quares estimates rather than the traditional order statistics used for quantile estimation. This makes the study of the tail expectile process a lot harder than that of the standard tail quantile process. Under the challenging model of heavy-tailed distributions, we derive joint weighted Gaussian approximations of the tail empirical expectile and quantile processes. We then use this powerful result to introduce and study new estimators of extreme expectiles and the standard quantile-based expected shortfall, as well as a novel expectile-based form of expected shortfall [22].

Both quantiles and expectiles were embedded in the more general class of L_p-quantiles [21] as the minimizers of a generic asymmetric convex loss function. It has been proved very recently that the only L_p-quantiles that are coherent risk measures are the expectiles. In [75], we work in a context of heavy tails, which is especially relevant to actuarial science, finance, econometrics and natural sciences, and we construct an estimator of the tail index of the underlying distribution based on extreme L_p-quantiles. We establish the asymptotic normality of such an estimator and in doing so, we extend very recent results on extreme expectile and L_p-quantile estimation. We provide a discussion of the choice of p in practice, as well as a methodology for reducing the bias of our estimator. Its finite-sample performance is evaluated on simulated data and on a set of real hydrological data. This work is submitted for publication.

7.2.3. Conditional extremal events

Participants: Stephane Girard, Antoine Usseglio Carleve.

Joint work with: G. Stupfler (Univ. Nottingham, UK), A. Ahmad, E. Deme and A. Diop (Université Gaston Berger, Sénégal).

The goal of the PhD thesis of Aboubacrene Ag Ahmad is to contribute to the development of theoretical and algorithmic models to tackle conditional extreme value analysis, *ie* the situation where some covariate information X is recorded simultaneously with a quantity of interest Y. In such a case, extreme quantiles and expectiles are functions of the covariate. In [13], we consider a location-scale model for conditional heavy-tailed distributions when the covariate is deterministic. First, nonparametric estimators of the location and scale functions are introduced. Second, an estimator of the conditional extreme-value index is derived. The asymptotic properties of the estimators are established under mild assumptions and their finite sample properties are illustrated both on simulated and real data.

As explained in Paragraph 7.2.2, expectiles have recently started to be considered as serious candidates to become standard tools in actuarial and financial risk management. However, expectiles and their sample versions do not benefit from a simple explicit form, making their analysis significantly harder than that of quantiles and order statistics. This difficulty is compounded when one wishes to integrate auxiliary information about the phenomenon of interest through a finite-dimensional covariate, in which case the problem becomes the estimation of conditional expectiles. In [74], we exploit the fact that the expectiles of a distribution F are in fact the quantiles of another distribution E explicitly linked to F, in order to construct nonparametric kernel estimators of extreme conditional expectiles. We analyze the asymptotic properties of our estimators in the context of conditional heavy-tailed distributions. Applications to simulated data and real insurance data are provided. The results are submitted for publication.

7.2.4. Estimation of the variability in the distribution tail

Participant: Stephane Girard.

Joint work with: L. Gardes (Univ. Strasbourg).

We propose a new measure of variability in the tail of a distribution by applying a Box-Cox transformation of parameter $p \geq 0$ to the tail-Gini functional. It is shown that the so-called Box-Cox Tail Gini Variability measure is a valid variability measure whose condition of existence may be as weak as necessary thanks to the tuning parameter p. The tail behaviour of the measure is investigated under a general extreme-value condition on the distribution tail. We then show how to estimate the Box-Cox Tail Gini Variability measure within the range of the data. These methods provide us with basic estimators that are then extrapolated using the extreme-value assumption to estimate the variability in the very far tails. The finite sample behavior of the estimators is illustrated both on simulated and real data. This work is submitted for publication [72].

7.2.5. Extrapolation limits associated with extreme-value methods

Participant: Stephane 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 [15], 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 [14], 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.

7.2.6. Bayesian inference for copulas

Participants: Julyan Arbel, Marta Crispino, Stephane Girard.

We study in [16] 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.

7.2.7. Approximations of Bayesian nonparametric models

Participant: Julyan Arbel.

Joint work with: Stefano Favaro and Pierpaolo De Blasi from Collegio Carlo Alberto, Turin, Italy, Igor Prunster from Bocconi University, Milan, Italy, Caroline Lawless from Université Paris-Dauphine, France, Olivier Marchal from Université Jean Monnet.

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 [25] an elementary proof of Pitman–Yor's Chinese Restaurant process from its stick-breaking representation.

In [17], we consider approximations to the popular Pitman-Yor process obtained by truncating the stick-breaking representation. The truncation is determined by a random stopping rule that achieves an almost sure control on the approximation error in total variation distance. We derive the asymptotic distribution of the random truncation point as the approximation error goes to zero in terms of a polynomially tilted positive stable random variable. The practical usefulness and effectiveness of this theoretical result is demonstrated by devising a sampling algorithm to approximate functionals of the-version of the Pitman–Yor process.

In [18], we approximate predictive probabilities of Gibbs-type random probability measures, or Gibbs-type priors, which are arguably the most "natural" generalization of the celebrated Dirichlet prior. Among them the Pitman–Yor process certainly stands out for the mathematical tractability and interpretability of its predictive probabilities, which made it the natural candidate in several applications. Given a sample of size n, in this paper we show that the predictive probabilities of any Gibbs-type prior admit a large n approximation, with an error term vanishing as $o(1/n)$, which maintains the same desirable features as the predictive probabilities of the Pitman–Yor process.

In [18], we prove a monotonicity property of the Hurwitz zeta function which, in turn, translates into a chain of inequalities for polygamma functions of different orders. We provide a probabilistic interpretation of our result by exploiting a connection between Hurwitz zeta function and the cumulants of the exponential-beta distribution.

7.2.8. Concentration inequalities

Participant: Julyan Arbel.

Joint work with: Olivier Marchal from Université Jean Monnet and Hien Nguyen from La Trobe University Melbourne Australia.

In [19], we investigate the sub-Gaussian property for almost surely bounded random variables. If sub-Gaussianity per se is de facto ensured by the bounded support of said random variables, then exciting research avenues remain open. Among these questions is how to characterize the optimal sub-Gaussian proxy variance? Another question is how to characterize strict sub-Gaussianity, defined by a proxy variance equal to the (standard) variance? We address the questions in proposing conditions based on the study of functions variations. A particular focus is given to the relationship between strict sub-Gaussianity and symmetry of the distribution. In particular, we demonstrate that symmetry is neither sufficient nor necessary for strict sub-Gaussianity. In contrast, simple necessary conditions on the one hand, and simple sufficient conditions on the other hand, for strict sub-Gaussianity are provided. These results are illustrated via various applications to a number of bounded random variables, including Bernoulli, beta, binomial, uniform, Kumaraswamy, and triangular distributions.

7.2.9. Extraction and data analysis toward "industry of the future"

Participants: Florence Forbes, Hongliang Lu, Fatima Fofana.

Joint work with: J. F. Cuccaro and J. C Trochet from Vi-Technology company.

The overall idea of this project with Vi-Technology is to work towards manufacturing processes where machines communicate automatically so as to optimize the process performance as a whole. Starting from the assumption that transmitted information is essentially of statistical nature, the role of MISTIS in this context was to identify what statistical methods might be useful for the printed circuits boards assembly industry. A first step was to extract and analyze data from two inspection machines in an industrial process making electronic cards. After a first extraction 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 had access to two databases at both ends (SPI and Component Inspection) of the assembly process. The goal was 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.

7.2.10. Tracking and analysis of large population of dynamic single molecules

Participant: Florence Forbes.

Joint work with: Virginie Stoppin-Mellet from Grenoble Institute of Neuroscience, Vincent Brault from Laboratoire Jean Kuntzmann, Emilie Lebarbier from Nanterre University and Guy Bendao from AgroParisTech.

In the last decade, the number of studies using single molecule approaches has increased significantly. Thanks to technological progress and in particular with the development of TIRFM (Total Internal Reflection Fluorescence Microscopy), biologists can now observe single molecules at work. However, real time single molecule approaches remain mastered by a limited number of labs, and challenging obstacles have to be overcome before it becomes more broadly accessible. One important issue is the efficient detection and tracking of individual molecules in noisy images (low signal-to-noise ratio, SNR). Considering for example a TIRFM movie where single molecules stochastically appear and disappear at random positions, the low SNR implies that each individual molecule has to be detected at sub-pixel resolution over its local background and that this operation has to be repeated on each frame of the movie, thus requiring considerable amount of calculations. Procedures to detect single molecules are available, but they are mostly applicable to immobile molecules, are not statistically robust, and they often require an image processing that alters the quantitative signal information. In particular the intensity of a signal might be modified so that it becomes difficult to know the number of molecules associated with a specific signal. Crucial information such as the stoichiometry of the molecular complexes are then lost. Another challenging issue concerns data processing. Molecule tracking generate traces of time-dependent intensity fluctuations for each molecule. But single traces contain limited amount of information, and thus a very large number of traces must be analysed to extract general rules. In this context, the first aim of the present project was to provide a general procedure to track in real time transient interactions of a large number of biological molecules observed with TIRF microscopy and to generate traces of time-dependent intensity fluctuations. The second aim was to define a robust statistical approach to detect discrete events in a noisy time-dependent signal and extract parameters that describe the kinetics of these events. For this task we gathered expertise from biology (Grenoble Institute of Neuroscience) and statistics (Inria Mistis, LJK and AgroParisTech) in the context of a multidisciplinary project funded by the Grenoble data institute for 2 years.

7.3. Graphical and Markov models

7.3.1. Structure learning via Hadamard product of correlation and partial correlation matrices

Participants: Sophie Achard, Karina Ashurbekova, Florence Forbes.

Structure learning is an active topic nowadays in different application areas, i.e. genetics, neuroscience. Classical conditional independences or marginal independences may not be sufficient to express complex relationships. This work [39] is introducing a new structure learning procedure where an edge in the graph corresponds to a non zero value of both correlation and partial correlation. Based on this new paradigm, we define an estimator and derive its theoretical properties. The asymptotic convergence of the proposed graph estimator and its rate are derived. Illustrations on a synthetic example and application to brain connectivity are displayed.

7.3.2. Optimal shrinkage for robust covariance matrix estimators in a small sample size setting

Participants: Sophie Achard, Karina Ashurbekova, Florence Forbes, Antoine Usseglio Carleve.

When estimating covariance matrices, traditional sample covariance-based estimators are straightforward but suffer from two main issues: 1) a lack of robustness, which occurs as soon as the samples do not come from a Gaussian distribution or are contaminated with outliers and 2) a lack of data when the number of parameters to estimate is too large compared to the number of available observations, which occurs as soon as the covariance matrix dimension is greater than the sample size. The first issue can be handled by assuming samples are drawn from a heavy-tailed distribution, at the cost of more complex derivations, while the second issue can be addressed by shrinkage with the difficulty of choosing the appropriate level of regularization. In this work [66] we offer both a tractable and optimal framework based on shrinked likelihood-based M-estimators. First, a closed-form expression is provided for a regularized covariance matrix estimator with an optimal shrinkage coefficient for any sample distribution in the elliptical family. Then, a complete inference procedure is proposed which can also handle both unknown mean and tail parameter, in contrast to most existing methods that focus on the covariance matrix parameter requiring pre-set values for the others. An illustration on synthetic and real data is provided in the case of the t-distribution with unknown mean and degrees-of-freedom parameters.

7.3.3. Robust penalized inference for Gaussian Scale Mixtures

Participants: Sophie Achard, Karina Ashurbekova, Florence Forbes.

The literature on sparse precision matrix estimation is rapidly growing. Many strong methods are valid only for Gaussian variables. One of the most commonly used approaches in this case is glasso which aims to minimize the negative L1-penalized log-likelihood function. In practice, data may deviate from normality in various ways, outliers and heavy tails frequently occur that can severely degrade the Gaussian models performance. A natural solution is to turn to heavier tailed distributions that remain tractable. For this purpose, we propose [51] a penalized version of the EM algorithm for Gaussian Scale Mixtures.

7.3.4. Non parametric Bayesian priors for graph structured data

Participants: Florence Forbes, Julyan Arbel, Hongliang Lu.

We consider the issue of determining the structure of clustered data, both in terms of finding the appropriate number of clusters and of modelling the right dependence structure between the observations. Bayesian nonparametric (BNP) models, which do not impose an upper limit on the number of clusters, are appropriate to avoid the required guess on the number of clusters but have been mainly developed for independent data. In contrast, Markov random fields (MRF) have been extensively used to model dependencies in a tractable manner but usually reduce to finite cluster numbers when clustering tasks are addressed. Our main contribution is to propose a general scheme to design tractable BNP-MRF priors that combine both features: no commitment to an arbitrary number of clusters and a dependence modelling. A key ingredient in this construction is the availability of a stick-breaking representation which has the threefold advantage to allowing us to extend standard discrete MRFs to infinite state space, to design a tractable estimation algorithm using variational approximation and to derive theoretical properties on the predictive distribution and the number of clusters of the proposed model. This approach is illustrated on a challenging natural image segmentation task for which it shows good performance with respect to the literature. This work [77] will be presented as a poster at BayesComp2020 in Gainesville, Florida, USA, [78].

7.3.5. Bayesian nonparametric models for hidden Markov random fields on count variables and application to disease mapping

Participants: Julyan Arbel, Fatoumata Dama, Jean-Baptiste Durand, Florence Forbes.

Hidden Markov random fields (HMRFs) have been widely used in image segmentation and more generally, for clustering of data indexed by graphs. Dependent hidden variables (states) represent the cluster identities and determine their interpretations. Dependencies between state variables are induced by the notion of neighborhood in the graph. A difficult and crucial problem in HMRFs is the identification of the number of possible states K. Recently, selection methods based on Bayesian non parametric priors (Dirichlet processes) have been developed. They do not assume that K is bounded a priori, thus allowing its adaptive selection with respect to the quantity of available data and avoiding costly systematic estimation and comparison of models with different fixed values for K. Our previous work [77] has focused on Bayesian nonparametric priors for HMRFs and continuous, Gaussian observations. In this work, we consider extensions to discrete observed data typically issued from counts. We define and implement Bayesian nonparametric models for HMRFs with Poisson distributed observations. As an illustration, we propose a new disease mapping model for epidemiology. The inference is done by Variational Bayesian Expectation Maximization (VBEM). Results on synthetic data sets suggest that our model is able to recover the true number of risk levels (clusters) and to provide a good estimation of the true risk level partition. Application on real data then also shows satisfying results.

As a perspective, Bayesian nonparametric models for hidden Markov random fields could be extended to non-Poissonian models (particularly to account for zero-inflated and over-/under-dispersed cases of application) and to regression models.

7.3.6. Hidden Markov models for the analysis of eye movements

Participants: Jean-Baptiste Durand, Brice Olivier, Sophie Achard.

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 strategies or 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) [86], but we used hidden semi-Markov models for better characterization of phase length distributions (Olivier *et al.*, 2017) [85]. The estimated HMM highlighted contrasted reading strategies, with both individual and document-related variability. 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.

Then, using the restored state values, statistical characteristics of EEGs were compared according to strategies, brain wave frequencies and EEG channels (i.e., location on scalp). Differences in variance and correlations related to strategy changes were highlighted. Dependency graphs interpreted as maps of functional brain connectivity were estimated for each strategy and frequency and their changes were interpreted.

These results were published in Brice Olivier's PhD manuscript [12]. Although the approach was sufficient to highlight significant discrimination of strategies, it suffered from somewhat overlapping eye-movement characteristics over strategies. As a result, high uncertainty in the phase changes arose, which could induce underestimation of EEG and eye movement abilities to discriminate strategies.

This is why we developed integrated models coupling EEG and eye movements within one single HMM for better identification of strategies. Here, the coupling incorporated some delay between transitions in both EEG and eye-movement state sequences, since EEG patterns associated to cognitive processes occur lately with respect to eye-movement state switches. 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.

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 reading strategies.

7.3.7. Comparison of initialization strategies in the EM algorithm for hidden Semi-Markov processes

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)

In Subsection 7.3.6, hidden semi-Markov models (HSMMs) were used to infer reading strategies from eye-movement and EEG time series. Model parameters were estimated by the EM algorithm. Its principle is to build a sequence of parameters with increasing likelihood values, starting from a starting point. The impact of this starting point has not been investigated in the case of HSMMs; this is why we aimed at developing and assessing an initialization method based on the available sequence lengths [48]. This consists in randomly choosing a number of transitions and then, uniformly-distributed transition times given the number of transitions. These transition times break the sequences into segements and assign uniformly-distributed states to each segment with the constraint that two consecutive states should be different.

The method was compared to other initialization strategies and was shown to be efficient on several data sets with multiple categorical sequences.

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 [40], 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. Bayesian neural networks

Participants: Julyan Arbel, Mariia Vladimirova.

Joint work with: Pablo Mesejo from University of Granada, Spain, Jakob Verbeek from Inria Grenoble Rhône-Alpes, France.

We investigate in [45] 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

Contract with EDF (2019). Stéphane Girard is the advisor of the internship of Valentin Chevalier founded by 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 50 keuros.

Contract with VALEO (2018-2019). Stéphane 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. ANR

MISTIS is involved in the 4-year ANR project ExtremReg (2019-2023) hosted by Toulouse University. This research project aims to provide new adapted tools for nonparametric and semiparametric modeling from the perspective of extreme values. Our research program concentrates around three central themes. First, we contribute to the expanding literature on non-regular boundary regression where smoothness and shape constraints are imposed on the regression function and the regression errors are not assumed to be centred, but one-sided. Our second aim is to further investigate the study of the modern extreme value theory built on the use of asymmetric least squares instead of traditional quantiles and order statistics. Finally, we explore the less-discussed problem of estimating high-dimensional, conditional and joint extremes

The financial support for MISTIS is about 15.000 euros.

9.1.2. Grenoble Idex projects

MISTIS is involved in a transdisciplinary 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.

9.1.3. 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.

9.1.4. 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.

9.2. European Initiatives

9.2.1. FP7 & H2020 Projects

VHIA ERC project (2015-19).

MISTIS is involved in R. Horaud's ERC advanced Grant entitled Vision and Hearing In Action. VHIA studies the fundamentals of audio-visual perception for human-robot interaction.

9.3. International Initiatives

9.3.1. Inria International Labs

International Laboratory for Research in Computer Science and Applied Mathematics
Associate Team involved in the International Lab:

9.3.1.1. SIMERG2E

Title: Statistical Inference for the Management of Extreme Risks, Genetics and Global Epidemiology

International Partner (Institution - Laboratory - Researcher):

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).

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

9.3.2.1. LANDER

Title: Latent Analysis, Adversarial Networks, and DimEnsionality Reduction

International Partner (Institution - Laboratory - Researcher):

La Trobe university, Melbourne (Australia) - Department of Mathematics - Hien Nguyen

Start year: 2019

See also: https://team.inria.fr/mistis/projects/lander/

The collaboration is based on three main points, in statistics, machine learning and applications: 1) clustering and classification (mixture models), 2) regression and dimensionality reduction (mixture of regression models and non parametric techniques) and 3) high impact applications (neuroimaging and MRI). Our overall goal is to collectively combine our resources and data in order to develop tools that are more ubiquitous and universal than we could have previously produced, each on our own. A wide class of problems from medical imaging can be formulated as inverse problems. Solving an inverse problem means recovering an object from indirect noisy observations. Inverse problems are therefore often compounded by the presence of errors (noise) in the data but also by other complexity sources such as the high dimensionality of the observations and objects to recover, their complex dependence structure and the issue of possibly missing data. Another challenge is to design numerical implementations that are computationally efficient. Among probabilistic models, generative models have appealing properties to meet all the above constraints. They have been studied in various forms and rather independently both in the statistical and machine learning literature with different depths and insights, from the well established probabilistic graphical models to the more recent (deep) generative adversarial networks (GAN). The advantages of the latter being primarily computational and their disadvantages being the lack of theoretical statements, in contrast to the former. The overall goal of the collaboration is to build connections between statistical and machine learning tools used to construct and estimate generative models with the resolution of real life inverse problems as a target. This induces in particular the need to help the models scale to high dimensional data while maintaining our ability to assess their correctness, typically the uncertainty associated to the provided solutions.

9.3.3. Inria International Partners

9.3.3.1. 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 other active international collaborations in 2019 are with:

- E. Deme and A. Diop from Gaston Berger University in Senegal.
- N. Wang and C-C. Tu from University of Michigan, Ann Arbor, USA.
- 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.
- Alex Petersen, University of California Santa Barbara, USA.
- Dimitri van de Ville, EPFL, University of Geneva, Switzerland.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Bernardo Nipoti, assistant professor at Milano Bicocca University, Italy, visited for a month in 2019 (three visits in February, April and September).
- Natalie Karavarsamis, assistant professor at La Trobe University in Melbourne, Australia, visited for a week in November 2019.
- Hien Nguyen, researcher at La Trobe University in Melbourne, Australia, visited for a month in November 2019.
- Darren Wraith, assistant professor at QUT, Brisbane, Australia, visited for 2 weeks in December 2019 and January 2020.
- Aboubacrène Ag Ahmad, PhD student at Univ. Gaston Berger, Senegal visited from September 2019 until November 2019.

9.4.1.1. Internships

Sharan Yalburgi did an internship of three months with Julyan Arbel on *Bayesian deep learning for model selection and approximate inference*.

9.4.1.2. Research Stays Abroad

Mariia Vladimirova visited David Dunson at Duke University for three months (Nov 2019 - Jan 2020).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- Stéphane Girard was chairman at the 2nd workshop on Multivariate Data and Software (Limassol, Cyprus) and at the International Workshop on Stress Test and Risk Management (Paris).

10.1.1.2. Member of the Organizing Committees

- Florence Forbes is a member of the scientific committees of the Bayes Comp 2020 conference in Gainesville, Florida, USA (January 2020) and of the Research school on Networks and molecular biology at CIRM in Marseille (March 2020).
- Sophie Achard was a member of the scientific committee of the Wavelet & Sparsity XVIII 2019 in San Diego and the organizer of a special session within this conference.
- Stéphane Girard and Julyan Arbel were members of the organizing committee of the 10th Statlearn international workshop "Challenging problems in Statistical Learning", Grenoble, http://statlearn.sfds.asso.fr. Stéphane Girard also co-organized (with D. Fraix-Burnet, IPAG) the 4th international school Stat4Astro "Variability and Time Series Analysis", Autrans, http://stat4astro2019.sciencesconf.org.
- Julyan Arbel was a member of the scientific committee of *Statistical Methods for Post Genomic Data analysis (SMPGD)*, link. Julyan Arbel organized the session entitled ‘Bayesian Machine Learning’ at the 12th International Conference of Computational and Methodological Statistics (CMStat), University of London, UK (14-16 December 2019).

Seminars organization

- MISTIS participates in the weekly statistical seminar of Grenoble. Several lecturers have been invited in this context.
- Julyan Arbel is organizing monthly reading group Bayes in Grenoble on Bayesian statistics.

10.1.2. Scientific Events Selection

10.1.2.1. Reviewer

- In 2019, Florence Forbes has been a reviewer for CAP 2019 in Toulouse and for ICDHT 2019 in Tunis.
- In 2019, Julyan Arbel has been a reviewer for the *Bayesian Young Statisticians Meeting proceedings* (BAYSM).
- In 2019, Florence Forbes and Julyan Arbel have been reviewers for the *Research School on Statistics and Data Science* (RSSDS2019).

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, Associate Editor of the *Journal of Multivariate Analysis* since 2016 and Associate Editor of *REVSTAT - Statistical Journal* since 2019. 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 *Bayesian Analysis* (BA) and of Statistics & Probability Letters (SPL) since 2019.
- Julyan Arbel and Florence Forbes are Associate Editors for the *Australian & New Zealand Journal of Statistics* (ANZJS), since 2018.
- Sophie Achard is Associate Editor of *Neural Processing Letters* and *Network Neuroscience* since 2016.

10.1.3.2. Reviewer - Reviewing Activities

- In 2019, Florence Forbes has been a reviewer for *IEEE Transactions on Pattern Analysis and Machine Intelligence* (TPAMI), *Statistics and Computing* (STCO), and *Neural Processing Letters*.
- In 2019, Stéphane Girard has been a reviewer for *Journal of the American Statistical Association* (JASA), *Journal of Statistical Planning and Inference* (JSPI), *Communications in Statistics - Theory and Methods, Spatial Statistics.*
- In 2019, Jean-Baptiste Durand has been a reviewer for *Behavior Research Methods* (BRM) and *Statistics and Computing* (STCO).
- In 2019, Julyan Arbel has been a reviewer for: *Annals of Applied Statistics* (AOAS), *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* (AIHP), *Bernoulli*, *Biometrika*, *Entropy*, *Journal of the American Statistical Association* (JASA), *Journal of Computational and Graphical Statistics* (JCGS), *Journal of Nonparametric Statistics* (JNS), *Sankhyā*, *Stats*, *Statistica Sinica*, *Statistics and Probability Letters* (SPL), *Stochastic Processes and their Applications* (SPA), *IEEE Transactions on Pattern Analysis and Machine Intelligence* (TPAMI).

10.1.4. Invited Talks

Florence Forbes has been invited to give talks at the following seminars and conferences:

- Glasgow Statistics Department, March 2019
- Conference on Applied Inverse Problems, AIP 2019, July 2019 [34]
- 51ème Journées de la Statistique, Nancy, France, June 2019 [35]
- Workshop on Model-based clustering, Vienna, Austria, July 15-19, [32]
- Research School on Statistics and Data Science, Melbourne, Australia, July 24-26, [33]

Julyan Arbel has been invited to give talks at the following seminars and conferences:

- Applied Inverse Problems conference, Grenoble, France, July 8-12, 2019. Invited talk: Understanding Priors in Bayesian Neural Networks at the Unit Level.
- 11th Workshop on Bayesian Inference in Stochastic Processes (BISP), Madrid, Spain, June 12-14, 2019. Invited talk: Understanding Priors in Bayesian Neural Networks at the Unit Level.
- Workshop on Multivariate Data Analysis, Limassol, Cyprus, April 14-16, 2019. Invited talk: Some distributional properties of Bayesian neural networks.
- Seminar, Laboratoire d'Informatique de Grenoble, Grenoble, France, September 19, 2019. Invited talk: On some theory for Bayesian neural networks.
- Grenoble R User Group, Grenoble, France, April 11, 2019. Invited talk: R Markdown.

Sophie Achard has been invited to give talks at the following seminars and conferences:

- Workshop ATLAS, GDR MADICS, November 2019, Grenoble http://ama.liglab.fr/ATLAS/Wksp-22112019.html. Invited talks: Assessing reliability of resting-state fMRI graph analysis: challenges in measuring brain connectivity networks alterations for clinical applications.
- NeuroSTIC, GDR BioComp et ISIS, October 2019, Nice http://www.gdr-isis.fr/neurostic/?p=452. Invited talks: Brain connectivity for patients with consciousness disorders: statistical and clinical challenges

Stéphane Girard has been invited to give talks at the following seminars and conferences:

- 2nd workshop on Multivariate Data and Software (Limassol, Cyprus) [37],
- Workshop "Appréhender la grande dimension" (Paris) [36],
- Seminar, Nottingham University, UK "Estimation of extreme risk measures based on L_p- quantiles".

Antoine Usseglio-Carleve was invited to give a talk [38] at the 12th International Conference of Computational and Methodological Statistics, London, UK.

Marta Crispino was invited to give a talk [31] at the 12th International Conference of Computational and Methodological Statistics, London, UK.

10.1.5. Scientific Expertise

Florence Forbes is Scientific Advisor since March 2015 for the Pixyl company.

10.1.6. Research Administration

- Stéphane Girard is a member of the "Comité des Emplois Scientifiques" at Inria Grenoble Rhône-Alpes since 2015.
- Since 2015, Stéphane 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.
- 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 Ensimag Grenoble and at Ecole Centrale Lille in 2019.
- Florence Forbes is a member of the advisory committee of the Helmholtz AI Cooperation Unit https://helmholtz.ai/, since 2019.
- Sophie Achard is co-director of pôle MSTIC whithin Université Grenoble Alpes, since 2017.
- Julyan Arbel is a scientific committee member of the Data Science axis of Persyval Labex (Machine learning: fundamentals and applications, and Data linking, sharing and privacy), link, since 2019.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Master : Stéphane Girard, *Statistique Inférentielle Avancée*, 18 ETD, M1 level, Ensimag. Grenoble-INP, France.
- PhD course: Julyan Arbel, *Bayesian nonparametrics*, Jyväskylä Summer School, Finland, August 2019, 25 ETD.
- 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 and PhD course: Julyan Arbel, *Bayesian machine learning*, Master Mathématiques Vision et Apprentissage Master MVA, École normale supérieure Paris-Saclay, 36 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.
- Sophie Achard M1 course Théorie des graphes et réseaux sociaux, M1 level, MIASHS, Université Grenoble Alpes (UGA), 14 ETD.

10.2.2. Supervision

- PhD defended: Karina Ashurbekova "*Robust Structure Learning*", December 2019, Sophie Achard and Florence Forbes, Université Grenoble Alpes.
- PhD defended: Brice Olivier "*Joint analysis of eye-movements and EEGs using coupled hidden Markov models*", June 2019, Jean-Baptiste Durand and Anne Guérin-Dugué, Université Grenoble Alpes.
- PhD defended: Chun-Chen Tu,"*Gaussian mixture sub-clustering/reduction refinement of Non-linear high-to-low dimensional mapping*", "Date", Florence Forbes and Naisyin Wang, University of Michigan, Ann Arbor.
- HDR: Julyan Arbel, Université Grenoble Alpes, “*Bayesian Statistical Learning and Applications*” [11], October 2019.
- 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: Mariia Vladimirova, “*Prior specification for Bayesian deep learning models and regularization implications*”, started on October 2018, Julyan Arbel and Jakob Verbeek.
- PhD in progress: Aboubacrène Ag Ahmad "*A new location-scale model for heavy-tailed distributions*", started on September 2016, Stéphane 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, Stéphane 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, Stéphane Girard and Mathieu Fauvel (Université Grenoble Alpes).
- PhD in progress: Daria Bystrova, "*Joint Species Distribution Modeling: Dimension reduction using Bayesian nonparametric priors*", started on October 2019, Julyan Arbel and Wilfried Thuiller.
- PhD in progress: Giovanni Poggiatto , "*Scalable Approaches for Joint Species Distribution Modeling*", started on November 2019, Julyan Arbel and Wilfried Thuiller.

10.2.3. Juries

- Julyan Arbel has been reviewer for the PhD thesis of Romain Mismer, LPSM, Sorbonne Université, Paris.
- Stéphane Girard has been reviewer for the PhD thesis of Maxime Baelde, Université de Lille.
- Stéphane Girard has been the president of the HDR committee of Julie Carreau, Université de Montpellier, and an examinator for the HDR of Julyan Arbel.
- Stéphane Girard has been a member of the PhD committee of Abdul-Fattah Abu-Awwad, Université de Lyon.
- Florence Forbes has been reviewer for the PhD thesis of Lê-Huu D. Khuê, Université Paris Saclay, CentraleSupelec, Cedric Meurée, Université de Rennes, Bao Tuyen Huynh Université de Caen and for the HDR thesis of Christine Keribin, Université Paris Orsay.
- Florence Forbes has been a member of the PhD committee of Charlotte Maugard, Université Grenoble Alpes and Esteban Bautista, ENS Lyon.
- Sophie Achard has been reviewer for the HDR of Julien Modolo, Université Rennes 1.

10.3. Popularization

10.3.1. Interventions

- Sophie Achard has been invited to Festival des Nouvelles Explorations https://nouvellesexplorations.com/.
- Julyan Arbel gave a presentation for ISN conference, March 2019.

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. DAOUIA, S. GIRARD, G. STUPFLER.*Estimation of Tail Risk based on Extreme Expectiles*, in "Journal of the Royal Statistical Society series B", 2018, vol. 80, p. 263–292

[7] 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

[8] 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

[9] 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

[10] 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

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] J. ARBEL.*Bayesian Statistical Learning and Applications*, Université grenoble Alpes, CNRS, Institut des Géosciences et de l'Environnement, October 2019, Habilitation à diriger des recherches, https://tel.archives-ouvertes.fr/tel-02429156

[12] B. OLIVIER.*Joint analysis of eye movements and EEGs using coupled hidden Markov*, Université Grenoble Alpes, June 2019, https://tel.archives-ouvertes.fr/tel-02311373

Articles in International Peer-Reviewed Journal

[13] A. A. AHMAD, E. H. DEME, A. DIOP, S. GIRARD.*Estimation of the tail-index in a conditional location-scale family of heavy-tailed distributions*, in "Dependence Modeling", 2019, vol. 7, p. 394–417, https://hal.inria.fr/hal-02132976

[14] C. ALBERT, A. DUTFOY, L. GARDES, S. GIRARD.*An extreme quantile estimator for the log-generalized Weibull-tail model*, in "Econometrics and Statistics ", 2019, p. 1-39, forthcoming [*DOI :* 10.1016/J.ECOSTA.2019.01.004], https://hal.inria.fr/hal-01783929

[15] C. ALBERT, A. DUTFOY, S. GIRARD.*Asymptotic behavior of the extrapolation error associated with the estimation of extreme quantiles*, in "Extremes", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-01692544

[16] J. ARBEL, M. CRISPINO, S. GIRARD.*Dependence properties and Bayesian inference for asymmetric multivariate copulas*, in "Journal of Multivariate Analysis", November 2019, vol. 174, p. 104530:1-20 [*DOI :* 10.1016/J.JMVA.2019.06.008], https://hal.archives-ouvertes.fr/hal-01963975

[17] J. ARBEL, P. DE BLASI, I. PRÜNSTER.*Stochastic approximations to the Pitman-Yor process*, in "Bayesian Analysis", June 2019, vol. 14, n^{o} 3, p. 753-771 [*DOI :* 10.1214/18-BA1127], https://hal.archives-ouvertes.fr/hal-01950654

[18] J. ARBEL, S. FAVARO.*Approximating predictive probabilities of Gibbs-type priors*, in "Sankhya A", September 2019, p. 1-21, https://hal.archives-ouvertes.fr/hal-01693333

[19] J. ARBEL, O. MARCHAL, H. T. NGUYEN.*On strict sub-Gaussianity, optimal proxy variance and symmetry for bounded random variables*, in "ESAIM: Probability and Statistics", December 2019, https://hal.archives-ouvertes.fr/hal-01998252

[20] M. CRISPINO, E. ARJAS, V. VITELLI, N. BARRETT, A. FRIGESSI.*A Bayesian Mallows Approach to Non-Transitive Pair Comparison Data: How Human are Sounds?*, in "Annals of Applied Statistics", June 2019, vol. 13, n^{o} 1, p. 492-519 [*DOI :* 10.1214/18-AOAS1203], https://hal.archives-ouvertes.fr/hal-01972952

[21] A. DAOUIA, S. GIRARD, G. STUPFLER.*Extreme M-quantiles as risk measures: From L1 to Lp optimization*, in "Bernoulli", February 2019, vol. 25, n^{o} 1, p. 264-309 [*DOI :* 10.3150/17-BEJ987], https://hal.inria.fr/hal-01585215

[22] A. DAOUIA, S. GIRARD, G. STUPFLER.*Tail expectile process and risk assessment*, in "Bernoulli", 2019, p. 1-27, forthcoming, https://hal.archives-ouvertes.fr/hal-01744505

[23] L. GARDES, S. GIRARD, G. STUPFLER.*Beyond tail median and conditional tail expectation: extreme risk estimation using tail Lp -optimisation*, in "Scandinavian Journal of Statistics", 2019, p. 1-69, forthcoming [*DOI :* 10.1111/SJOS.12433], https://hal.inria.fr/hal-01726328

[24] M. JALBERT, F. ZHENG, A. WOJTUSCISZYN, F. FORBES, S. BONNET, K. SKAARE, P.-Y. BENHAMOU, S. LABLANCHE.*Glycemic variability indices can be used to diagnose islet transplantation success in type 1 diabetic patients*, in "Acta Diabetologica", October 2019, p. 1-11 [*DOI :* 10.1007/S00592-019-01425-3], https://hal.archives-ouvertes.fr/hal-02328170

[25] C. LAWLESS, J. ARBEL.*A simple proof of Pitman-Yor's Chinese restaurant process from its stick-breaking representation*, in "Dependence Modeling", 2019, vol. 7, n^{o} 1, p. 45-52 [*DOI :* 10.1515/DEMO-2019-0003], https://hal.archives-ouvertes.fr/hal-01950653

[26] Q. LIU, M. CRISPINO, I. SCHEEL, V. VITELLI, A. FRIGESSI.*Model-based learning from preference data*, in "Annual Reviews of Statistics and its Application", March 2019, vol. 6, n^{o} 1, p. 329-354 [*DOI :* 10.1146/ANNUREV-STATISTICS-031017-100213], https://hal.archives-ouvertes.fr/hal-01972948

[27] H. D. NGUYEN, F. CHAMROUKHI, F. FORBES.*Approximation results regarding the multiple-output Gaussian gated mixture of linear experts model*, in "Neurocomputing", November 2019, vol. 366, p. 208-214 [*DOI :* 10.1016/J.NEUCOM.2019.08.014], https://hal.archives-ouvertes.fr/hal-02265793

[28] H. D. NGUYEN, F. FORBES, G. J. MCLACHLAN.*Mini-batch learning of exponential family finite mixture models*, in "Statistics and Computing", 2019, p. 1-40, forthcoming, https://hal.archives-ouvertes.fr/hal-02415068

[29] C.-C. TU, F. FORBES, B. LEMASSON, N. WANG.*Prediction with high dimensional regression via hierarchically structured Gaussian mixtures and latent variables*, in "Journal of the Royal Statistical Society:

Series C Applied Statistics", 2019, p. 1-23 [*DOI :* 10.1111/RSSC.12370], https://hal.archives-ouvertes.fr/hal-02263144

[30] F. ZHENG, M. JALBERT, F. FORBES, S. BONNET, A. WOJTUSCISZYN, S. LABLANCHE, P.-Y. BENHAMOU.*Characterization of Daily Glycemic Variability in Subjects with Type 1 Diabetes Using a Mixture of Metrics*, in "Diabetes Technology and Therapeutics", 2019, p. 1-17, forthcoming [*DOI :* 10.1089/DIA.2019.0250], https://hal.archives-ouvertes.fr/hal-02415078

Invited Conferences

[31] M. CRISPINO, S. GIRARD, J. ARBEL.*Dependence properties and Bayesian inference for asymmetric multivariate copulas*, in "CMStatistics 2019 - 12th International Conference of the ERCIM WG on Computational and Methodological Statistics", London, United Kingdom, December 2019, https://hal.archives-ouvertes.fr/hal-02413948

[32] F. FORBES, A. ARNAUD, B. LEMASSON, E. L. BARBIER.*Bayesian mixtures of multiple scale distributions*, in "2019 - 26th Summer Working Group on Model-Based Clustering", Vienna, Austria, July 2019, https://hal.archives-ouvertes.fr/hal-02423638

[33] F. FORBES, A. ARNAUD, B. LEMASSON, E. L. BARBIER.*Component elimination strategies to fit mixtures of multiple scale distributions*, in "RSSDS 2019 - Research School on Statistics and Data Science", Melbourne, Australia, Proceedings of the Research School on Statistics and Data Science 2019, July 2019, p. 1-15, https://hal.archives-ouvertes.fr/hal-02415090

[34] F. FORBES, A. DELEFORGE, R. HORAUD, E. PERTHAME.*Robust non-linear regression approach for generalized inverse problems in a high dimensional setting*, in "AIP 2019 - Applied Inverse Problem conference", Grenoble, France, July 2019, https://hal.archives-ouvertes.fr/hal-02415115

[35] F. FORBES, D. WRAITH.*Robust mixture modelling using skewed multivariate distributions with variable amounts of tailweight*, in "JdS 2019 - 51èmes Journées de Statistique", Nancy, France, Proceedings des 51èmes Journées de Statistique 2019, June 2019, https://hal.archives-ouvertes.fr/hal-02423639

[36] S. GIRARD.*Un aperçu des méthodes statistiques pour la classification et la régression en grande dimension*, in "Workshop "Appréhender la grande dimension" 2019", Paris, France, June 2019, https://hal.inria.fr/hal-02149891

[37] S. GIRARD, G. STUPFLER.*Estimation of high-dimensional extreme conditional expectiles*, in "CRoNoS & MDA 2019 - Final CRoNoS meeting and 2nd workshop on Multivariate Data Analysis", Limassol, Cyprus, April 2019, https://hal.inria.fr/hal-02099370

[38] A. USSEGLIO-CARLEVE, S. GIRARD, G. STUPFLER.*Nonparametric extreme conditional expectile estimation*, in "CMStatistics 2019 - 12th International Conference of the ERCIM WG on Computational and Methodological Statistics", London, United Kingdom, December 2019, https://hal.archives-ouvertes.fr/hal-02413682

International Conferences with Proceedings

[39] K. ASHURBEKOVA, S. ACHARD, F. FORBES.*Structure Learning via Hadamard Product of Correlation and Partial Correlation Matrices*, in "EUSIPCO 2019 - 27th European Signal Processing Conference", A Coruña, Spain, IEEE, September 2019, p. 1-5 [*DOI :* 10.23919/EUSIPCO.2019.8902948], http://hal.univ-grenoble-alpes.fr/hal-02290847

[40] 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, SIAM, 2019, p. 39-53, https://arxiv.org/abs/1810.10860 [*DOI :* 10.1137/1.9781611975499.4], https://hal.archives-ouvertes.fr/hal-01294013

[41] 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*, in "CCGrid 2019 - International Symposium in Cluster, Cloud, and Grid Computing", Larcana, Cyprus, May 2019, p. 1-10 [*DOI :* 10.1109/CCGRID.2019.00026], https://hal.inria.fr/hal-02110868

[42] S. GIRARD, G. STUPFLER, A. USSEGLIO-CARLEVE.*Nonparametric extreme conditional expectile estimation*, in "EVA 2019 - 11th International Conference on Extreme Value Analysis", Zagreb, Croatia, July 2019, 1, https://hal.inria.fr/hal-02186705

[43] V. MUÑOZ RAMÍREZ, F. FORBES, J. ARBEL, A. ARNAUD, M. DOJAT.*Quantitative MRI characterization of brain abnormalities in de novo Parkinsonian patients*, in "ISBI 2019 - IEEE International Symposium on Biomedical Imaging", Venice, Italy, Proceedings of IEEE International Symposium on Biomedical Imaging, April 2019, p. 1-4 [*DOI :* 10.1109/ISBI.2019.8759544], https://hal.archives-ouvertes.fr/hal-01970682

[44] V. MUÑOZ RAMÍREZ, F. FORBES, P. COUPÉ, M. DOJAT.*No Structural Differences Are Revealed by VBM in 'de novo' Parkinsonian Patients*, in "MEDINFO 2019 - 17th World Congress On Medical And Health Informatics", Lyon, France, August 2019, p. 268-272 [*DOI :* 10.3233/SHTI190225], https://hal.inria.fr/hal-02426273

[45] M. VLADIMIROVA, J. VERBEEK, P. MESEJO, J. ARBEL.*Understanding Priors in Bayesian Neural Networks at the Unit Level*, in "ICML 2019 - 36th International Conference on Machine Learning", Long Beach, United States, Proceedings of the 36th International Conference on Machine Learning, June 2019, vol. 97, p. 6458-6467, https://arxiv.org/abs/1810.05193 - 10 pages, 5 figures, ICML'19 conference, https://hal.archives-ouvertes.fr/hal-02177151

National Conferences with Proceeding

[46] C. ALBERT, A. DUTFOY, S. GIRARD.*Etude de l'erreur relative d'extrapolation associée à l'estimateur de Weissman pour les quantiles extrêmes*, in "JdS 2019 - 51èmes Journées de Statistique", Nancy, France, Société Française de Statistique, June 2019, p. 1-6, https://hal.inria.fr/hal-02149905

[47] J.-B. DURAND.*Compétitions d'analyse des données à l'Université Grenoble Alpes : motivations, organisation et retours d'expérience*, in "CFIES 2019 - Colloque francophone international sur l'enseignement de la statistique", Strasbourg, France, September 2019, p. 1-6, https://hal.inria.fr/hal-02298606

[48] B. OLIVIER, A. GUÉRIN-DUGUÉ, J.-B. DURAND.*Assessment of various initialization strategies for the Expectation-Maximization algorithm for Hidden Semi-Markov Models with multiple categorical sequences*, in "JdS 2019 - 51èmes Journées de Statistique", Vandœuvre-lès-Nancy, France, June 2019, p. 1-7, https://hal.inria.fr/hal-02129122

[49] F. ZHENG, S. BONNET, F. FORBES, M. JALBERT, S. LABLANCHE, P.-Y. BENHAMOU.*Caractérisation de la variabilité glycémique par analyse statistique multivariée*, in "GRETSI 2019 - XXVIIème Colloque francophonede traitement du signal et des images", Lille, France, Proceedings du XXVIIème Colloque francophone de traitement du signal et des images, August 2019, p. 1-4, https://hal.archives-ouvertes.fr/hal-02415082

[50] 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 2019 - Congrès annuel de la Société Francophone du Diabète", Marseille, France, Proceedings du Congrès annuel de la Société Francophone du Diabète, March 2019, https://hal.archives-ouvertes.fr/hal-01971621

Conferences without Proceedings

[51] K. ASHURBEKOVA, S. ACHARD, F. FORBES.*Robust penalized inference for Gaussian Scale Mixtures*, in "SPARS 2019 - Workshop on Signal Processing with Adaptive Sparse Structured Representations", Toulouse, France, July 2019, p. 1-2, http://hal.univ-grenoble-alpes.fr/hal-02291576

[52] *Best Paper*
M. BOUSEBATA, G. ENJOLRAS, S. GIRARD.*Bayesian estimation of natural extreme risk measures. Application to agricultural insurance*, in "IDRiM 2019 - 10th conference of the international society for Integrated Disaster Risk Management", Nice, France, October 2019, https://hal.archives-ouvertes.fr/hal-02276292.

[53] F. BOUX, F. FORBES, J. ARBEL, E. L. BARBIER.*Dictionary learning via regression: vascular MRI application*, in "CNIV 2019 - 3e Congrès National d'Imagerie du Vivant", Paris, France, February 2019, p. 1-12, https://hal.archives-ouvertes.fr/hal-02428647

[54] F. BOUX, F. FORBES, J. ARBEL, E. L. BARBIER.*Estimation de paramètres IRM en grande dimension via une régression inverse*, in "SFRMBM 2020 - 4e congrés de la Société Française de Résonance Magnétique en Biologie et Médecine", Strasbourg, France, March 2020, 1, https://hal.archives-ouvertes.fr/hal-02428679

[55] A. CONSTANTIN, M. FAUVEL, S. GIRARD, S. IOVLEFF, Y. TANGUY.*Classification de Signaux Multidimensionnels Irrégulièrement Echantillonnés*, in "2019 - Journée Jeunes Chercheurs MACLEAN du GDR MADICS", Paris, France, December 2019, p. 1-2, https://hal.archives-ouvertes.fr/hal-02394120

[56] P. S. DKENGNE, S. GIRARD, S. AHIAD.*Estimation of the extrapolation range associated with extreme-value models: Application to the assessment of sensors reliability*, in "EMS 2019 - 32nd European Meeting of Statisticians", Palerme, Italy, July 2019, https://hal.archives-ouvertes.fr/hal-02278051

[57] V. MUÑOZ RAMÍREZ, M. DOJAT, F. FORBES.*Mixture Models for the characterization of brain abnormalities in "de novo" Parkinsonian patients*, in "CNIV 2019 - 3e Congrès National d'Imagerie du Vivant", Paris, France, February 2019, p. 1-16, https://hal.inria.fr/hal-02436886

[58] V. MUÑOZ RAMÍREZ, F. FORBES, A. ARNAUD, M. DOJAT.*Brain abnormalities detection in de Novo Parkinsonian patients*, in "OHBM 2019 - 25th Annual Meeting of the Organization for Human Brain Mapping", Rome, Italy, June 2019, p. 1-11, https://hal.archives-ouvertes.fr/hal-02415101

[59] V. MUÑOZ RAMÍREZ, F. FORBES, P. COUPÉ, M. DOJAT.*No structural Brain differences in 'de novo' Parkinsonian patients*, in "OHBM 2019 - 25th Annual Meeting of the Organization for Human Brain Mapping", Rome, Italy, June 2019, p. 1-5, https://hal.archives-ouvertes.fr/hal-02192447

[60] S. SALHI, F. BONNEFOY, S. GIRARD, M. BERNIER, N. BARBOT, R. SIRAGUSA, E. PERRET, F. GARET.*Enhanced THz tags authentication using multivariate statistical analysis*, in "IRMMW-THz 2019 - 44th International Conference on Infrared, Millimeter, and Terahertz Waves", Paris, France, September 2019, p. 1-2, https://hal.archives-ouvertes.fr/hal-02282841

Scientific Books (or Scientific Book chapters)

[61] K. K. MENGERSEN, E. DUNCAN, J. ARBEL, C. ALSTON-KNOX, N. WHITE.*Applications in Industry*, in "Handbook of mixture analysis", S. FRUHWIRTH-SCHNATTER, G. CELEUX, C. P. ROBERT (editors), CRC press, January 2019, p. 1-21, https://hal.archives-ouvertes.fr/hal-01963798

Scientific Popularization

[62] V. MUÑOZ RAMÍREZ, F. FORBES, A. ARNAUD, E. MORO, M. DOJAT.*Anomaly detection in the MRI data of newly diagnosed Parkinsonian patients*, March 2019, 4e congrès de la Société Française de Résonance Magnétique en Biologie et Médecine - SFRMBM 2019, Poster, https://hal.inria.fr/hal-02436613

Other Publications

[63] J. ARBEL, R. CORRADIN, B. NIPOTI.*Dirichlet process mixtures under affine transformations of the data*, January 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01950652

[64] J. ARBEL, O. MARCHAL, B. NIPOTI.*On the Hurwitz zeta function with an application to the exponential-beta distribution*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02400451

[65] A. ARNAUD, F. FORBES, R. STEELE, B. LEMASSON, E. L. BARBIER.*Bayesian mixtures of multiple scale distributions*, September 2019, working paper or preprint, https://hal.inria.fr/hal-01953393

[66] K. ASHURBEKOVA, A. USSEGLIO-CARLEVE, F. FORBES, S. ACHARD.*Optimal shrinkage for robust covariance matrix estimators in a small sample size setting*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02378034

[67] M. BOUSEBATA, G. ENJOLRAS, S. GIRARD.*Bayesian estimation of natural extreme risk measures. Application to agricultural insurance*, June 2019, Global Challenges Science Week: International interdisciplinary days of Grenoble Alpes, Poster, https://hal.archives-ouvertes.fr/hal-02150604

[68] M. BOUSEBATA, S. GIRARD, G. ENJOLRAS.*Estimation bayésienne des mesures de risques naturels extrêmes. Application à l'assurance du risque agricole*, March 2019, 1, Assises Nationales des Risques Naturels 2019, Poster, https://hal.archives-ouvertes.fr/hal-02092358

[69] F. BOUX, F. FORBES, J. ARBEL, E. L. BARBIER.*Inverse regression in MR Fingerprinting: reducing dictionary size while increasing parameters accuracy*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02314026

[70] A. CONSTANTIN, M. FAUVEL, S. GIRARD, S. IOVLEFF.*Classification de Signaux Multidimensionnels Irrégulièrement Échantillonnés*, August 2019, GRETSI 2019 - 27e Colloque francophone de traitement du signal et des images, Poster, https://hal.archives-ouvertes.fr/hal-02276255

[71] A. CONSTANTIN, M. FAUVEL, S. GIRARD, S. IOVLEFF.*Supervised classification of multidimensional and irregularly sampled signals*, April 2019, 1, Statlearn 2019 - Workshop on Challenging problems in Statistical Learning, Poster, https://hal.archives-ouvertes.fr/hal-02092347

[72] L. GARDES, S. GIRARD.*On the estimation of the variability in the distribution tail*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02400320

[73] S. GIRARD.*Deux méthodes statistiques pour la classification et la régression en grande dimension*, June 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02156948

[74] S. GIRARD, G. STUPFLER, A. USSEGLIO-CARLEVE.*Nonparametric extreme conditional expectile estimation*, May 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02114255

[75] S. GIRARD, G. STUPFLER, A. USSEGLIO-CARLEVE.*An Lp-quantile methodology for tail index estimation*, January 2020, working paper or preprint, https://hal.inria.fr/hal-02311609

[76] B. KUGLER, F. FORBES, S. DOUTÉ.*Massive hyperspectral images analysis by inverse regression of physical models*, April 2019, StatLearn 2019 Workshop on Challenging problems in Statistical Learning, Poster, https://hal.archives-ouvertes.fr/hal-02423640

[77] H. LU, J. ARBEL, F. FORBES.*Bayesian nonparametric priors for hidden Markov random fields*, June 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02163046

[78] H. LU, F. FORBES, J. ARBEL.*Bayesian Nonparametric Priors for Graph Structured Data: Application to Image Segmentation*, January 2020, Bayes Comp 2020, Poster, https://hal.archives-ouvertes.fr/hal-02423642

[79] H. D. NGUYEN, J. ARBEL, H. LU, F. FORBES.*Approximate Bayesian computation via the energy statistic*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02399934

References in notes

[80] 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

[81] P. EMBRECHTS, C. KLÜPPELBERG, T. MIKOSH.*Modelling Extremal Events*, Applications of Mathematics, Springer-Verlag, 1997, vol. 33

[82] F. FERRATY, P. VIEU.*Nonparametric Functional Data Analysis: Theory and Practice*, Springer Series in Statistics, Springer, 2006

[83] 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

[84] K. LI.*Sliced inverse regression for dimension reduction*, in "Journal of the American Statistical Association", 1991, vol. 86, p. 316–327

[85] 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

[86] J. SIMOLA, J. SALOJÄRVI, I. KOJO.*Using hidden Markov model to uncover processing states from eye movements in information search tasks*, in "Cognitive Systems Research", Oct 2008, vol. 9, n^{o} 4, p. 237-251

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 [Univ Grenoble Alpes, Associate Professor]
Jérôme David [Univ Grenoble Alpes, Associate Professor]

PhD Students

Yasser Bourahla [Inria, PhD Student, from Oct 2019]
Khadija Jradeh [Univ Grenoble Alpes, PhD Student]
Line Van Den Berg [Univ Grenoble Alpes, PhD Student]

Administrative Assistants

Alexandra Fitzgerald [Inria, Administrative Assistant, until Sep 2019]
Julia Di Toro [Inria, Administrative Assistant, from Oct 2019]

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 from 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.) [4].

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 [4], [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. An example of a link key expression is:

$$\{\langle \mathsf{auteur}, \mathsf{creator}\rangle\}\{\langle \mathsf{titre}, \mathsf{title}\rangle\}\ linkkey\ \langle \mathsf{Livre}, \mathsf{Book}\rangle$$

stating that whenever an instance of the class Livre has the same values for the property auteur as an instance of class Book has for the property creator and they share at least one value for their property titre and title, then they denote the same entity. 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 considers how culture spreads and evolves with human societies [21]. It applies an an idealised version of the theory of evolution to culture. In computer science, cultural evolution experiments are performed through multi-agent simulation: a society of agents adapts its culture through a precisely defined protocol [16]: 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 [12], [23]. 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.

Work has recently been developed for evolving alignments between ontologies. It can be used to repair alignments better than blind logical repair [19], to create alignments based on entity descriptions [13], to learn alignments from dialogues framed in interaction protocols [14], [18], or to correct alignments until no error remains [17][3] and to start with no alignment [2]. Each study provides new insights and opens perspectives.

We adapt this experimental strategy to knowledge representation [3]. 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 their 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 that have operators that preserve diverse knowledge recover faster from the changes than those that 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. Highlights of the Year

4.1. Highlights of the Year

- We published our work on relational concept analysis applied to link key extraction in *Discrete applied mathematics* [5].
- Jérôme Euzenat was invited to deliver a keynote talk at the International semantic web conference (ISWC), in Aukland (NZ). The title of the talk was a call to brains: *For knowledge!*
- The teams leads the *Knowledge communication and evolution* chair of the Multidisciplinary Institute of Artificial Intelligence awarded in Grenoble.

5. New Software and Platforms

5.1. Lazylav

Lazy lavender

KEYWORDS: Reproducibility - Multi-agent - Simulation

SCIENTIFIC DESCRIPTION: Lazy lavender aims at supporting mOeX's research on simulating knowledge evolution. It is not a general purpose simulator. However, it features some methodological innovations in term of facilitating publication, recording, and replaying of experiments.

FUNCTIONAL DESCRIPTION: Lazy Lavender is a simulation environment for cultural knowledge evolution, i.e. running randomised experiments with agent adjusting their knowledge while attempting to communicate. It can generate detailed report and data from the experiments and directions to repeat them.

NEWS OF THE YEAR: In 2019, we implemented facilities for dealing with population of agents and designed several synchronisation approaches. This led to seriously refactor the code.

- Participant: Jérôme Euzenat
- Contact: Jérôme Euzenat
- Publications: Crafting ontology alignments from scratch through agent communication - Interaction-based ontology alignment repair with expansion and relaxation - First experiments in cultural alignment repair (extended version)
- URL: http://lazylav.gforge.inria.fr

5.2. Alignment API

KEYWORDS: Ontologies - Alignment - Ontology engineering - Knowledge representation

SCIENTIFIC DESCRIPTION: The api itself is a Java description of tools for accessing the common format. It defines five main interfaces (OntologyNetwork, Alignment, Cell, Relation and Evaluator).

We provide an implementation for this api which can be used for producing transformations, rules or bridge axioms independently from the algorithm that produced the alignment. The proposed implementation features: - a base implementation of the interfaces with all useful facilities, - a library of sample matchers, - a library of renderers (XSLT, RDF, SKOS, SWRL, OWL, C-OWL, SPARQL), - a library of evaluators (various generalisation of precision/recall, precision/recall graphs), - a flexible test generation framework that allows for generating evaluation data sets, - a library of wrappers for several ontology APIs , - a parser for the format.

To instanciate the API , it is sufficient to refine the base implementation by implementing the align() method. Doing so, the new implementation will benefit from all the services already implemented in the base implementation.

FUNCTIONAL DESCRIPTION: Using ontologies is the privileged way to achieve interoperability among heterogeneous systems within the Semantic web. However, as the ontologies underlying two systems are not necessarily compatible, they may in turn need to be reconciled. Ontology reconciliation requires most of the time to find the correspondences between entities (e.g. classes, objects, properties) occurring in the ontologies. We call a set of such correspondences an alignment.

NEWS OF THE YEAR: Link keys are fully supported by the EDOAL language. In particular it can transform them into SPARQL queries.

- Participants: Armen Inants, Chan Le Duc, Jérôme David, Jérôme Euzenat, Jérôme Pierson, Luz Maria Priego-Roche and Nicolas Guillouet
- Contact: Jérôme Euzenat
- Publications: An API for ontology alignment - The Alignment API 4.0
- URL: http://alignapi.gforge.inria.fr/

6. New Results

6.1. 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 [4], 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 [3] and revealed that, by playing simple interaction games, agents can effectively repair random networks of ontologies or even create new alignments.

6.1.1. Modelling in dynamic epistemic logic

Participants: Manuel Atencia, Jérôme Euzenat, Line Van Den Berg [Correspondent].

We explored how closely these operators resemble logical dynamics. We developed a variant of Dynamic Epistemic Logic to capture the dynamics of the cultural alignment repair game. The ontologies are modelled as knowledge and alignments as beliefs in a variant of plausibility-based dynamic epistemic logic. The dynamics of the game is achieved through (public) announcement of the game issue and the adaptation operators are defined through conservative upgrades, i.e. modalities that transform models by reordering world-plausibility. This allowed us to formally establish some limitations and redundancy of the operators [9]. More precisely, for a complete logical reasoner, the operators are redundant and some may be inconsistent with the agent knowledge.

These results hold for one agent in the game but not necessarily for the other that may not know the classes by which the alignment is repaired, nor the relations between them. The former can be dealt with by declaring that agents are aware of the signature of both ontologies (public signature assumption) but this does not allow ontologies to evolve. We are currently investigating partial semantics as a more dynamic alternative solution to this problem.

This work is part of the PhD thesis of Line van den Berg.

6.1.2. Populations

Participants: Manuel Atencia, Fatima Danash, Jérôme Euzenat [Correspondent].

We started taking the population standpoint on experimental cultural evolution. For that purpose we introduced the concept of population within the experiments. So far, a population is characterised as a set of agents sharing the same ontology. Such agents play the same alignment repair games as before with agents of other populations.

The notion of population enables to experiment with different transmission mechanisms found in cultural evolution: vertical transmission, in which culture spreads, like genes, from parents to siblings, and horizontal transmission, in which it spreads among all members of a population. We implemented explicit horizontal transmission through a synchronisation procedure in which, at a given interval, agents of the same population exchange their knowledge, i.e. alignments.

6.1.3. Link with interactor-replicator

Participant: Jérôme Euzenat [Correspondent].

Cultural evolution may be studied at a 'macro' level, inspired from population dynamics, or at a 'micro' level, inspired from genetics. The replicator-interactor model generalises the genotype-phenotype distinction of genetic evolution. We considered how it can be applied to cultural knowledge evolution experiments [8]. More specifically, we consider knowledge as the replicator and the behaviour it induces as the interactor. We showed that this requires to address problems concerning transmission. We discussed the introduction of horizontal transmission within the replicator-interactor model and/or differential reproduction within cultural evolution experiments.

6.1.4. Experiment reproducibility

Participants: Jimmy Avae, Robin Couret, Jérôme Euzenat [Correspondent].

Experiments are described and performed in our *Lazy lavender* platform which offers scripts to specify, run, and analyse experiments. This year, we investigated expressing experiment descriptions, i.e. design, results and analysis, in RDF. This facilitates the search of experiments based on structured queries that can be expressed in SPARQL: "which experiments have been performed but not analysed?", "which experiments are derived from another specific experiment?", "which hypotheses have not been confirmed since a precise release?", "which experiments test F-measure increase?". This also suggest a better organisation of our experiment reports.

6.2. Link keys

Link keys (§3.2) are explored following two directions:

- Extracting link keys;
- Reasoning with link keys.

6.2.1. Link key extraction with relational concept analysis

Participants: Manuel Atencia, Jérôme David [Correspondent], Jérôme Euzenat.

We first described our extraction approach [1] in the framework of formal context analysis (FCA, [20]). We recently showed that link keys extracted by formal concept analysis are equivalent to an extension of those which were extracted by our former algorithm [15]. We also used pattern structures, an extension of FCA with ordered structures, to reformulate this problem [6].

Furthermore, we used relational concept analysis (RCA, [22]), an extension of FCA taking relations between concepts into account. We showed that it is possible to encode the link key extraction problem in RCA to extract the optimal link keys even in the presence of cyclic dependencies [5]. Moreover, the proposed process does not require information about the alignments between the ontologies to find out from which pairs of classes to extract link keys.

We implemented these methods and evaluated them by reproducing the experiments made in previous studies. This shows that the method extracts the expected results as well as (also expected) scalability issues.

6.2.2. Combining link keys

Participants: Manuel Atencia, Alice Caporali, Jérôme David [Correspondent], Jérôme Euzenat, Basile Legal.

For certain data sets, it may be necessary to use several link keys, even on the same pair of classes, for retrieving a more complete link set. We introduced operators to combine link keys over the same pair of classes, investigated their relations and extended measures to evaluate their quality.

We specifically proposed strategies to extract disjunctions from RDF data and apply existing quality measures to evaluate them. We experimented with these strategies showing their benefits [7].

6.2.3. 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. As such, they can contribute to infer ABox axioms, such as links, terminological axioms, or other link keys. This has important practical applications, such as link key inference, link key consistency and link key redundancy checking. Yet, no reasoning support existed for link keys.

We previously extended the tableau method designed for the $\mathcal{ALC}$ description logic to support reasoning with link keys in $\mathcal{ALC}$. We showed how this extension enables combining link keys with classical terminological reasoning with and without ABox and TBox and generating non-trivial link keys. We further extended the method and have proven that this extended method terminates, is sound, complete, and that its complexity is 2ExpTime [11].

This work is part of the PhD thesis of Khadija Jradeh, co-supervised with Chan Le Duc (LIASD).

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. ANR Elker

Program: ANR-PRC

Project acronym: Elker

Project title: Extending link keys: extraction and reasoning

Web site: https://project.inria.fr/elker/

Duration: October 2017 - September 2021

Coordinator: LIG/Manuel Atencia

Participants: Manuel Atencia Arcas, Jérôme David, Jérôme Euzenat

Other partners: Inria Lorraine, Université de Vincennes+Université Paris 13

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.

7.1.2. PEPS RegleX-LD

Program: Projets Exploratoires Premier Soutien (CNRS, INS2I)

Project acronym: RegleX-LD

Project title: Découverte de règles expressives de correspondances complexes et de liage de données

Duration: January 2019 – December 2019

Coordinator: IRIT/Cássia Trojahn

Participants: Manuel Atencia Arcas, Jérôme David, Jérôme Euzenat

Other partners: IRIT Toulouse, INRA Paris, LRI Orsay

Abstract: RegleX-LD aims at discovering expressive ontology correspondences and data interlinking patterns using unsupervised or weakly supervised methods.

7.2. International Research Visitors

7.2.1. Visits of International Scientists

7.2.1.1. Internships

- Nacira Abbas (U. Lorraine) visited mOeX between 2019-02-04 and 2019-02-15 in the framework of the Elker project, working on link keys extraction with formal concept analysis.
- Hiba Belhadi, PhD student at Université des Sciences et de la Technologie Houari Boumediene (UTHB), Algiers, visited mOeX between 2019-10-15 and 2019-11-15 to work on selecting and matching properties for data interlinking.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events: Organisation

8.1.1.1. Member of the Organizing Committees

- Jérôme Euzenat was organiser of the 2nd Workshop on Interaction-Based Knowledge Sharing (WINKS) of the 5th Joint Ontology Workshop (JOWO), Graz (AT), 2019 (with Adrian Kemo, Dagmar Gromann, Ernesto Jiménez Ruiz, Marco Schorlemmer and Valentina Tamma) [10]
- Jérôme Euzenat had been organiser of the 14th Ontology matching workshop of the 19th ISWC, Auckland (NZ), 2019 (with Pavel Shvaiko, Ernesto Jiménez Ruiz, Cássia Trojahn dos Santos and Oktie Hassanzadeh)

8.1.2. Scientific Events: Selection

8.1.2.1. Member of the Conference Program Committees

- Manuel Atencia, Jérôme David and Jérôme Euzenat have been programme committee members of the “International Joint Conference on Artificial Intelligence (IJCAI)”.
- Manuel Atencia, Jérôme David and Jérôme Euzenat have been programme committee member of the “Web Conference (WWW)”.
- Manuel Atencia, Jérôme David and Jérôme Euzenat have been programme committee members of the “International semantic web conference (ISWC)”.
- Jérôme Euzenat has been programme committee member of the “International Conference on Conceptual Structures (ICCS)”.
- Jérôme Euzenat has been programme committee member of the “International conference on semantic systems (Semantics)”.
- Jérôme Euzenat has been programme committee member of the “Journées Françaises d’intelligence artificielle fondamentale (JIAF)”.
- Manuel Atencia and Jérôme David have been programme committee members of the “Extraction et Gestion des connaissances (EGC)”.

8.1.3. Journal

8.1.3.1. 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*.

8.1.3.2. Reviewer - Reviewing Activities

- Manuel Atencia had been reviewer for *Journal of web semantics*.
- Jérôme David had been reviewer for *Journal of web semantics*.
- Jérôme Euzenat had been reviewer for *Knowledge engineering reviews*.

8.1.4. Invited Talks

- "For knowledge", ISWC keynote speech, Auckland (NZ), 2019-10-29 (Jérôme Euzenat)

8.1.5. Leadership within the Scientific Community

- Jérôme Euzenat is member of the scientific council of the CNRS GDR on Formal and Algorithmic Aspects of Artificial intelligence.
- Jérôme Euzenat is EurAI fellow.
- Jérôme David is member of the board of the Extraction and gestion des connaissances (Knowledge extraction and management) conference series.

8.1.6. Scientific Expertise

- Jérôme Euzenat had been member of the HCERES visiting committee for the CRIL research laboratory.
- Jérôme Euzenat had been member of the recruitment committee of Université Grenoble Alpes for the associate professor position 27MCF330, 2019.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

8.2.1.1. Responsibilities

- 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 2nd year of Master "Mathematiques et informatiques appliquées aux sciences humaines et sociales" (Univ. Grenoble Alpes);
- Manuel Atencia is coordinator of the "Web, Informatique et Connaissance" option of the master M2 "Mathematicques et informatiques appliquées aux sciences humaines et sociales" (Univ. Grenoble Alpes);

8.2.1.2. Lectures

- Licence: Jérôme David, Algorithmique et programmation par objets, 70h/y, L2 MIASHS, UGA, France
- Licence: Jérôme David, Système, L3 MIASHS, 18h/y, UGA, France
- Licence: Manuel Atencia, Introduction aux technologies du Web, 60h/y, L3 MIASHS, UGA, France
- Master: Jérôme David, Programmation Java 2, 30h/y, M1 MIASHS, UGA, France
- Master: Jérôme David, JavaEE, 30h/y, M2 MIASHS, UGA, France
- Master: Jérôme David, Développement Web Mobile, 30h/y, M2 MIASHS, UGA, France
- Master: Jérôme David, Web sémantique, 3h/y, M2 MIASHS, UGA, France
- Master: Manuel Atencia, Formats de données du web, 30h/y, M1 MIASHS, UGA, France

- Master: Manuel Atencia, Introduction à la programmation web, 30h/y, M1 MIASHS, UGA, France
- Master: Manuel Atencia, Intelligence artificielle, 7.5h/y, M1 MIASHS, UGA, France
- Master: Manuel Atencia, Web sémantique, 27h/y, M2 MIASHS, UGA, France
- Master: Manuel Atencia, Semantic web: from XML to OWL, 22.5h/y, M2R MoSIG, UGA, France
- Master: Jérôme David, Stage de programmation, 10h/y, M2 MIASHS, UGA, France

8.2.2. Supervision

- Nacira Abbas, "Link key extraction and relational concept analysis", in progress since 2018-10-01 (Jérôme David and Amedeo Napoli)
- Khadija Jradeh, "Reasoning with link keys", in progress since 2018-10-01 (Manuel Atencia and Chan Le Duc)
- Line van den Berg, "Knowledge Evolution in Agent Populations", in progress since 2018-10-01 (Manuel Atencia and Jérôme Euzenat)
- Yasser Bourahla, "Evolving ontologies through communication", in progress since 2019-10-01 (Manuel Atencia and Jérôme Euzenat)

8.2.3. Juries

- Jérôme Euzenat had been reviewer and panel chair of the computer science PhD of Élodie Thiéblin (Université Toulouse 3 Paul Sabatier) "Automatic generation of complex ontology alignments" supervised by Ollivier Haemmerlé and Cássia Trojahn, 2019
- Jérôme Euzenat had been member of the computer science habilitation (HDR) panel of Cássia Trojahn (Université Toulouse 2 Jean Jaurès) "Towards ontology matching maturity: contributions to complex, holistic and foundational ontology matching", 2019

8.3. Popularization

8.3.1. Interventions

- Présentation "L'intelligence artificielle en perspective", Université Inter-Âges du Dauphiné (UIAD), Grenoble (FR), 2019-03-06 (Jérôme Euzenat)
- Introduction of the *Class?* game to a fourth graders (CM1-CM2) class, Montbonnot (FR), 2019-06-25 (Line van den Berg and Jérôme Euzenat)
- Introduction of the *Class?* game to a tenth graders (2^{nd} MathC2+) group, Montbonnot (FR), 2019-06-25 (Line van den Berg and Jérôme Euzenat)
- Introduction of the *Class?* game to seventh and eighth graders (5^{e} et 4^{e}) classes within the Fête de la science (Science fair), Montbonnot (FR), 2019-10-10 (Line van den Berg and Jérôme Euzenat)
- Introduction of the *Class?* game to general public within the Fête de la science (Science fair), Montbonnot (FR), 2019-10-12 (Line van den Berg and Jérôme Euzenat)
- Presentation of the *Class?* game to the meeting of the Société Francophone de Classification, Nancy (FR), 2019-09-03–04 (Jérôme Euzenat)

8.3.2. Creation of media or tools for science outreach

Class? We are developing mediation material for explaining to the general public what knowledge representation is and how it may evolve. Its main goal is to show children that the same individuals may be classified in different and evolving ways and that it is possible to communicate such classifications without expressing them. For that purpose, we have designed a card game called *Class?*[0] which allows players to guess the hidden ontology of another player. It has been presented to school classes from year 5 of primary school (fourth graders) to year 11 (tenth graders), albeit shows features of interest for a wider audience.

[0] https://moex.inria.fr/mediation/class/

9. 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.*Crafting ontology alignments from scratch through agent communication*, in "PRIMA 2017: Principles and Practice of Multi-Agent Systems", Nice, France, B. AN, A. BAZZAN, J. LEITE, S. VILLATA, L. VAN DER TORRE (editors), Springer Verlag, 2017, p. 245-262

[3] 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

[4] 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

[5] M. ATENCIA, J. DAVID, J. EUZENAT, A. NAPOLI, J. VIZZINI.*Link key candidate extraction with relational concept analysis*, in "Discrete Applied Mathematics", 2019, p. 1-19 [*DOI :* 10.1016/J.DAM.2019.02.012], https://hal.archives-ouvertes.fr/hal-02196757

International Conferences with Proceedings

[6] N. ABBAS, J. DAVID, A. NAPOLI.*Linkex: A Tool for Link Key Discovery Based on Pattern Structures*, in "ICFCA 2019 - workshop on Applications and tools of formal concept analysis", Frankfurt, Germany, Proc. ICFCA workshop on Applications and tools of formal concept analysis, 2019, p. 33-38, abbas2019a, https://hal.archives-ouvertes.fr/hal-02168775

[7] M. ATENCIA, J. DAVID, J. EUZENAT.*Several link keys are better than one, or extracting disjunctions of link key candidates*, in "K-CAP 2019 - 10th ACM international conference on knowledge capture (K-Cap)", Marina del Rey, United States, No commercial editor., 2019, p. 61-68 [*DOI :* 10.1145/3360901.3364427], https://hal.archives-ouvertes.fr/hal-02395703

[8] J. EUZENAT.*Replicator-interactor in experimental cultural knowledge evolution*, in "WINKS 2019 - 2nd JOWO workshop on Interaction-Based Knowledge Sharing", Graz, Austria, No commercial editor., 2019, p. 1-11, https://hal.archives-ouvertes.fr/hal-02414736

[9] L. VAN DEN BERG.*Epistemic alignment repair*, in "ESSLLI 2019 - 31st edition European Summer School in Logic, Language and Information - Student Session", Riga, Latvia, No commercial editor., 2019, p. 1-12, https://hal.archives-ouvertes.fr/hal-02395710

Scientific Books (or Scientific Book chapters)

[10] *Proc. 2nd JOWO workshop on Interaction-Based Knowledge Sharing (WINKS)*, No commercial editor., Graz, Austria, 2019, p. 1-48, https://hal.archives-ouvertes.fr/hal-02414685

Research Reports

[11] M. ATENCIA, J. EUZENAT, C. LE DUC, K. JRADEH.*Reasoning for the description logic ALC with link keys*, Laboratoire d'Informatique de Grenoble ; Inria Grenoble Rhône-Alpes ; Université Paris 8, April 2019, p. 1-36, https://hal.archives-ouvertes.fr/hal-02090087

References in notes

[12] L. STEELS (editor). *Experiments in cultural language evolution*, John Benjamins, Amsterdam (NL), 2012

[13] M. ANSLOW, M. ROVATSOS.*Aligning experientially grounded ontologies using language games*, in "Proc. 4th international workshop on graph structure for knowledge representation, Buenos Aires (AR)", 2015, p. 15–31

[14] M. ATENCIA, M. SCHORLEMMER.*An interaction-based approach to semantic alignment*, in "Journal of web semantics", 2012, vol. 13, n^{o} 1, p. 131–147

[15] MANUEL. ATENCIA, JÉRÔME. DAVID, JÉRÔME. EUZENAT.*Data interlinking through robust linkkey extraction*, in "Proc. 21st ECAI, Praha (CZ)", 2014, p. 15–20

[16] 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

[17] P. CHOCRON, M. SCHORLEMMER.*Attuning ontology alignments to semantically heterogeneous multi-agent interactions*, in "Proc. 22nd European conference on artificial intelligence (ECAI), The Hague (NL)", 2016, p. 871–879

[18] P. CHOCRON, M. SCHORLEMMER.*Vocabulary alignment in openly specified interactions*, in "Proc. 16th International conference on autonomous agents and multi-agent systems (AAMAS), Saõ Paolo (BR)", 2017, p. 1064–1072, http://www.aamas-conference.org/Proceedings/aamas2017/pdfs/p1064.pdf

[19] JÉRÔME. EUZENAT.*First experiments in cultural alignment repair (extended version)*, in "Proc. ESWC 2014 satellite events revised selected papers", Lecture notes in computer science, 2014, n^{o} 8798, p. 115-130

[20] B. GANTER, R. WILLE.*Formal concept analysis: mathematical foundations*, Springer, 1999

[21] A. MESOUDI, A. WHITEN, K. LALAND.*Towards a unified science of cultural evolution*, in "Behavioral and brain sciences", 2006, vol. 29, n^{o} 4, p. 329–383

[22] M. ROUANE HACENE, M. HUCHARD, A. NAPOLI, P. VALTCHEV.*Relational Concept Analysis: mining concept lattices from multi-relational data*, in "Annals of mathematics and artificial intelligence", 2013, vol. 67, n^{o} 1, p. 81–108

[23] 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, until Mar 2019]
Stefanie Wuhrer [Inria, Researcher, HDR]

Faculty Members

Jean-Sébastien Franco [Institut polytechnique de Grenoble, Associate Professor]
Sergi Pujades Rocamora [Univ Grenoble Alpes, Associate Professor]

Technical Staff

Eymeric Amselem [Inria, Engineer, until Jun 2019]
Laurence Boissieux [Inria, Engineer]
Julien Pansiot [Inria, Engineer, from Oct 2019]
Tomas Svaton [Inria, Engineer]

PhD Students

Matthieu Armando [Inria, PhD Student]
Jean Basset [Inria, PhD Student]
Sanae Dariouche [Univ Grenoble Alpes, PhD Student, from Nov 2019]
Victoria Fernandez Abrevaya [Univ Grenoble Alpes, PhD Student]
Claude Goubet [Inria, PhD Student, until Sep 2019]
Roman Klokov [Inria, PhD Student]
Vincent Leroy [Inria, PhD Student, until Mar 2019]
Mathieu Marsot [Inria, PhD Student, from Nov 2019]
Di Meng [Inria, PhD Student]
Abdullah Haroon Rasheed [Inria, PhD Student]

Nitika Verma [Univ Grenoble Alpes, PhD Student]
Boyao Zhou [Inria, PhD Student]
Pierre Zins [Inria, PhD Student, from Oct 2019]
Jinlong Yang [Inria, PhD Student, until Mar 2019]

Visiting Scientist

Stephane Durocher [Université du Manitoba (Canada), from Jul 2019]

Administrative Assistant

Nathalie Gillot [Inria, Administrative Assistant]

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 is 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 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, 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 emerge 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

5.1.1. Awards

BEST PAPERS AWARDS :

[15]

R. KLOKOV, J. VERBEEK, E. BOYER.*Probabilistic Reconstruction Networks for 3D Shape Inference from a Single Image*, in "BMVC 2019 - British Machine Vision Conference", Cardiff, United Kingdom, September 2019, p. 1-15, https://arxiv.org/abs/1908.07475 - Awarded with Best Science Paper Honourable Mention Award at BMVC'19., https://hal.inria.fr/hal-02268466

6. New Software and Platforms

6.1. 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/

6.2. 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

6.3. 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/

6.4. 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/

6.5. Adaptive mesh texture

KEYWORDS: 3D - Geometry Processing - Texturing

FUNCTIONAL DESCRIPTION: Tool for computing appearance information on a 3D scene acquired with a multi-view stereo (MVS) pipeline. Appearance information is sampled in an adaptive way so as to maximize the entropy of stored information. This is made possible through a homemade representation of appearance, different from the more traditional texture maps. This tool laso includes a compression module, so as to optimize disk space.

RELEASE FUNCTIONAL DESCRIPTION: 1st version

- Authors: Matthieu Armando, Edmond Bover, Jean-Sébastien Franco and Vincent Leroy
- Partner: Microsoft
- Contact: Matthieu Armando
- URL: https://gitlab.inria.fr/marmando/adaptive-mesh-texture

7. New Results

7.1. Surface Motion Capture Animation Synthesis

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 2).

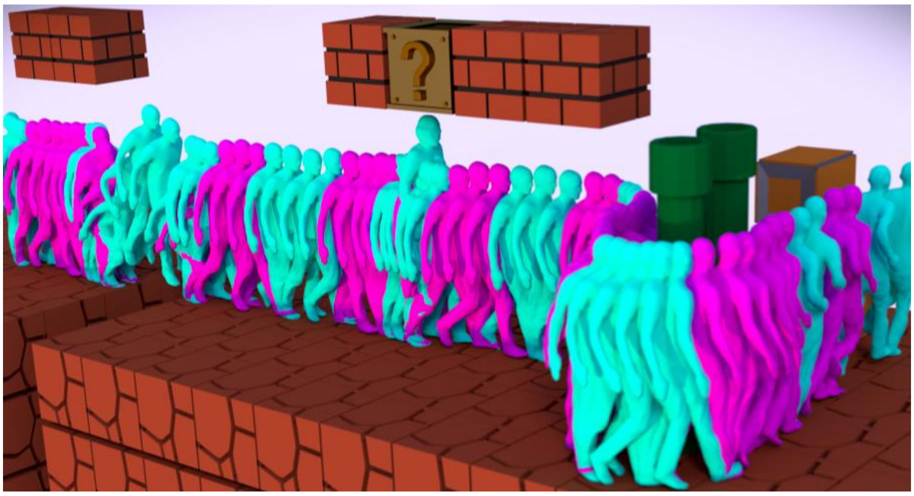


Figure 2. Animation Synthesis

This result was published in a prominent computer graphics journal, IEEE Transactions on Visualization and Computer Graphics [7].

7.2. CBCT of a Moving Sample from X-rays and Multiple Videos

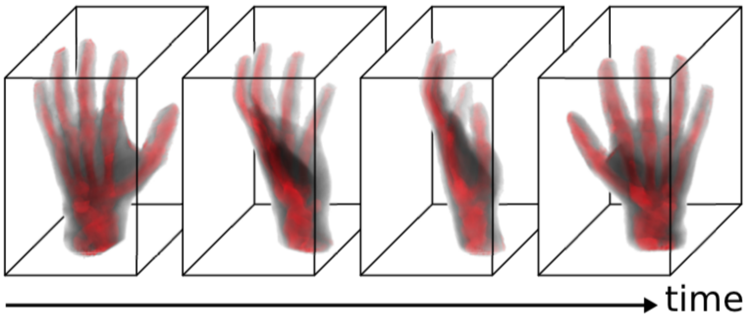


Figure 3. Dense volumetric attenuation reconstruction from a rigidly moving sample captured by a single planar X-ray imaging device and a surface motion capture system. Higher attenuation (here bone structure) is highlighted in red.

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 3.

This result was published in a prominent medical journal, IEEE Transactions on Medical Imaging [9].

7.3. Learning and Tracking the 3D Body Shape of Freely Moving Infants from RGB-D sequences

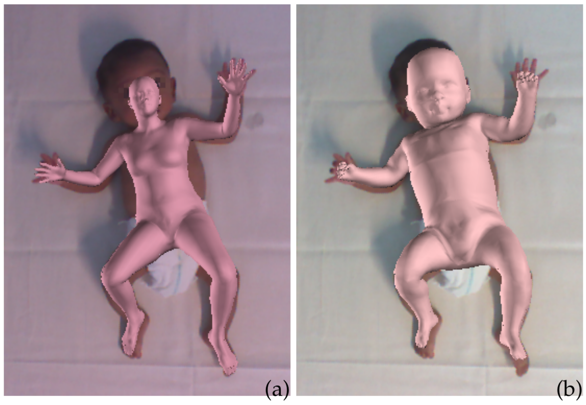


Figure 4. (a) Simply scaling a generic adult body model and fitting it to an infant does not work as body proportions significantly differ. (b) The proposed SMIL model properly captures the infants' shape and pose

Statistical models of the human body surface are generally learned from thousands of high-quality 3D scans in predefined poses to cover the wide variety of human body shapes and articulations. Acquisition of such data requires expensive equipment, calibration procedures, and is limited to cooperative subjects who can understand and follow instructions, such as adults. We presented a method for learning a statistical 3D Skinned Multi-Infant Linear body model (SMIL) from incomplete, low-quality RGB-D sequences of freely moving infants. Quantitative experiments show that SMIL faithfully represents the RGB-D data and properly factorizes the shape and pose of the infants. To demonstrate the applicability of SMIL, we fitted the model to RGB-D sequences of freely moving infants and show, with a case study, that our method captures enough motion detail for General Movements Assessment (GMA), a method used in clinical practice for early detection of neurodevelopmental disorders in infants. SMIL provides a new tool for analyzing infant shape and movement and is a step towards an automated system for GMA. This result was published in a prominent computer vision journal, IEEE Transactions on PAMI [8].

7.4. The Virtual Caliper: Rapid Creation of Metrically Accurate Avatars from 3D Measurements

Creating metrically accurate avatars is important for many applications such as virtual clothing try-on, ergonomics, medicine, immersive social media, telepresence, and gaming. Creating avatars that precisely represent a particular individual is challenging however, due to the need for expensive 3D scanners, privacy issues

Figure 5. Using the wand controllers of the HTC Vive, the Virtual Caliper produces a rigged 3D model with exactly the dimensions of the measured person.

with photographs or videos, and difficulty in making accurate tailoring measurements. We overcome these challenges by creating “The Virtual Caliper”, which uses VR game controllers to make simple measurements. First, we establish what body measurements users can reliably make on their own body. We find several distance measurements to be good candidates and then verify that these are linearly related to 3D body shape as represented by the SMPL body model. The Virtual Caliper enables novice users to accurately measure themselves and create an avatar with their own body shape. We evaluate the metric accuracy relative to ground truth 3D body scan data, compare the method quantitatively to other avatar creation tools, and perform extensive perceptual studies. We also provide a software application to the community that enables novices to rapidly create avatars in fewer than five minutes. Not only is our approach more rapid than existing methods, it exports a metrically accurate 3D avatar model that is rigged and skinned.

This result was published in a prominent computer graphics journal, IEEE Transactions on Visualization and Computer Graphics [10].

7.5. Adaptive Mesh Texture for Multi-View Appearance Modeling

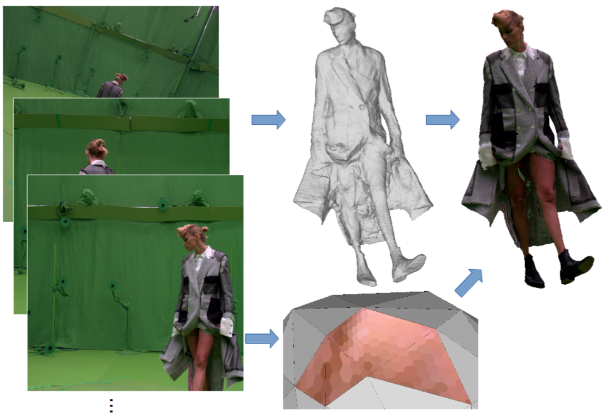


Figure 6. Texturing 3D models: given a set of input photographs (left), a geometric mesh is computed (top), along with an appearance function stored within the surface mesh structure (bottom).

Most applications in image based 3D modeling resort to texture maps, a 2D mapping of shape color information into image files. Despite their unquestionable merits, in particular the ability to apply standard image tools, including compression, image textures still suffer from limitations that result from the 2D mapping of information that originally belongs to a 3D structure. This is especially true with 2D texture atlases, a generic 2D mapping for 3D mesh models that introduces discontinuities in the texture space and plagues many 3D appearance algorithms. Moreover, the per-triangle texel density of 2D image textures cannot be individually adjusted to the corresponding pixel observation density without a global change in the atlas mapping function. To address these issues, we have proposed a new appearance representation for image-based 3D shape modeling, which stores appearance information directly on 3D meshes, rather than a texture atlas. We have shown this representation to allow for input-adaptive sampling and compression support. Our experiments demonstrated that it outperforms traditional image textures, in multi-view reconstruction contexts, with better visual quality and memory foot- print, which makes it a suitable tool when dealing with large amounts of data as with dynamic scene 3D models.

This result was published in the international conference on 3D Vision (3DV'19) [11].

7.6. Contact Preserving Shape Transfer for Motion Retargeting

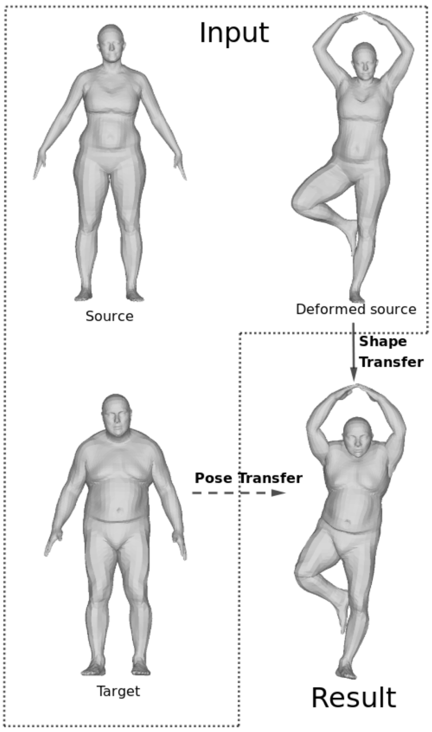


Figure 7. Motion Retargeting: Instead of transferring the pose from a source to a target shape, we propose to transfer the shape of the target to the deformed source character.

Retargeting a motion from a source to a target character is an important problem in computer animation, as it allows to reuse existing rigged databases or transfer motion capture to virtual characters. Surface based pose transfer is a promising approach to avoid the trial-and-error process when controlling the joint angles. In this work we investigated whether shape transfer instead of pose transfer would better preserve the original contextual meaning of the source pose. To this end, we proposed an optimization-based method to deform the source shape+pose using three main energy functions: similarity to the target shape, body part volume preservation, and collision management (preserve existing contacts and prevent penetrations). The results show that this strategy is able to retarget complex poses, including several contacts, to very different morphologies. In particular, we introduced new contacts that are linked to the change in morphology, and which would be difficult to obtain with previous works based on pose transfer that aim at distance preservation between body parts.

This result was published in the ACM SIGGRAPH Conference on Motion Interaction and Games (MIG'19) [12].

7.7. A Decoupled 3D Facial Shape Model by Adversarial Training

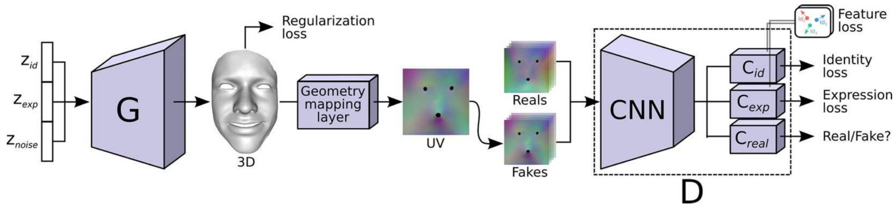


Figure 8. The face generator. Identity and expression codes zid , zexp are used to control the generator, and classification losses are added to decouple between the two. A feature loss is introduced to ensure consistency over features with fixed identities or expressions

Data-driven generative 3D face models are used to compactly encode facial shape data into meaningful parametric representations. A desirable property of these models is their ability to effectively decouple natural sources of variation, in particular identity and expression. While factorized representations have been proposed for that purpose, they are still limited in the variability they can capture and may present modeling artifacts when applied to tasks such as expression transfer. In this work, we explored a new direction with Generative Adversarial Networks and showed that they contribute to better face modeling performances, especially in decoupling natural factors, while also achieving more diverse samples. To train the model we introduced a novel architecture that combines a 3D generator with a 2D discriminator that leverages conventional CNNs, where the two components are bridged by a geometry mapping layer. We further presented a training scheme, based on auxiliary classifiers, to explicitly disentangle identity and expression attributes. Through quantitative and qualitative results on standard face datasets, we illustrated the benefits of our model and demonstrate that it outperforms competing state of the art methods in terms of decoupling and diversity.

This result was published in the international conference on computer vision (ICCV'19) [13]

7.8. Non-parametric 3D Human Shape Estimation from Single Images

In this work, we tackle the problem of 3D human shape estimation from single RGB images. While the recent progress in convolutional neural networks has allowed impressive results for 3D human pose estimation, estimating the full 3D shape of a person is still an open issue. Model-based approaches can output precise meshes of naked under-cloth human bodies but fail to estimate details and un-modelled elements such as hair or clothing. On the other hand, non-parametric volumetric approaches can potentially estimate complete shapes but, in practice, they are limited by the resolution of the output grid and cannot produce detailed estimates. In this work, we propose a non-parametric approach that employs a double depth map to represent the 3D shape of a person: a visible depth map and a "hidden" depth map are estimated and combined, to reconstruct

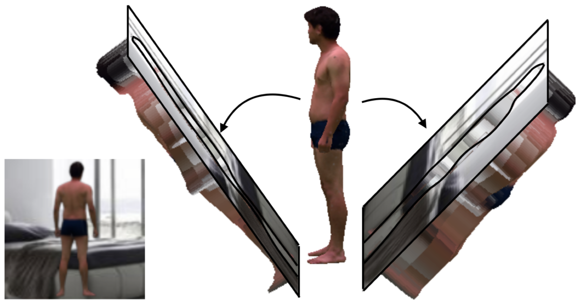


Figure 9. Given a single image, we estimate the "visible" and the "hidden" depth maps from the camera point of view. The two depth maps can be seen as the two halves of a virtual "mould".

the human 3D shape as done with a "mould". This representation through 2D depth maps allows a higher resolution output with a much lower dimension than voxel-based volumetric representations. Additionally, our fully derivable depth-based model allows us to efficiently incorporate a discriminator in an adversarial fashion to improve the accuracy and "humanness" of the 3D output. We train and quantitatively validate our approach on SURREAL and on 3D-HUMANS, a new photorealistic dataset made of semi-synthetic in-house images annotated with 3D ground truth surfaces.

This work was published in the international conference on computer vision (ICCV'19) [14]

7.9. Probabilistic Reconstruction Networks

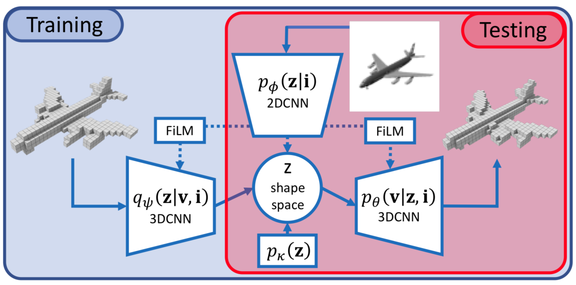


Figure 10. Probabilistic Reconstruction Networks for 3D shape inference from a single image. Arrows show the computational flow through the model, dotted arrows show optional image conditioning. The inference network q_Ψ is only used during training for variational inference

We study end-to-end learning strategies for 3D shape inference from images, in particular from a single image. Several approaches in this direction have been investigated that explore different shape representations and suitable learning architectures. We focus instead on the underlying probabilistic mechanisms involved and contribute a more principled probabilistic inference-based reconstruction framework, which we coin Probabilistic Reconstruction Networks. This framework expresses image conditioned 3D shape inference through a family of latent variable models, and naturally decouples the choice of shape representations from the inference itself. Moreover, it suggests different options for the image conditioning and allows training in two regimes, using either Monte Carlo or variational approximation of the marginal likelihood. Using our Probabilistic Reconstruction Networks we obtain single image 3D reconstruction results that set a new state of the art on the ShapeNet dataset in terms of the intersection over union and earth mover's distance evaluation metrics. Interestingly, we obtain these results using a basic voxel grid representation, improving over recent work based on finer point cloud or mesh based representations.

This work was published in the British machine vision conference (BMVC'19) [15] where it won the runner-up best paper award.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

1. The Morpheo Inria team and Microsoft research set up a collaboration on the capture and modelling of moving shapes using multiple videos. Two PhD works are part of this collaboration with the objective to make contributions on 4D Modeling. The PhDs take place at Inria Grenoble Rhône-Alpes and involve visits and stays at Microsoft in Cambridge (UK) and Zurich (CH). The collaboration is part of the Microsoft Research - Inria Joint Centre.
2. The Morpheo Inria team has another collaboration with Facebook reality lab in San Francisco. The collaboration involves one PhD who is currently at the Inria Grenoble Rhône-Alpes working on the estimation of shape and appearance from a single image. The collaboration started in 2019.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Data Driven 3D Vision

Edmond Boyer obtained a chair in the new Multidisciplinary Institute in Artificial Intelligence (MIAI) of Grenoble Alpes University. The chair entitled Data Driven 3D Vision is for 4 years and aims at investigating deep learning for 3D artificial vision in order to break some of the limitations in this domain. Applications are especially related to humans and to the ability to capture and analyze their shapes, appearances and motions, for upcoming new media devices, sport and medical applications.

9.2. National Initiatives

9.2.1. ANR

9.2.1.1. 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 and is currently handled by Tomas Svaton, recruited as an engineer in April 2018.

9.2.1.2. ANR JCJC SEMBA – Shape, Motion and Body composition to Anatomy

Existing medical imaging techniques, such as Computed Tomography (CT), Dual Energy X-Ray Absorption (DEXA) and Magnetic Resonance Imaging (MRI), allow to observe internal tissues (such as adipose, muscle, and bone tissues) of in-vivo patients. However, these imaging modalities involve heavy and expensive equipment as well as time consuming procedures. External dynamic measurements can be acquired with optical scanning equipment, e.g. cameras or depth sensors. These allow high spatial and temporal resolution acquisitions of the surface of living moving bodies. The main research question of SEMBA is: "can the internal observations be inferred from the dynamic external ones only?". SEMBA's first hypothesis is that the quantity and distribution of adipose, muscle and bone tissues determine the shape of the surface of a person. However, two subjects with a similar shape may have different quantities and distributions of these tissues. Quantifying adipose, bone and muscle tissue from only a static observation of the surface of the human might be ambiguous. SEMBA's second hypothesis is that the shape deformations observed while the body performs highly dynamic motions will help disambiguating the amount and distribution of the different tissues. The dynamics contain key information that is not present in the static shape. SEMBA's first objective is to learn statistical anatomic models with accurate distributions of adipose, muscle, and bone tissue. These models are going to be learned by leveraging medical dataset containing MRI and DEXA images. SEMBA's second objective will be to develop computational models to obtain a subject-specific anatomic model with an accurate distribution of adipose, muscle, and bone tissue from external dynamic measurements only.

9.2.1.3. ANR JCJC 3DMOVE - Learning to synthesize 3D dynamic human motion

It is now possible to capture time-varying 3D point clouds at high spatial and temporal resolution. This allows for high-quality acquisitions of human bodies and faces in motion. However, tools to process and analyze these data robustly and automatically are missing. Such tools are critical to learning generative models of human motion, which can be leveraged to create plausible synthetic human motion sequences. This has the potential to influence virtual reality applications such as virtual change rooms or crowd simulations. Developing such tools is challenging due to the high variability in human shape and motion and due to significant geometric and topological acquisition noise present in state-of-the-art acquisitions. The main objective of 3DMOVE is to automatically compute high-quality generative models from a database of raw dense 3D motion sequences for human bodies and faces. To achieve this objective, 3DMOVE will leverage recently developed deep learning techniques. The project also involves developing tools to assess the quality of the generated motions using perceptual studies. This project currently involves one Ph.D. student who was hired in November 2019.

9.2.2. Competitivity Clusters

9.2.2.1. 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.

9.3. International Research Visitors

The Morpheo team is hosting Professor Stephane Durocher during his sabbatical from July 2019 to June 2020. He is involved in the team research activities, in particular on the development of efficient algorithms to cluster a set of moving objects based on their trajectories, as obtained using the Kinovis platform. This will allow to perform motion analysis tasks, such as clustering objects into components that follow similar motions, which can help in analyzing the relative motion of body parts.

9.3.1. Visits to International Teams

9.3.1.1. Research Stays Abroad

1. Victoria Fernández Abrevaya did an internship with a British company in London, from July 2019 until September 2019.
2. Nitika Verma did an intership with Google at New York, from May 2019 until September 2019.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Selection

10.1.1.1. Chair of Conference Program Committees

- Stefanie Wuhrer was program co-chair of 3DV 2019, held in Québec City, Canada

10.1.1.2. Member of the Conference Program Committees

- Edmond Boyer was area chair for BMVC and ICCV 2019.

10.1.1.3. Reviewer

- Stefanie Wuhrer reviewed for CVPR, ICCV, and SIGGRAPH
- Jean-Sébastien Franco reviewed for CVPR, ICCV, and 3DV
- Edmond Boyer reviewed for CVPR.

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 (IJCV, Springer).

10.1.2.2. Reviewer - Reviewing Activities

- Edmond Boyer reviewed for IJCV and PAMI.
- Jean-Sébastien Franco reviewed for IJCV and Computer & Graphics.

10.1.3. Invited Talks

- Stefanie Wuhrer was invited to present at Dagstuhl seminar 19102
- Edmond Boyer gave invited talks at: ETH Zurich (February), Prague university (April), Naverlabs Grenoble (May), Lille University (June), Microsoft Zurich (October).

10.1.4. Scientific Expertise

- Jean-Sebastien Franco was a member of the Ensimag Engineering school - Grenoble INP steering commitee (Conseil d'École) in 2019.
- Jean-Sebastien Franco was a member of the recruiting commitee of Ensimag - Grenoble INP Engineering school for temporary research and teaching associates (ATER) in 2019.

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 : Sergi Pujades, Algorithmique et programmation fonctionnelle, 41H équivalent TD, L1, Université Grenoble Alpes, France.

Licence : Sergi Pujades, Modélisation des structures informatiques: aspects formels, 45H équivalent TD, L1, Université Grenoble Alpes, France.

Master : Sergi Pujades, Numerical Geometry, 15H équivalent TD, M1, Université Grenoble Alpes, France

Licence: Jean-Sébastien Franco, Introduction to Imperative Programming, 57h, Ensimag 1st year, Grenoble INP.

Licence: Jean-Sébastien Franco, C Programming project, 20h, 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, 45h, 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), 11h, Ensimag 3rd year, Grenoble INP.

Master: Edmond Boyer, 3D Modeling, 23h, M2R Mosig GVR, Grenoble INP.

Master: Edmond Boyer, Introduction to Visual Computing, 42h, M1 MoSig, Université Grenoble Alpes.

Master: Stefanie Wuhrer, 3D Graphics, 13.5h, M1 MoSig and MSIAM, Université Grenoble Alpes.

10.2.2. Supervision

PhD: Jinlong Yang, Learning shape space of dressed human in motion, Université Grenoble Alpes (France), March 2019, Franck Hétroy-Wheeler and Stefanie Wuhrer.

PhD: Vincent Leroy, 4D shape reconstruction from photoconsistency cues, Université Grenoble Alpes, October 2019, supervised by Edmond Boyer and Jean-Sébastien Franco.

PhD in progress : Sanae Dariouche, Anatomic Statistical Models of adipose, bone and muscle tissue, 01/11/2019, Sergi Pujades and Edmond Boyer

PhD in progress : Marilyn Keller, Ribs motion models for personalized 3D printed implants, 23/09/2019, Sergi Pujades and Michael Black

PhD in progress : Di Meng, , 23/09/2019, Edmond Boyer and Sergi Pujades

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: Boyao Zhou, Augmenting User Self-Representation in VR Environments, Université Grenoble Alpes (France), started 01/02/2019, supervised by Edmond Boyer and Jean-Sébastien Franco.

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: Pierre Zins, Learning to infer human motion, Université Grenoble Alpes (France), started 01/10/2019, supervised by Edmond Boyer, Tony Tung (Facebook) and Stefanie Wuhrer.

PhD in progress: Mathieu Marsot, Generative modeling of 3D human motion, Université Grenoble Alpes (France), started 01/11/2019, supervised by Jean-Sébastien Franco and Stefanie Wuhrer.

10.2.3. Juries

- Stefanie Wuhrer was jury member for the PhD defense of Alban Fichet, Université Grenoble Alpes, December 2019.
- Edmond Boyer was president of the PhD jury of Romain Rombourg, Université Grenoble Alpes, December 2019.
- Edmond Boyer was jury member for the PhD defense of Arthur Crenn, Université de Lyon, December 2019.

10.3. Popularization

10.3.1. Articles and contents

- Edmond Boyer was interviewed by the journal *Vivre à Grenoble* en février 2019. An article mentioning his work with the Kinovis platform was published in april 2019.

11. Bibliography

Major publications by the team in recent years

[1] A. Boukhayma, E. Boyer.*Surface Motion Capture Animation Synthesis*, in "IEEE Transactions on Visualization and Computer Graphics", June 2019, vol. 25, n^o 6, p. 2270-2283 [*DOI :* 10.1109/TVCG.2018.2831233], https://hal.inria.fr/hal-01781164

[2] V. Leroy, J.-S. Franco, E. Boyer.*Shape Reconstruction Using Volume Sweeping and Learned Photoconsistency*, in "European Conference on Computer Vision", Munich, Germany, Lecture Notes in Computer Science, Springer, September 2018, vol. 11213, p. 796-811 [*DOI :* 10.1007/978-3-030-01240-3_48], https://hal.archives-ouvertes.fr/hal-01849286

[3] J. Pansiot, E. Boyer.*CBCT of a Moving Sample from X-rays and Multiple Videos*, in "IEEE Transactions on Medical Imaging", February 2019, vol. 38, n^o 2, p. 383-393 [*DOI :* 10.1109/TMI.2018.2865228], https://hal.inria.fr/hal-01857487

[4] 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, IEEE, 2018, p. 2598-2606 [*DOI :* 10.1109/CVPR.2018.00275], https://hal.inria.fr/hal-01540389

Publications of the year

Doctoral Dissertations and Habilitation Theses

[5] V. Leroy.*Fast and Accurate 4D Modeling of Large Multi-Camera Sequences*, Université Grenoble Alpes (2016-2019) ; Inria Grenoble, October 2019, https://hal.archives-ouvertes.fr/tel-02435385

[6] J. YANG. *Learning shape spaces of dressed 3D human models in motion*, Université Grenoble Alpes, March 2019, https://tel.archives-ouvertes.fr/tel-02091727

Articles in International Peer-Reviewed Journal

[7] A. BOUKHAYMA, E. BOYER. *Surface Motion Capture Animation Synthesis*, in "IEEE Transactions on Visualization and Computer Graphics", June 2019, vol. 25, n^o 6, p. 2270-2283 [*DOI :* 10.1109/TVCG.2018.2831233], https://hal.inria.fr/hal-01781164

[8] N. HESSE, S. PUJADES, M. J. BLACK, M. ARENS, U. HOFMANN, S. SCHROEDER. *Learning and Tracking the 3D Body Shape of Freely Moving Infants from RGB-D sequences*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", June 2019, 12 [*DOI :* 10.1109/TPAMI.2019.2917908], https://hal.inria.fr/hal-02162171

[9] J. PANSIOT, E. BOYER. *CBCT of a Moving Sample from X-rays and Multiple Videos*, in "IEEE Transactions on Medical Imaging", February 2019, vol. 38, n^o 2, p. 383-393 [*DOI :* 10.1109/TMI.2018.2865228], https://hal.inria.fr/hal-01857487

[10] S. PUJADES, B. MOHLER, A. THALER, J. TESCH, N. MAHMOOD, N. HESSE, H. H. BÜLTHOFF, M. J. BLACK. *The Virtual Caliper: Rapid Creation of Metrically Accurate Avatars from 3D Measurements*, in "IEEE Transactions on Visualization and Computer Graphics", May 2019, vol. 25, n^o 5, p. 1887-1897 [*DOI :* 10.1109/TVCG.2019.2898748], https://hal.inria.fr/hal-02162159

International Conferences with Proceedings

[11] M. ARMANDO, J.-S. FRANCO, E. BOYER. *Adaptive Mesh Texture for Multi-View Appearance Modeling*, in "3DV 2019 - 7th International Conference on 3D Vision", Quebec City, Canada, IEEE, September 2019, p. 1-9, https://hal.inria.fr/hal-02284101

[12] J. BASSET, S. WUHRER, E. BOYER, F. MULTON. *Contact Preserving Shape Transfer For Rigging-Free Motion Retargeting*, in "MIG 2019 - ACM SIGGRAPH Conference Motion Interaction and Games", Newcastle Upon Tyne, United Kingdom, ACM, October 2019, p. 1-10 [*DOI :* 10.1145/3359566.3360075], https://hal.archives-ouvertes.fr/hal-02293308

[13] V. FERNÁNDEZ ABREVAYA, A. BOUKHAYMA, S. WUHRER, E. BOYER. *A Decoupled 3D Facial Shape Model by Adversarial Training*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, p. 1-10, https://hal.archives-ouvertes.fr/hal-02064711

[14] V. GABEUR, J.-S. FRANCO, X. MARTIN, C. SCHMID, G. ROGEZ. *Moulding Humans: Non-parametric 3D Human Shape Estimation from Single Images*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, p. 1-10, https://hal.inria.fr/hal-02242795

[15] *Best Paper*
R. KLOKOV, J. VERBEEK, E. BOYER. *Probabilistic Reconstruction Networks for 3D Shape Inference from a Single Image*, in "BMVC 2019 - British Machine Vision Conference", Cardiff, United Kingdom, September 2019, p. 1-15, https://arxiv.org/abs/1908.07475 - Awarded with Best Science Paper Honourable Mention Award at BMVC'19., https://hal.inria.fr/hal-02268466.

[16] 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 "APS 2019 - American Physical Society March Meeting", Boston, United States, March 2019, 1, https://hal.inria.fr/hal-01982257

Other Publications

[17] B. EGGER, W. A. P. SMITH, A. TEWARI, S. WUHRER, M. ZOLLHOEFER, T. BEELER, F. BERNARD, T. BOLKART, A. KORTYLEWSKI, S. ROMDHANI, C. THEOBALT, V. BLANZ, T. VETTER.*3D Morphable Face Models - Past, Present and Future*, September 2019, https://arxiv.org/abs/1909.01815 - working paper or preprint, https://hal.inria.fr/hal-02280281

Project-Team MOSAIC

MOrphogenesis Simulation and Analysis In siliCo

IN PARTNERSHIP WITH:
CNRS

Ecole normale supérieure de Lyon

INRA

Institut national de recherche pour l'agriculture, l'alimentation et l'environnement

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Computational Biology

Table of contents

Project-Team MOSAIC

Creation of the Project-Team: 2019 July 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

Christophe Godin [Team leader, Inria, Senior Researcher, HDR]
Olivier Ali [Inria, Researcher]
Romain Azaïs [Inria, Researcher]

External Collaborators

Arezki Boudaoud [Ecole Normale Supérieure Lyon]
Frederic Boudon [Centre de coopération internationale en recherche agronomique]
Emmanuel Faure [CNRS]
Patrick Lemaire [CNRS, HDR]
François Parcy [CNRS, HDR]

Jan Traas [INRA, HDR]
Samuel Teva Vernoux [CNRS, HDR]

Technical Staff
Guillaume Cerutti [INRA, Engineer]
Florian Gacon [Inria, Engineer]
Jonathan Legrand [CNRS, Engineer]

PhD Students
Florian Ingels [Inria, from Sep 2019]
Katia Mirande [Inria]
Hadrien Oliveri [Inria, until May 2019]
Manuel Petit [Inria, from Nov 2019]
Anne Schneider [INRA, until Nov 2019]

Post-Doctoral Fellows
Ayan Chaudhury [Inria, from Feb 2019]
Anuradha Kar [INRA, from Jun 2019]
Bruno Leggio [Ecole Normale Supérieure Lyon]

Visiting Scientists
Elva Carolina Chavez Hernandez
Farah Djelloul
Salah Eddine Habibeche
Gabriella Mosca

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.

[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.

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.

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?

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.

[0] Raff, R. A. (1996). The Shape of Life: Genes, Development, and the Evolution of Form. Univ. Chicago Press.

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

- MOSAIC has been promoted to Inria project-team in July 2019.
- In collaboration with CNRS (LIRMM and CRBM units in Montpellier), the team published a new web browser-based computational tool, Morphonet, to interactively explore complex 3D+time biological structures in silico, [8].

5. New Software and Platforms

5.1. cellcomplex

KEYWORDS: Polyhedral meshes - 3D

FUNCTIONAL DESCRIPTION: The cellcomplex library is a Python library that allows manipulating 2D or 3D multicellular complexes, with the study of plant tissues as a main application. It is mostly structured around a data structure that is used to represent such complexes as incidence graphs of dimension 2 or 3, and provides several key functionalities:

* The creation of structures from more basic representation (polygons of points for instance), from some geometrical primitives (2D or 3D) and the generation of synthetic regular or irregular grids, allowing notably the simulation of tissues.

* The computation of topological and geometrical properties on the multicellular complex structures, including notably useful computations on triangle meshes, a specific case of complexes with simplicial faces (areas, normals, triangle eccentricity, curvature estimator).

* The edition of structures by local topological operations, notably in th case of triangle meshes (edge flip, subdivision, vertex insertion) and multi-criteria geometrical optimization processes and isotropic remeshing.

* The import and export in various standard file formats for geometries (.obj, .ply, .msh) and notably in the standard formad defined by the community of plant tissue modelling (PLY, Saisnbury Computational Workshop 2015).

RELEASE FUNCTIONAL DESCRIPTION: * Major restructuration involving a change of namespace and a simplification of module architecture. * Inclusion of 3D visualization functionalities based on VTK.

- Participant: Guillaume Cerutti
- Contact: Guillaume Cerutti

5.2. draco_stem

DRACO-STEM : Dual Reconstruction by Adjacency Complex Optimization & SAM Tissue Enhanced Mesh

KEYWORDS: Meshing - Image segmentation - Computational biology - Optimization

FUNCTIONAL DESCRIPTION: Draco-stem provides a computational pipeline that allows going from multi-label segmented images of living tissue (typically resulting from a watershed segmentation of 3D microscopy image stacks) to topologically consistent, FEM-ready triangular meshes of all cell interfaces in the tissue.

It relies on an original topological optimization method that aims at reconstructing the simplicial complex of cellular adjacencies from the image, and on dualization and geometrical optimization to obtain a triangle mesh that satisfies simultaneously several quality criteria (triangle regularity, adequation to image, biological priors). The library provides implementations for 3D tissue reconstruction, single-layer 2.5D reconstruction and advanced 2D reconstruction.

RELEASE FUNCTIONAL DESCRIPTION: * Major refactoring, python3 compatibility * Addition of Draco2D functionalities

- Contact: Guillaume Cerutti
- Publication: DRACO-STEM: An Automatic Tool to Generate High-Quality 3D Meshes of Shoot Apical Meristem Tissue at Cell Resolution
- URL: https://gitlab.inria.fr/mosaic/draco_stem.git

5.3. 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 a plugin-based computational platform for the analysis and simulation of morphogenesis. It relies on a scalable software architecture based on the dtk kernel developed by the group of software engineers (SED) from the Sophia-Antipolis Inria Center. The development of Gnomon aims at answering four main challenges:

* Provide an easily accessible computational tool for the exploration of morphogenesis, by focusing on the deployability of the software (using conda), on the ergonomics of the user interface and the availability of the documentation.

* Give access to powerful tools for the exploration of dynamical forms, through an interactive visualization framework allowing the exploration in space in time and the access to algorithmic resources developed by the team for image sequences of multicellular tissues or collections of branching forms.

* Ensure the interoperability of computational libraries within the platform and its extensibility by a generalized plugin-based architecture (facilitated by the dtk framework) for algorithms, visualizations and data structures, enabling the members of the team and future users to feed the platform with their own C++ and Python libraries.

* Bridge the gap between experimental data and computational simulations by offering the possibility to go from one to the other in the same platform in a nearly transparent way, thanks to a common dynamical system framework integrated to the core of the platform.

Gnomon project organization: * Project leader: Christophe Godin * Software development coordinator: Guillaume Cerutti * DTK coordinators: Julien Wintz, Thibaud Kloczko * Plugin coordinators: Jonathan Legrand, Romain Azais, Olivier Ali, Frédéric Boudon. * Diffusion coordinator: 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 major technical update has been carried out on the Gnomon platform with the switch of the platform core to dtk2, implying an upgrade of all involved Python code to python 3.6+. The newer version 0.13 displays a clearer and more customizable interface and all functionalities from the version 0.9. In terms of functionalities, a major effort has been put on recovering former applications of the team (LPy and PlantGL) and including them in a robust way in the scope of the platform. A release of the version 1.0 is planned for early 2020 including validated algorithms for the quantitative analysis of 3D images of tissue, a much more robust user interface and an extensive user and developer documentation made available online.

- Participants: Olivier Ali, Frédéric Boudon, Guillaume Cerutti, Florian Gacon, Christophe Godin, Jonathan Legrand and Grégoire Malandain
- Contact: Christophe Godin

5.4. MorphoNet

Keywords: 3D web - Morphogenesis - Big data - 3D reconstruction

Functional Description: MorphoNet is an open-source web-based morphological browser. It consists of a web application, exploiting the Unity3D gaming engine, which offers the user a comprehensive palette of interactions with the data, in order to explore the structure, the dynamics and the variability of biological systems. Users can also project quantitative and genetic properties onto the morphological scaffold, allowing for instance to easily explore the correlation between shape dynamics and gene expression patterns. On top of that, datasets and associated information can be shared with other selected users or with entire groups. This possibility of directly sharing results within and between research communities, together with the use of a unified, human readable format, makes MorphoNet a unique tool for multidisciplinary research. Its web-based, user-friendly and open-source structure is also ideal for science dissemination and teaching.

- Partner: CRBM - Centre de Recherche en Biologie cellulaire de Montpellier
- Contact: Emmanuel Faure
- URL: http://www.morphonet.org

5.5. TimageTK

Tissue Image ToolKit

Keywords: 3D - Image segmentation - Fluorescence microscopy - Image registration - Image processing - Image filter

Functional Description: TimageTK (Tissue Image Toolkit) is a Python package dedicated to image processing of multicellular architectures such as plants or animals and is intended for biologists and modelers. It provides grayscale or labeled image filtering and mathematical morphology algorithms, as well as image registration and segmentation methods.

- Contact: Jonathan Legrand
- URL: https://mosaic.gitlabpages.inria.fr/timagetk/index.html

5.6. treex

Keywords: Graph algorithmics - Data structures - Combinatorics - Machine learning

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: 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 - Kernels for ordered or not, labeled or not, trees - Computation of coding processes (Harris path, Lukasiewicz walk and height process) - Visualization algorithms in Matplotlib or in LaTeX

RELEASE FUNCTIONAL DESCRIPTION: In 2019, treex has been published in JOSS (Journal of Open Source Software). The subtree kernel has been released to accompany an article submitted to Journal of Machine Learning Research. In addition, the DAG class and the kernel class have been extensively redesigned to be more user-friendly.

- Participants: Romain Azais, Guillaume Cerutti, Didier Gemmerle and Florian Ingels
- Contact: Romain Azais
- Publication: treex: a Python package for manipulating rooted trees
- URL: https://gitlab.inria.fr/azais/treex

6. New Results

6.1. Dynamical characterization of morphogenesis at cellular scale

Participants: Guillaume Cerutti, Emmanuel Faure [External Collaborator], Christophe Godin, Anuradha Kar, Bruno Leggio, Jonathan Legrand, Patrick Lemaire [External Collaborator], Grégoire Malandain [External Collaborator], Florent Papini, Manuel Petit, Jan Traas [External Collaborator].

- Related Research Axes: RA1 (Representation of biological organisms and their forms in silico) & RA3 (Plasticity & robustness of forms)
- Related 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:

Cell-scale atlases of development. One fundamental question linked to morphogenesis is at which level and timescale tissue or organ development is reproducible and stereotyped. To answer this question, variability must be quantitatively assessed. In the team we have created to this end 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.

Thanks to the invariant cellular lineage of early development of *P. mammillata* embryos and to 3D reconstruction of their development at cellular resolution, quantitative comparison of their properties from cell to tissue scale has been performed. After fluorescent membrane labelling, several embryos have been imaged for several hours by light-sheet microscopy. These images were then reconstructed through the segmentation pipeline ASTEC, which also automatically tracked each cell over several rounds of cell division. This large amount of data allowed us to create an atlas of geometrical and topological properties at cellular resolution, which gives unprecedented depth of information on the variability of ascidian development. In addition this atlas, coupled to previous knowledge on gene-expression dynamics from the ascidian genetic database (ANISEED), made it possible for us to develop a mathematical and computational model to explore the main drivers of early ascidian development, identified as area-of-contact-mediated cell-cell communications. This model was also validated by experimental manipulations and mutations induced in ascidian embryos. This work is currently under review [26].

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 in early phases, the development of most plant organs is under the influence of robust genetic patterns 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 Arabidopsis 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.

Robustness of ascidian embryonic development. The image segmentation pipeline ASTEC developed by the team in collaboration with the Inria Morpheme project-team in Sophia Antipolis and the CRBM team in Montpellier, allows the 3D reconstruction and tracking of each cell during early ascidian embryogenesis. This methods allowed us to reconstruct over 50 ascidian embryos, both wild-type and mutants. Exploiting this large database and the fixed cellular lineage of ascidian embryos, we extracted and compared geometrical and topological cellular properties. This allowed us to compare the intra-embryonic (left/right) to the inter-embryonic level of variability of several properties, including cell volume, cell-cell contacts and the structure of the tree seeded by each cell. This study demonstrated that the genetic-induced variability is comparable to the stochastic one, quantitatively showing that ascidian embryonic development is highly canalized, and that the high reproducibility of shapes observed during embryogenesis is rooted in the robustness of cellular geometry and topology. To look for the origin of this canalisation, we developed a mathematical model exploiting our quantitative geometric database and the previously-existing ascidian genetic database ANISEED. This model suggests that the main driver of ascidian development is the cell-cell communication mediated by direct physical contact, and hence dependent of the area-of-contact between neighbouring cells. This means that the robustness of cell topology and geometry is necessary for cell-cell biochemical interactions to give rise to the correct fate restriction events, which in turn we showed to be responsible for major changes of embryo geometry. We also tested and validated this feedback loop between cell contacts, fate restriction events and embryonic geometry predicted by the model by manipulations and mutations induced in ascidian embryos.These results are reported in a paper which is currently under review [26].

Robust extraction and characterization of cellular lineages. The quantification of temporal properties at cellular scale such as volumetric growth rate or strain patterns relies extensively on the identification of cellular lineages in time-lapse acquisitions of living tissues. In the case of plant tissues where the deformations between two consecutive time points can be very important in post-embryonic morphogenesis processes such as early flower development, it remains a real challenge to compute those lineages automatically, and manual user annotation is generally required to produce reliable results.

Building on the previous expertise of the team [25], [28] and on the state-of-the-art computational library for image analysis, timagetk, developed in collaboration with the Morpheme team, we currently develop a set of robust automatic cell lineaging methods for cases ranging from small to highly non-linear deformations. In the course of a M2 internship and the first months of a starting PhD work (Manuel Petit), a first so-called “naive” lineaging method has been implemented and validated on synthetic data with limited deformations. Methods involving optimal flow algorithms on graph structures and iterative image registration are being developed to provide robust results in the case of faster growing tissues. The output of these methods will allow to use the tools developed by the team for the analysis of spatio-temporal properties of growing cells at a much larger scale. This work is part of the Inria IPL Naviscope.

Reconstruction of Arabidopsis ovule development. 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. Given the technical difficulty of producing live-imaging acquisition sequences of ovules, we developed a method to perform a spatial registration of multiple individual ovules at various developmental states and in different global poses. Using the global cylindrical symmetry of the organ and the surface curvature as a key geometrical feature, we aligned individuals on their main axes and on their junction with the underlying placental tissue. Jointly with the 3D segmentation of cells in images, this will allow to evidence the invariant features of ovule development at cellular scale, and to study the robustness of the dynamics of the megaspore mother cell (MMC) across indivduals. This work was part of the Imago project.

6.2. Reconstruction of macroscopic forms from images and characterization of their variability

Participants: Ayan Chaudhury, Christophe Godin, Jonathan Legrand, Katia Mirande.

- Related Research Axes: RA1 (Representations of forms in silico) & RA3 (Plasticity & robustness of forms)
- Related 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. In collaboration with the ROMI partners from Sony CSL, Paris, 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. We aim at releasing both hardware schematics and the developed software for image reconstruction to be used as cheap open-source solution to phenotype plants. In addition, 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:

- RGB images were generated from the model and used as input of the pipeline;
- a physical version of the model has been obtained using 3D printing techniques;

In both cases, knowing the generated phenotypic traits or the model shape allow to test the pipeline ability to reconstruct the plant and quantify its traits of interest

The developed reconstruction and quantification pipeline is not made from scratch but aggregate a number of available third party libraries and codes in addition to three active research topics: spectral clustering, skeleton extraction, and ML segmentation. 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.

6.3. Analysis of tree data

Participants: Romain Azaïs, Christophe Godin, Salah Eddine Habibeche [External Collaborator], Florian Ingels.

- 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 2019, 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 into 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 [11], 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. In [3], we present two approximation algorithms that are optimal but assume that the approximation can be obtained by only adding vertices to the initial data (or by only deleting vertices from the initial data). In [11], 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. Another algorithm based on the efficient simulation of conditioned random walks on the space of trees is currently under development. This work should result in the submission of a paper next year.

It should be noted that the aforementioned strategy and algorithms can only be applied to topological trees. In 2019, we also began a new project on approximation of trees with geometrical attributes on their vertices and with possibly a controlled loss of information during the compression.

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 [5], 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 [5]. We have introduced new estimators and stated their consistency. Our techniques improve the existing results both theoretically and numerically.

We continue to explore these questions for subcritical but surviving Galton-Watson trees. The conditioning is a source of bias that must be taken into account to build efficient estimators of the birth distribution. This work should be submitted to a journal next year.

Kernel methods for tree data. Standard statistical techniques – such as SVMs for supervised learning – are usually designed to process Euclidean data. However, trees are typically non-Euclidean, thus preventing using these methods. Kernel methods allow this problem to be overcome by mapping trees in Hilbert spaces. However, the choice of kernel determines the feature space obtained, and thus greatly influences the performance of the different statistical algorithms. Our work is therefore focused on the question of how to build a good kernel.

We first looked in [17] at a kernel of the literature, the subtree kernel, and showed that the choice of the weight function – arbitrarily fixed so far – was crucial for prediction problems. By proposing a new framework to calculate this kernel, based on the DAG compression of trees, we were able to propose a new weight, learned from the data. In particular, on 8 data sets, we have empirically shown that this new weight improves prediction error in 7 cases, and with a relative improvement of more than 50% in 4 of these cases. This work was presented at a national conference [15].

We then tried to generalize our framework by proposing a kernel that is no longer based on subtrees, but on more general structures. To this end, we have developed an algorithm for the exhaustive enumeration of such structures, namely the forest of subtrees with a uniform fringe. This work will be submitted for pre-publication early in the coming year.

6.4. Mechanics of tissue morphogenesis

Participants: Olivier Ali, Arezki Boudaoud [External Collaborator], Guillaume Cerutti, Ibrahim Cheddadi [External Collaborator], Florian Gacon, Christophe Godin, Bruno Leggio, Jonathan Legrand, Hadrien Oliveri, Jan Traas [External Collaborator].

- Related Research Works: RW2 (*Data-driven models*) & RW3 (*Plasticity & robustness of forms*)
- Related 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. Integrating cellular behaviors such as mechano-sensitivity, intercellular fluxes of materials and cell division into a macroscopic mechanical picture of morphogenesis is the topic of this section.

Flattening mechanism during organogenesis in plants. Many plant species have thin leaf blades and axisymmetric elongating organs, such as stems and roots. From a morphoelastic perspective, such complex shapes are currently believed to emerge from the coordination between strain-based growth and stress-based stiffening at the cellular level.

To study the plausibility of such an hypothesis, we conducted numerical simulations where both a stress-based stiffening mechanism of cell walls [29] and a strain-based growth mechanism [24] have been implemented. We performed such simulations on multicellular and multilayered ellipsoidal structures and track their aspect ratio as they developed under various parametrization sets. One key aspect we wanted to investigate was the effect of an heterogeneous stress-based stiffening mechanism on the overall dynamics: Starting from a given initial shape, can we get significantly different shapes by assuming the stress-based stiffening mechanism active only in specific parts of the structures?

Our results, in accordance with experimental measurements conducted simultaneously by biologist colleagues, showed that: (i) Stress-based stiffening was mandatory to grow flat and axisymmetric organs; (ii) in order to grow flat structures, stress-based stiffening should only be active on anticlinal inner walls.

This work was part of Jan Traas's ERC grant *Morphodynamics*. This work is currently under review, see preprint version [23].

Influence of cell division during flat organogenesis in plants. One key limitation of our 3D modeling approach of leaf-like organogenesis is the lack of cell division implementation. This can be seen as a major flaw in the mechanical understanding of flattening since cell divisions, by increasing the number of load bearing walls, impact significantly the redistribution of mechanical stresses within the tissue.

To alleviate this limitation, we developed a 2D modeling approach to complement the 3D one. This 2D model encompasses the same biophysical processes as the 3D one (described in the previous subsection): a stress-based stiffening and a strain-based growth mechanisms of cell walls; augmented with a cell division module. We used this 2D framework to investigate the flattening dynamics of structures mimicking ellipsoid cross sections of growing organs. Such cross section were described as vertex-based, multicellular and multilayered structures.

We first reproduced the results obtained with the 3D approach to ensure that both models agreed on similar situations, where no cell division was implemented. We tested then several rules of cell division orientation and check which one(s) produced the most efficient flattening process. We were able to show that heterogeneity in the division rule between the epidermis and the inner tissues led to the more efficient flattening process and that a stress-based division rule was the most efficient to produce flat structure.

This analysis is part of the manuscript currently under review and available online in a preprint version [23].

Influence of mechanical stress anisotropy on the orientation of cell divisions in animal tissues. Tight regulation of cell division orientation is fundamental for tissue development. Recently, a great effort has been put into biophysical understanding of the *long-axis* division rules (Hertwig's rule for animal cells, Errera's rule for plant cells) and the systematic deviations from these rules observed *in vivo*. In both plants and animals, such deviations often correlate with anisotropic tensions within the tissue. To what extend these deviations are regulated or simply the result of stochasticity?

To address these questions in animal cells, we modeled theoretically and numerically cell division as an active process in a many-body system. We showed that under isotropic tension a cell's long axis emerges as the energetically optimal division orientation and that anisotropic stresses biased the energetics, leading to systematic deviations form Hertwig's rule. These deviations, as reported experimentally, are correlated to the main direction of stress anisotropy.

Our model successfully predicted division orientation distributions within two experimental systems: epidermis of the ascidian *Phallusia mammillata* (where deviations from Hertwig's rule have been so far eluding explanation) and of the pupal epithelium of the dorsal thorax of *D. melanogaster*.

This work was part of the *Digem* project and was presented in two international conferences: *Mechanobiology and Physics of Life* (Lyon) and *Developmental and Cell Biology of the Future* (Paris); and at the yearly *InriaBio* meeting in Lyon. A paper is currently under review and a preprint is available on bioRxiv [22].

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, between experiments and modeling, address the influence of turgor pressure heterogeneities on relative growth rate between cells. We showed that the coupling between fluxes and mechanics allows to predict observed morphological heterogeneities without any *ad hoc* assumption.

This work was part of the Agropolis fundation project *MecaFruit3D* and Arezki Boudaoud's ERC *PhyMorph*. It resulted in a publication in PLoS Computational Biology [7] that introduces the theoretical model and studies some of its properties. Another paper [27] presents the comparisons with experiments and is currently under review.

Development of *de novo* finite element (F.E.) library dedicated to mechanical simulations performed on complex cellularized structures. In order to compute accurately the mechanical stress field borne by multicellular pressurized 3D structures (such as plant tissues), we needed to update our existing library (*tissueMeca*, see [24]). Three key aspects had to be upgraded (i) the control over the F.E. solver, (ii) tracking of its precision and (iii) integration of the F.E. framework with the rest of our pipeline.

To that end, we decided to switch from *Sofa* to *FEniCS* (https://fenicsproject.org/) as the core F.E. framework used within our simulation pipeline. We started to develop a dedicated library, called *CellFem*, to solve F.E. problems on *PropertyTopomesh* instances (the data structure we developed within the team to describe multicellular plant tissues). *CellFem* provides a high level API to define and resolve variational problems to solve linear as well as non-linear elastic and elasto-plastic problems related to plant tissue morphogenesis.

In parallel, we also started the development of a meshing library (based on the GMSH library (http://gmsh.info/)) called *CellMesh* and dedicated to the triangulation of simplicial complexes. This work is currently under development.

6.5. Signaling and transport for tissue patterning

Participants: Romain Azaïs, Guillaume Cerutti, Christophe Godin, Bruno Leggio, Jonathan Legrand, Teva Vernoux [External Collaborator].

- Related Research Axes: RA1 (Representations of forms in silico) & RA2 (Data-driven models)
- Related Key Modeling 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 the spatial precision of the targeting. 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 govern the rhythmic patterning of organs, and the consequent emergence of phyllotaxis.

Using *Arabidopsis thaliana* as a model system, we develop 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 an original method to quantify in 3D the polarization of auxin transport for every anticlinal wall of the first layer of cells in confocal images. The developed method [13] was thoroughly evaluated against super-resolution acquisitions of the same tissue obtained using radial fluctuations (SRRF), and show to provide highly consistent results (less than 10% incorrect polarities, 80% of cells with a polarity vector error lesser than 30°). The digitally reconstructed networks evidenced an overall stable convergence of PIN1 polarities towards the center of the meristem, with a local convergence and divergence pattern that could explain the dynamics of auxin distributions in the meristem [19].

Landmark-based registration for the averaging of meristem patterning. To perform statistics of meristem patterning at the scale of a population, we developed a series of tools to compute a rigid 3D transformation that registers any individual meristem into a common cylindrical reference frame in which point-wise comparison is meaningful. The orignal method relies on the identification of biological landmarks (apex and main symmetry axis of the meristematic dome, position of the lastly emerged organ primordium and direction of the phyllotactic spiral) to compute this transform. These landmarks can be extracted from image acquisitions of meristems carrying the right fluorescent bio-markers (*CLV3* central zone marker for the apex, *DIIV* auxin bio-sensor for the organ primordia) using an original method that relies on the computation of 2D continuous maps of epidermal signal from discrete point clouds. The use of this registration method allowed to evidence key features of the transcriptional response of mersitematic cells to auxin [19].

In a second time, we aim to generalize the method to images without specific bio-markers, using only the geometry of the tissue to identify the relevant landmarks. To do so, machine learning approaches making use of the data processed for [19] are being developed and evaluated. This new landmark-based registration method would drastically improve the ability of comparing different individual meristems, open the way to spatial statistics over of multiple genetic and molecular signals, and contribute to an integrated tissue-level view of meristem patterning.

Computational models of integrated transport and signaling. Guided by new discoveries on auxin patterning dynamics in the shoot apical meristem (SAM) of *A. thaliana*, we developed a theoretical model of active and passive auxin transport. This model, built on existing view of auxin active transport [30], [31], naturally integrates the role of deeper cellular layers in the SAM and the mutual feedbacks between different components of the auxin-transport machinery. Through numerical simulation, the consequences of competing theories on PIN polarisation mechanism on auxin dynamics were explored. These results will serve, in quantitative comparisons with *in vivo* observation, to validate hypotheses on molecular mechanisms of auxin transport and to provide information on the role of memory effects and information fluxes during patterning.

These works were part of the *BioSensors* HFSP project and are carried out in the *Phyllo* ENS-Lyon project. These works gave rise to a journal article which is currently under review and have been partly presented at the *International Worskhop on Image Analysis Methods for the Plant Sciences* in Bron in July 2019.

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 activation/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 [6] and [10] for a 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.

6.7. Miscellaneous

Participants: Romain Azaïs, Christophe Godin, Bruno Leggio.

Measurements and nonlocal correlations in quantum mechanics. Based on a long standing collaboration between Christophe Godin and Przemyslaw Prusinkiewicz from the University of Calgary on the analysis of connections between computer simulation paradigms and quantum mechanics, we theoretically investigated with the quantum mechanics expertize of Bruno Leggio in the team effects of measurements on quantum systems, mostly in connection with quantum non-locality and entanglement. At the same time, we exploit formal and conceptual analogies between quantum theory and biologically-inspired structures to study the latter under new paradigms.

One fruitful line of research deals with the inherent non-locality of correlations between measurement outcomes, characterizing the quantum world. These phenomena are described by the celebrated Bell inequalities. We study ways to generalize such inequalities to better capture non-local correlations, at the same time shedding light on the origin of the discrepancy between quantum and classical stochasticity. In parallel, we develop and profit from formal analogies between the theory of non-locality and the exploration of fractal structures in the context of simulation of arborescent systems.

Another research line sees the application of parameter-estimation techniques for piecewise deterministic Markovian processes (PDMP), developed by members of the team, to the special case of quantum dynamics: under certain conditions, the evolution of an open quantum system can be described as a PDMP, with a specific and non-trivial structure marking its departure from classical behaviour. We show [21] that approaches to appraise parameter values of the evolving systems, developed in the context of classical dynamics, can be successfully applied to the specific case of quantum systems.

Finally, a third research topic consists of the study of the structure of typical quantum correlations, called entanglement, and its relation to thermal noise induced in a quantum system by its unavoidable interaction with its surrounding environment. We show [9] that the quantitative amount of noise represents a tight upper bound on the amount of bipartite quantum correlation two systems can establish between them.

Statistical analysis and stochastic modelling of penguin diving. The activity at sea of penguins can be reconstructed from measurement devices equipped on the animals during their trips. We study the relative behavior of the time under water with respect to the time spent at the surface from a dataset of about 100 thousands dives of little penguins. We show that dives that form a bout in which the penguin explores a patch of preys show a type of stationarity. We have built a mathematical model of sequences of dives that can be optimized in terms of number of preys catched by the animal under physiological constraints. This reproduces the stationary behavior observed in the data.

7. Partnerships and Cooperations

7.1. Regional Initiatives

7.1.1. 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.

7.1.2. IDEX Lyon Impulsion - MecaField (2019 - 2020)

Participants: Christophe Godin, Bruno Leggio, 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.

7.2. National Initiatives

7.2.1. 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

7.2.2. 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.

7.2.3. 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

7.2.4. 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

7.2.5. 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

7.2.6. MITI - MISGIVING (2019)

Participant: Romain Azaïs.

The diving performance of lung-breathing vertebrates, such as seabirds, can be quantified using measurement devices equipped on animals that allow us to reconstruct their activity at sea. During a classic dive, diving animals are faced with a dilemma: on the one hand, they want to optimize the time spent in contact with prey and therefore increase the time spent in diving; but, on the other hand, they are forced to return to the surface to breathe and will want to minimize this duration which remains however constrained by physiological rules. In addition, the dives are gathered in sequences because the prey are generally grouped in patches. In this project, we propose to use specific mathematical models to understand the complexity of the multi-scale decision processes that condition not only the optimal duration of the dive but also dives within a bout and therefore the total duration of the bout.

Partners:

- Centre d'Etudes Biologiques de Chizé
- Inria team CQFD in Bordeaux

7.3. European Initiatives

7.3.1. 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.

7.3.2. Collaborations with Major European Organizations

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)

7.4. International Research Visitors

7.4.1. Visits of International Scientists

- Farah Ben Naoum, associate professor in computer science at the University of Sidi Bel Abbes, Algeria, visited the team in March 2019 for 3 weeks and worked with Romain Azais and Christophe Godin on the definition of a strategy to make efficient random walks in spaces of trees.
- Gabriela Mosca was a visiting researcher from Celia Baroux's Lab (U. Zurich, Switzerland) in the context of the ANR project IMAGO. She spent 3 weeks in the team working with Guillaume Cerutti, Jonathan Legrand, Olivier Ali and Christophe Godin to set up a protocol to reconstruct ovule development from confocal imaging.

7.4.1.1. Internships

- Salah Eddine Habibeche is a PhD student supervised by Farah Ben Naoum from the University of Sidi Bel Abbes. The PhD subject of Salah consists of developing compressing schemes for semi-ordered trees. During his visit, he will study methods of compression of trees with loss of information.
- Caro Chavez Hernandez is a PhD student from Elena Alvarez-Buylla, UNAM University, Mexico. Caro visited the MOSAIC group to work with Christophe Godin to integrate the extensive gene regulatory network she assembled of key molecular processes involved at different phases of plant development into a model of plant architecture development written in LPy.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events: Organisation

8.1.1.1. Member of the Organizing Committees

- Romain Azaïs was a member of the organizing committee of the Journées de Statistique 2019 in Nancy.

- Romain Azaïs organized with Hélène Leman (NUMED) the second InriaBio@Lyon day, which gathered 6 teams of the Inria center interested in biological questions.

8.1.2. Scientific Events: Selection

8.1.2.1. Reviewer

- Romain Azaïs: NeurIPS 2019.

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

- Christophe Godin:
  - Asssociate editor in Plant biophysics and modeling of the journal Frontiers in Plant Sciences.
  - Member of the Editorial Advisory Board of the journal in silico Plants.
  - Member of the board of the functional-structural plant modeling (FSPM) series of conferences (next to come at fall 2020).

8.1.3.2. Reviewer - Reviewing Activities

- Olivier Ali: Frontiers in plant sciences.
- Romain Azaïs: Stochastic Processes and their Applications, Journal of Statistical Theory and Practice, Bernoulli.
- Christophe Godin: Frontiers in plant sciences, Annals of Botany, Bulletin of mathematical Biology (co-coordination of a special issue).
- Bruno Leggio: Physical Review Letters, Physical Review A, Physical Review B, Physical Review E, Frontiers in Plant Science, Entropy, Symmetry, Energies.

8.1.4. Invited Talks

- Oliver Ali was invited to give a seminar at the UPSC Lab in Umea (Sweden) in December 2019.
- Romain Azaïs was invited to give a presentation at the UTC in Compiègne, at the Laboratoire Paul Painlevé in Lille, and at the Laboratoire de Biométrie et Biologie Evolutive in Lyon.
- Christophe Godin was invited to give presentations at the research unit IGFL (Lyon), University of Calgary (Canada), University of Heidelberg (Germany) and ENS de Lyon.

8.1.5. Scientific Expertise

- Christophe 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).

8.1.6. Research Administration

- Christophe 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.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

- Olivier Ali:

  - Jury for the evaluation of a practical course on computational modeling for developmental biology (Licence 3 Biology ENS Lyon).
  - Course (3h) and practical course (3h) about mechanical modeling of plant morphogenesis (Master plant biology for sustainable production, Umea university and Swedish University of Agricultural Science, Sweden).
- Romain Azaïs:
  - Colles de mathématiques, CPGE PCSI, Lycée Jean Perrin, Lyon
  - Cours de Master 2 Apprentissage à partir de données arborescentes, Master Maths en Action, Université Lyon 1
- Christophe Godin:
  - Cours Les plantes dans tous leurs états pour non-specialistes, ENS de Lyon: *Phyllotaxis.* (2h).
  - Cours Master Sysbio, U. de Lyon: *A journey in Phyllotaxis.* (2h).
  - Master Stem Cells and Development ENS de Lyon (3h): *Introduction to Phyllotaxis*.
- Florian Ingels:
  - Séances de TP Probabilités-Statistiques, Licence Mathématiques et Economie, Université Lyon 1
  - Séances de TP Statistiques pour l'informatique, Licence Informatique, Université Lyon 1

8.2.2. Supervision

PhD (2016 – 2019): Florine Greciet (IECL, Université de Lorraine and Safran). Régression polynomiale par morceaux pour la propagation de fissures. Supervisors: Anne Gégout-Petit (Inria team BIGS, IECL, Université de Lorraine) and Romain Azaïs.

PhD in progress (2019 - 2022): Florian Ingels (MOSAIC, Université de Lyon). Supervisors: Romain Azaïs, Christophe Godin.

Hadrien Oliveri PhD (2015-2019) Montpellier University, co-supervision Christophe Godin, Jan Traas,

Katia Mirande PhD (2018-2021) Strasbourg University, co-supervision Franck Hetroy, Christophe Godin

Anne Schneider PhD (2016-2020) Angers University, co-supervision Jesssica Bertheloot, Christophe Godin

Manuel Petit PhD (2019-2022), ENS de Lyon, co-supervision Christophe Godin, Grégoire Malandain.

Master 2 (6 months): Florian Ingels. Statistical learning from trees: toward the definition of an optimal kernel for tree data. Supervisor: Romain Azaïs.

Master 2 (6 months): Florent Papini. Tracking cells in confocal images of meristems. Supervisor: Christophe Godin.

8.2.3. Juries

Christophe Godin was a member (as a supervisor) of the PhD defense of Hadrien Oliveri - University of Montpellier (May 2019).

Christophe Godin was the President of the PhD defense of Valeria Hernandez at ENS de Lyon (December 2019).

8.3. Popularization

- Olivier Ali:

  - Participated to the Declic initiative: Discussions about of scientific research to high school students (3h, Lycée Jean Perrin, Lyon).
  - Invited conference about mechanical modeling of plant morphogenesis for the Université Ouverte in Lyon.
- Christophe Godin participated to the Declic initiative: Introductory talk and discussions about scientific research to high school students (3h, Lycée Frédéric Fays, Villeurbanne).

8.3.1. Interventions

- Christophe Godin gave:
  - a lecture at Inria Grenoble in the context of a training program for Informatique et Sciences du Numérique dedicated to Math and Physics highschool teachers. see video (1h30)
  - a visio-seminar to highschool students of the French Lycée in Laos (1h30)

9. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] M. ALBER, C. GODIN, P. K. MAINI, R. M. MERKS, E. MJOLSNESS.*Introduction*, in "Bulletin of Mathematical Biology", July 2019 [*DOI :* 10.1007/S11538-019-00649-2], https://hal.archives-ouvertes.fr/hal-02424621

[2] O. ALI, H. OLIVERI, J. TRAAS, C. GODIN.*Simulating Turgor-Induced Stress Patterns in Multilayered Plant Tissues*, in "Bulletin of Mathematical Biology", May 2019, p. 1-23 [*DOI :* 10.1007/S11538-019-00622-Z], https://hal.inria.fr/hal-02154814

[3] R. AZAÏS.*Nearest embedded and embedding self-nested trees*, in "Algorithms", August 2019, vol. 12, n^{O} 9, p. 1-16, https://arxiv.org/abs/1709.02334 [*DOI :* 10.3390/A12090180], https://hal.archives-ouvertes.fr/hal-01584078

[4] R. AZAÏS, G. CERUTTI, D. GEMMERLÉ, F. INGELS.*treex: a Python package for manipulating rooted trees*, in "Journal of Open Source Software", June 2019, vol. 4, n^{O} 38, p. 1-2 [*DOI :* 10.21105/JOSS.01351], https://hal.archives-ouvertes.fr/hal-02164364

[5] R. AZAÏS, A. GENADOT, B. HENRY.*Inference for conditioned Galton-Watson trees from their Harris path*, in "ALEA : Latin American Journal of Probability and Mathematical Statistics", 2019, vol. 16, n^{O} 1, p. 1-45, https://arxiv.org/abs/1609.01057 [*DOI :* 10.30757/ALEA.V16-21], https://hal.archives-ouvertes.fr/hal-01360650

[6] J. BERTHELOOT, F. BARBIER, F. BOUDON, M. DOLORES PEREZ-GARCIA, T. PERON, S. CITERNE, E. DUN, C. BEVERIDGE, C. GODIN, S. SAKR.*Sugar availability suppresses the auxin-induced strigolactone pathway to promote bud outgrowth*, in "New Phytologist", September 2019, p. 1-14 [*DOI :* 10.1111/NPH.16201], https://hal-agrocampus-ouest.archives-ouvertes.fr/hal-02402977

[7] I. CHEDDADI, M. GÉNARD, N. BERTIN, C. GODIN.*Coupling water fluxes with cell wall mechanics in a multicellular model of plant development*, in "PLoS Computational Biology", June 2019, vol. 15, n^{O} 6, e1007121 [*DOI :* 10.1371/JOURNAL.PCBI.1007121], https://hal.archives-ouvertes.fr/hal-02196768

[8] B. LEGGIO, J. LAUSSU, A. CARLIER, C. GODIN, P. LEMAIRE, E. FAURE.*MorphoNet: an interactive online morphological browser to explore complex multi-scale data*, in "Nature Communications", June 2019, vol. 10, n° 2812, p. 1-8 [*DOI :* 10.1038/S41467-019-10668-1], https://hal.archives-ouvertes.fr/hal-01938153

[9] B. LEGGIO, A. NAPOLI, H. NAKAZATO, A. MESSINA.*Bounds on mixed state entanglement*, in "Entropy", 2020, vol. 22, n° 1, 62 [*DOI :* 10.3390/E22010062], https://hal.archives-ouvertes.fr/hal-02395355

[10] A. SCHNEIDER, C. GODIN, F. BOUDON, S. DEMOTES-MAINARD, S. SAKR, J. BERTHELOOT.*Light Regulation of Axillary Bud Outgrowth Along Plant Axes: An Overview of the Roles of Sugars and Hormones*, in "Frontiers in Plant Science", 2019, vol. 10, p. 1-17, This review was conducted in the framework of a PhD thesis funded by the Environment and Agronomy department of the French National Institute for Agricultural Research (INRA) [*DOI :* 10.3389/FPLS.2019.01296], https://hal.archives-ouvertes.fr/hal-02327576

International Conferences with Proceedings

[11] 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, SIAM, 2019, p. 39-53, https://arxiv.org/abs/1810.10860 [*DOI :* 10.1137/1.9781611975499.4], https://hal.archives-ouvertes.fr/hal-01294013

[12] B. HENRY, S. R. CHOWDHURY, A. LAHMADI, R. AZAÏS, J. FRANÇOIS, R. BOUTABA.*SPONGE: Software-Defined Traffic Engineering to Absorb Influx of Network Traffic*, in "44th IEEE Conference on Local Computer Networks (LCN)", Osnabrück, Germany, October 2019, https://hal.archives-ouvertes.fr/hal-02403616

Conferences without Proceedings

[13] G. CERUTTI, C. GALVAN-AMPUDIA, J. LEGRAND, C. GODIN, T. VERNOUX.*Quantifying cell polarities in confocal images using 3D wall meshes*, in "IAMPS 2019 - 7th International Workshop on Image Analysis Methods in the Plant Sciences", Bron, France, July 2019, p. 1-2, https://hal.archives-ouvertes.fr/hal-02176398

[14] F. GRECIET, R. AZAÏS, A. GÉGOUT-PETIT.*Régression polynomiale par morceaux sous contrainte de régularité pour la propagation de fissures*, in "JdS 2019 - 51èmes Journées de Statistique", Nancy, France, June 2019, p. 1-5, https://hal.archives-ouvertes.fr/hal-02172747

[15] F. INGELS, R. AZAÏS.*De l'importance de la fonction de poids dans le noyau des sous-arbres*, in "JdS 2019 - 51èmes Journées de Statistique", Nancy, France, June 2019, p. 1-6, https://hal.archives-ouvertes.fr/hal-02129866

Software

[16] C. PRADAL, C. FOURNIER, F. BOUDON, P. VALDURIEZ, E. PACITTI, Y. Y. GUÉDON, C. GODIN.*OpenAlea*, April 2019, OpenAlea 2.0 adds to OpenAlea 1.0 a high-level formalism dedicated to the modeling of morphogenesis that makes it possible to use several modeling paradigms (Blackboard, L-systems, Agents, Branching processes, Cellular Automata) expressed with different languages (Python, L-Py, R, Visual Programming, ...) to analyse and simulate shapes and their development, https://hal.inria.fr/hal-02100181

Other Publications

[17] R. AZAÏS, F. INGELS.*The Weight Function in the Subtree Kernel is Decisive*, November 2019, https://arxiv.org/abs/1904.05421v2 - 35 pages, https://hal.archives-ouvertes.fr/hal-02097593

[18] A. CHAUDHURY, C. GODIN.*Geometry Reconstruction of Plants*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02414486

[19] C. GALVAN-AMPUDIA, G. CERUTTI, J. LEGRAND, G. BRUNOUD, R. MARTIN AREVALILLO, R. AZAÏS, V. BAYLE, S. MOUSSU, C. WENZL, Y. JAILLAIS, J. U. LOHMANN, C. GODIN, T. VERNOUX.*Temporal integration of auxin information for the regulation of patterning*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02368529

[20] C. GODIN, C. GOLÉ, S. DOUADY.*Phyllotaxis: a remarkable example of developmental canalization in plants*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02370969

[21] B. LEGGIO, R. AZAÏS.*Estimation of Piecewise-Deterministic Trajectories in a Quantum Optics Scenario*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02389225

[22] B. LEGGIO, J. LAUSSU, E. FAURE, P. LEMAIRE, C. GODIN.*Multiscale mechanical model for cell division orientation in developing biological systems*, November 2019, working paper or preprint [*DOI :* 10.1101/785337], https://hal.archives-ouvertes.fr/hal-02367600

[23] F. ZHAO, F. DU, H. OLIVERI, L. ZHOU, O. ALI, W. CHEN, S. FENG, Q. WANG, S. LÜ, M. LONG, R. SCHNEIDER, A. SAMPATHKUMAR, C. GODIN, J. TRAAS, Y. JIAO.*A microtubule-mediated mechanical feedback controls leaf blade development in three dimensions*, November 2019, working paper or preprint [*DOI :* 10.1101/604710], https://hal.archives-ouvertes.fr/hal-02370615

References in notes

[24] 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

[25] R. FERNANDEZ, P. DAS, V. MIRABET, E. MOSCARDI, J. TRAAS, J.-L. VERDEIL, G. MALANDAIN, C. GODIN.*Imaging plant growth in 4D : robust tissue reconstruction and lineaging at cell resolution*, in "Nature Methods", 2010, vol. 7, p. 547–553

[26] 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

[27] 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

[28] G. MICHELIN, Y. REFAHI, R. WIGHTMAN, H. JONSSON, J. TRAAS, C. GODIN, G. MALANDAIN.*Spatio-temporal registration of 3D microscopy image sequences of arabidopsis floral meristems*, in "Proceedings - International Symposium on Biomedical Imaging", 2016, vol. 2016-June, p. 1127–1130 [*DOI :* 10.1109/ISBI.2016.7493464]

[29] 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

[30] 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

[31] 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

Team NANO-D

Algorithmes pour la Modélisation et la Simulation de Nanosystèmes

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 NANO-D

Creation of the Team: 2018 October 01

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

Sergei Grudinin [Team leader, CNRS, Researcher]
Leonard Jaillet [Inria, Starting Research Position, until Jan 2019]

PhD Students

Maria Kadukova [Univ Grenoble Alpes, PhD Student]
Guillaume Pages [Inria, PhD Student, until May 2019]

Post-Doctoral Fellow

Didier Devaurs [Univ Grenoble Alpes, Post-Doctoral Fellow]

Visiting Scientists

Karina Dos Santos Machado [FURG Brazil, until Oct 2019]
Vadim Strijov [MIPT Moscow, from Jun 2019 until Aug 2019]

Administrative Assistant

Maria Immaculada Presseguer [Inria, Administrative Assistant]

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* [20], *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 [62], for neighbor search between large rigid molecules [19], and for bond-order reactive force-fields [23]. 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 [21] and over 500,000 structures of small molecules stored in the Cambridge Structural Database [15]. 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 [41]. 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 [59]. 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 [61].

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 [32]. 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 [33], a wheelbarrow molecule [43], a nano-car [65], a Morse molecule [16], etc. Typically, these designs are optimized using semi-empirical quantum mechanics calculations, such as the semi-empirical ASED+ calculation technique [17].

While impressive, these are but two examples of the nanoscience revolution already impacting numerous fields, including electronics and semiconductors [50], textiles [48], [38], energy [52], food [27], drug delivery [37], [68], chemicals [39], materials [28], the automotive industry [14], aerospace and defense [34], medical devices and therapeutics [30], medical diagnostics [69], etc. According to some estimates, the world market for nanotechnology-related products and services will reach one trillion dollars by 2015 [60]. 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 [71], removing some degrees of freedom (e.g. keeping torsion angles only [46], [66]), coarse-graining [70], [63], [18], [64], multiple time step methods [57], [58], fast multipole methods [31], parallelization [45], averaging [26], multi-scale modeling [25], [22], reactive force fields [24], [74], interactive multiplayer games for predicting protein structures [29], 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 [72], [73].

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 [53], [54], [55], [56], the adaptive multiscale method [51], and the adaptive partitioning of the Lagrangian method [40]. 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).

4. Highlights of the Year

4.1. Highlights of the Year

- The work on first-principle simulation has been completed. The aim was to use the restrained dynamical model ARPS previously developed by the team to speed-up dynamical simulations using a first-principle interaction model. We have chosen Orbital-Free Density Functional Theory (OF-DFT), a fast scheme of DFT, as interaction model. We have developed a new OF-DFT code adapted to restrained particle simulations and have compared the accuracy and speed of our method to the state of the art OF-DFT code, PROFESS. The results were published in the Journal of Computational Chemistry [11] and the code is available in SAMSON. The thesis at the origin of this research has been defended in October.
- The proof-of-concept orientation-dependent potential for small molecules was developed and tested.
- With the advance of experimental procedures, obtaining sparse experimental data of proteins in solution (chemical crosslinking and small-angle scattering) is becoming a fast and routine practice. These can greatly enhance the accuracy of protein structure modeling. We participated in reviewing the current state of the art in modeling protein structures with the assistance of experimentally

determined chemical crosslinks and small-angle scattering profiles within the framework of the 13th meeting of Critical Assessment of Structure Prediction approaches [2], [4].

5. New Software and Platforms

5.1. 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 Sergei Grudinin
- Partner: MIPT Moscow
- Contact: Sergei 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/

5.2. 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 Sergei Grudinin
- Contact: Sergei 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/

5.3. 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.

RELEASE FUNCTIONAL DESCRIPTION: Version 0.7 from 3 Sep 2018: mmCIF support added. A file can be gzipped. Version 0.8 from 30 July 2019: Checks the similarity between chains. Explicitly output when an incomplete assembly is detected as symmetric. Version 0.9 from 8 Nov 2019: Added lists of excluded chains for both PDB and CIF formats.

- Participants: Guillaume Pages and Sergei Grudinin
- Contact: Sergei 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/

5.4. 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.

RELEASE FUNCTIONAL DESCRIPTION: Version 1.0 from 6th March 2018: Added computation of P(r). Version 1.1 from 19th April 2018: Excluded solvent radii updated. Many more ions added. Version 1.2 from 24th April 2018: Added automatic identification of labile and nonlabile explicit hydrogens. Added bulk SLD option. Version 1.3 from 30th July 2018: Added initial prototype of 2D dcattering images. Version 1.4 from 1st August 2018: Added docking module. Version 2.0 from January 2019: Added flexible optimization of conformations. A bug with absolute contrast fixed. Default behaiviour of Pepsi-SAXS changed, now it reads all the hydrogen atoms if the –hyd flag is set. Otherwise, it reads only explicit hydrogens on non-standard residues. Version 2.1 from February 2019: First implimentation of the multi-structure fitting.

- Participant: Sergei Grudinin
- Partner: MIPT Moscow
- Contact: Sergei 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/

5.5. 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.

RELEASE FUNCTIONAL DESCRIPTION: Version 1.2 from January 2019: Exclusion interaction list added. Version 1.3 from April 2019: Sequence and structural alignment fixed thanks to Guillaume Pages. More advanced iterative structural alignment added. Version 1.4 from June 2019: Firt version of domain identification added. Please use the "–nDomains" option Version 1.5 from July 2019: Rigid block format updated. JSON and Text output formats added. Version 1.6 from December 2019: Added initial Voronota support.

- Participants: Sergei Grudinin and Alexandre Hoffmann
- Contact: Sergei 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/

5.6. 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 Sergei Grudinin
- Contact: Sergei Grudinin
- Publication: Smooth orientation-dependent scoring function for coarse-grained protein quality assessment
- URL: https://team.inria.fr/nano-d/software/sbrod/

5.7. 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 Sergei Grudinin
- Contact: Sergei Grudinin
- Publication: Protein model quality assessment using 3D oriented convolutional neural networks
- URL: https://team.inria.fr/nano-d/software/ornate/

6. New Results

6.1. Reconstructing molecular shapes from SAXS data

We are working on a novel method to reconstruct the three-dimensional shape of a molecule from low-resolution experimental data. The structural data we are currently focusing on is obtained through small-angle X-ray scattering (SAXS) experiments, but we plan to also consider small-angle neutron scattering (SANS) data. Our *ab initio* reconstruction method is inspired by iterative phase-retrieval algorithms that can produce an image for an object when only the amplitudes of its Fourier transform are known and the phases are unknown. In our context, the X-ray scattering amplitudes associated with a molecule are the Fourier transform of its electron density. The novelty of our approach resides in the use of spherical harmonics expansions, which will allow performing the whole reconstruction process in Fourier space—contrary to existing methods that iterate between Fourier space and real space—for an improved computational efficiency. Our method is being implemented within the software PEPSI (polynomial expansions of protein structures and interactions).

6.2. Docking of cyclic molecules

In 2018 we participated in a docking Grand Challenge 4, which was organized by the Drug Design Data Resource (D3R) group. The goal was to predict correct poses and affinities of ligands binding the beta secretase 1 (BACE) receptor. Most of the ligands were macrocyclic. The challenge answers and results were fully released in February 2019. Upon that we started analyzing our protein-ligand docking strategy and wrote a corresponding paper [5].

6.3. Convex-PL and entropy

During the 2019 we continued developing our knowledge-based scoring function for protein-ligand interactions called Convex-PL [44]. We introduced a new descriptor characterizing side-chain entropy, tried two ways of computing the solvent-accessible surface area (SASA) descriptors with either using buried SASA, or SASA of those atoms that are contributing to the protein-ligand interaction according to the scoring function design (i.e which are within the cutoff distance of an interaction potential). Using these descriptors alongwith Convex-PL score and ligand side chain entropy descriptor, we trained a ridge regression model to predict binding affinities. This modification of Convex-PL was assessed on the CASF Benchmark 2016 released in the end of 2018, and on a subset of the DUD benchmark [42]. We are currently preparing the corresponding manuscript for submission.

6.4. Orientational potential for small molecules

With Pablo Chacon and Karina Dos Santos Machado we have developed a novel statistical protein-ligand scoring function called KORP-PL. KORP-PL is based on the coarse grained backbone-only representation of protein that is very common in protein structure modeling. It is based on ideas implemented by Pablo Chacon in the preceding scoring function for protein quality assessment called KORP [49], where each amino acid residue was characterized with a 3D oriented frame. We kept this representation for protein residues and computed the distances and angles between the oriented frames and points that describe ligand atoms. Using this data, we then derived the scoring function statistically in the same way, as it was done for protein-protein interactions in KORP. We also ran an optimization procedure to compute weights for each residue-atom interaction that would minimize the difference between predicted and experimental binding constants. We have assessed KORP-PL on several benchmarks [47], [67], one of which was manually derived from the user-submitted data of the D3R Challenges 2 [35], 3 [36], and 4. On all of them it performed exceptionally in pose prediction despite being a coarse-grained scoring function. We are currently preparing the corresponding manuscript for submission.

6.5. Analysis of a deep-learning architecture for fold recognition

Deep learning has recently demonstrated outstanding capabilities in classical pattern recognition problems. It has also obtained a tremendous success in the very recent protein structure prediction tasks. This work studies recurrent structural patterns in protein structures recognized by a deep neural network. We demonstrated that neural networks can automatically learn a vast amount of chemo-structural features with only a very little amount of human supervision. For example, our architecture correctly learns atomic, amino acid, and also higher-level molecular descriptors. The network architecture and the results are available at https://team.inria.fr/nano-d/software/Ornate/.

6.6. Controlled-advancement rigid-body optimization of nanosystems

In this study, we proposed a novel optimization algorithm, with application to the refinement of molecular complexes. Particularly, we considered optimization problem as the calculation of quasi-static trajectories of rigid bodies influenced by the inverse-inertia-weighted energy gradient and introduce the concept of advancement region that guarantees displacement of a molecule strictly within a relevant region of conformational space. The advancement region helps to avoid typical energy minimization pitfalls, thus, the algorithm is suitable to work with arbitrary energy functions and arbitrary types of molecular complexes without necessary tuning of its hyper-parameters. Our method, called controlled-advancement rigid-body optimization of nanosystems (Carbon), is particularly useful for the large-scale molecular refinement, as for example, the putative binding candidates obtained with protein–protein docking pipelines. Implementation of Carbon with user-friendly interface is available in the SAMSON platform for molecular modeling at https://www.samson-connect.net. The method was published in [10].

6.7. SAXS- and SANS-assisted modeling of proteins

We collaborated on data-assisted modeling of a KRAB-domain associated protein 1. Our work sheds light on its overall organization and combins solution scattering diffraction data, integrative modeling and single-molecule experiments [3].

We also participated in a combination of coarse-grained molecular dynamics simulations with previously measured small-angle scattering data to study the conformation of three-domain protein TIA-1 in solution. More precisely, we contributed with a specifically developed version of the Pepsi-SANS code. Our results suggest a general strategy for studying the conformation of multi-domain proteins in solution that combines coarse-grained simulations with small-angle X-ray scattering data that are generally most easy to obtain [13].

6.8. Predicting protein functional motions

Large macromolecules, including proteins and their complexes, very often adopt multiple conformations. Some of them can be seen experimentally, for example with X-ray crystallography or cryo-electron microscopy. This structural heterogeneity is not occasional and is frequently linked with specific biological function. Thus, the accurate description of macromolecular conformational transitions is crucial for understanding fundamental mechanisms of life's machinery.

We report on a real-time method to predict such transitions by extrapolating from instantaneous eigen-motions, computed using the normal mode analysis, to a series of twists. We demonstrate the applicability of our approach to the prediction of a wide range of motions, including large collective opening-closing transitions and conformational changes induced by partner binding. We also highlight particularly difficult cases of very small transitions between crystal and solution structures. Our method guaranties preservation of the protein structure during the transition and allows to access conformations that are unreachable with classical normal mode analysis. We provide practical solutions to describe localized motions with a few low-frequency modes and to relax some geometrical constraints along the predicted transitions. This work opens the way to the systematic description of protein motions, whatever their degree of collectivity. Our method is available as a part of the NOn-Linear rigid Block (NOLB) package at https://team.inria.fr/nano-d/software/nolb-normal-modes/ [12].

6.9. Protein structure prediction experiments

We participated in the CAPRI Round 46, the third joint CASP-CAPRI protein assembly prediction challenge. The Round comprised a total of 20 targets including 14 homo-oligomers and 6 heterocomplexes. Eight of the homo-oligomer targets and one heterodimer comprised proteins that could be readily modeled using templates from the Protein Data Bank, often available for the full assembly. The remaining 11 targets comprised 5 homodimers, 3 heterodimers, and two higher-order assemblies. These were more difficult to model, as their prediction mainly involved "ab-initio" docking of subunit models derived from distantly related templates [7].

7. Partnerships and Cooperations

7.1. Regional Initiatives

- An IDEX UGA grant is covering post-doc of Didier Devaurs, starting from December.
- Inria CORDI-S post-doctoral fellowship was obtained for Agnieszka Karczynska.

7.2. National Initiatives

7.2.1. ANR

In 2019, NANO-D had funding from one ANR program:

- **ANR PRCI**: covered the end of the PhD thesis of Guillaume Pages.

7.3. European Initiatives

7.3.1. Collaborations with Major European Organizations

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.

The Institute Laue-Langevin (ILL), the bioSANS team, Grenoble (France)

We are collaborating on the development of neutron small-angle scattering software

7.4. International Initiatives

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

7.4.1.1. FlexMol

Title: Algorithms for Multiscale Macromolecular Flexibility

International Partner (Institution - Laboratory - Researcher):

Rocasolano Institute of Physical Chemistry (IQFR-CSIC), Madrid, Spain (Spain) - Pablo Chacon

Start year: 2019

See also: https://team.inria.fr/nano-d/research/flexmol/

Molecular flexibility is essential to link structure and function of many biological macromolecules. Changes in protein conformation play a vital role in biochemical processes, from biopolymer synthesis to membrane transport. Many proteins can drastically alter their architecture and display considerable interdomain flexibility, as found in their 3D structures. For example, proteins rely on flexibility to respond to environmental changes, ligand binding and chemical modifications. Also, protein flexibility is tightly bound to their stability and is fundamental for drugs to exert biological effects.

Thus, one of the main challenges in the field of computational structural biology is to predict and explain molecular flexibility and corresponding conformational changes. For example, currently there are no methods that can reliably predict structural changes in proteins upon their binding. However, these are crucial to predict the structure of protein complexes with large conformational changes upon binding. To give another example, flexibility of the protein binding pocket is the major hurdle in reliable prediction of protein-ligand interactions for computer-aided drug design. Finally, intrinsic flexibility of macromolecules is nowadays the limiting factor for high-resolution experimental structure determination.

The partners of this associate team proposal comprise world-leading teams working with sound mathematical representations and techniques in the field of structural bioinformatics. These include spherical harmonics, normal modes analysis, high-order fast Fourier transforms, and more. The partners have very similar interests, but complimentary expertise. The goal of this collaboration is to mutually explore novel computational techniques for emerging problems in structural biology and bioinformatics related to molecular flexibility. This problem can be tackled at different scales. Large-scale flexibility of macromolecules can be efficiently described using collective coordinates. We will try to represent these in polynomial spaces, such that a practical flexible docking method can be based on this representation. Other applications include 3D shape reconstruction and scattering problems. Local molecular flexibility can be modelled using various techniques, including robotics-inspired methods, fragment libraries, etc. Here, our goal will be to rapidly sample the conformational space, and to construct a potential energy function applicable to flexible molecules. The ultimate goal of the project is to combine multiple levels of representation of molecular flexibility together. The project outcome will be built around innovative computer-aided drug-design algorithm with applications to prediction and computational design of important pharmaceutical targets such as antibodies.

7.4.2. Inria International Partners

7.4.2.1. 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.

7.4.2.2. 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 collaborated on the integrative structural biology approaches.
- Francis Crick Institute, London (UK), Biomolecular Modelling Laboratory. We collaborate on the development of flexible protein docking methods.
- University of Oslo. Ongoing collaboration on modeling protein systems guided by small-angle Xray and neutron small-angle scattering.
- University of Bergen, Norway. Ongoing collaboration on novel methods for normal mode analysis of protein structures.
- Nagoya University and RIKEN Center for Computational Science, Kobe, Japan. We collaborated on novel algorithms for scattering methods.
- University of Kansas, bioinformatics unit, USA. We have been collaborating on modeling protein-protein interactions.

7.4.3. 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.
- Vadim Strijov, professor at the department of Intelligent Systems, MIPT Moscow MIPT Moscow, July-August 2019.

7.5.1.1. Internships

- Khalid Mustafin (MIPT Moscow, Russia), Sep 2018 - Feb 2019.
- Ilia Igashov (MIPT Moscow, Russia), Nov 2018 - Apr 2020.

- Dmitrii Zhemchuzhnikov (UGA Grenoble), May 2019 - Sep 2019.

7.5.2. Visits to International Teams

- Sergei Grudinin visited the team of Ilia Vakser at Kansas University, Oct 15-31, 2019.

7.5.2.1. Explorer programme

- Sergei Grudinin visited Florence Tama and Osamu Miyashita, Nagoya University and RIKEN Center for Computational Science, Kobe, Japan. This was supported by the Exploration Japon 2019 program.
- Sergei Grudinin visited the team of Reidar Lund at University of Oslo, and the team of Nathalie Reuter at University of Bergen, Norway. Supported by the ÅSGARD 2019 program.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events: Organisation

8.1.1.1. Member of the Organizing Committees

- Sergei Grudinin coorganized CECAM workshop simSAS 2019, April 8-11, 90 participants, at ILL Grenoble, with Jean-Louis Barrat (UGA), Anne Martel (ILL), Sylvain Prévost (ILL), https://workshops.ill. fr/event/143/overview.

8.1.2. Scientific Events: Selection

8.1.2.1. Member of the Conference Program Committees

International Conference on Bioinformatics and Computational Biology (Devaurs)

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

- Biophysical Reviews (Devaurs)
- Current Proteomics (Devaurs)

8.1.3.2. Reviewer - Reviewing Activities

- Applied Sciences (Devaurs)
- Autonomous Robots (Devaurs)
- Bioinformatics (Oxford Press) (Devaurs, Grudinin)
- Current Proteomics (Devaurs)
- IEEE Transactions on Automation Science and Engineering (Devaurs)
- Journal of Bioinformatics and Computational Biology (Devaurs)
- Sensors (Devaurs)
- PLOS Computational Biology (Grudinin)
- Journal of Computational Chemistry (Grudinin)
- BMC Bioinformatics (Grudinin)
- Computational and Structural Biotechnology Journal (Elsevier) (Grudinin)
- Accounts of Chemical Research (ACS) (Grudinin)
- Computational Biology and Chemistry (Grudinin)
- Journal of Chemical Information and Modeling (Grudinin)
- Journal of Computer-Aided Molecular Design (Grudinin)

- Nature (Grudinin)
- Proteins: Structure, Function, and Bioinformatics (Grudinin)

8.1.4. Invited Talks

- D. Devaurs. Efficient strategies to explore the conformational space of proteins and molecular complexes. Seminar of the Functionality and Protein Engineering Unit, University of Nantes, France, 09/2019
- D. Devaurs. Efficient strategies to explore the conformational space of proteins and molecular complexes. CAPSID Seminar, Loria (Inria, CNRS, University of Lorraine), Nancy, France, 09/2019
- D. Devaurs. Efficient strategies to explore the conformational space of proteins and molecular complexes. ABS Seminar, Inria Sophia Antipolis - Méditerranée, France, 11/2019
- S. Grudinin. What does deep learning see in 3d protein structures? In 5th Korean Polish Conference on Protein Folding, September 16-18, 2019, Seoul, 2019.
- S. Grudinin. What does artificial intelligence see in 3d protein structures? In Workshop on Artificial Intelligence Applied to Photon and Neutron Science, 12-14 November, Grenoble, France, https://workshops.ill.fr/event/209/, 2019.
- S. Grudinin. Symmetry in protein complexes. In 7th CAPRI evaluation meeting, EMBL-EBI, Hinxton, UK, April 3-5, 2019, 2019.
- S. Grudinin. SAXS/SANS-assisted flexible fitting and docking. In EMBO Practical course on small angle neutron and X-ray scattering from biomolecules in solution, http://meetings.embo.org/event/19-small-angle-scattering, 2019.
- S. Grudinin. The least constraint approach for automatic coarse-graining of macromolecules. In The First International Conference on Mathematical Multiscale Modeling in Biology, Guanacaste, October 21-25, Costa Rica, 2019.
- S. Grudinin. Deep learning and artificial intelligence applied to the prediction of protein structure and interactions. In Aviesan ITMO Molecular and Structural Basis of Life Sciences : Deciphering The Functional Mechanisms Of Biological Macromolecules: Upcoming Challenges In Bioinformatics, Modelling And Experimental Validations, October 7-8, Paris, France, 2019.
- S. Grudinin. Challenges of modern structural bioinformatics. At the biological department, MSU Minsk, December 26, Minsk, Belarus, 2019.
- S. Grudinin. Novel software tools for small angle scattering. Seminar at the Synrotron SOLEIL, 13th May, 2019.
- S. Grudinin. On the analysis of macromolecular flexibility. Seminar at STFC Rutherford Appleton Laboratory, UK, March 2019.
- S. Grudinin. Novel algorithms for rapid modeling and analysis of flexibility and symmetry in macromolecules. Seminar at Cambridge Institute for Medical Research, UK, April 2019.
- S. Grudinin. Modeling and analysis of flexibility and symmetry in macromolecules. Seminar at Computational Biophysics department, Nagoya University, Japan, March 2019.
- S. Grudinin. On the nonlinear normal mode analysis. Seminar at the department of Chemistry, University of Bergen, Norway, August 2019.
- S. Grudinin. Novel computational tools for biomolecular small-angle scattering. Seminar at the department of Chemistry, University of Oslo, Norway, August 2019.
- S.Grudinin. Integrative Approaches for Current Problems in Structural Biology. Seminar at Laboratory of Theoretical Biochemistry, Institut de Biologie Physico-Chimique, Paris, 31 Jan 2019.
- S.Grudinin. New methods for protein-protein docking. Seminar at the bioinformatics unit, Kansas University, Lawrence, USA, October 2019.

8.1.5. Research Administration

- Sergei Grudinin is a member of IA working group at LINX, Lund Institute Of Advanced Neutron And X-Ray Science, Sweden.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

- Sergei Grudinin gave 2 public lecture and tutorials for Master and PhD level students on structural bioinformatics at the biology department of BSU Minsk, Belarus in December 2019.
- Sergei Grudinin prepared 2 tutorials for the EMBO school and CECAM workshop on small-angle scattering.

8.2.2. Supervision

PhD : Phd thesis defence of Guillaume Pagès, Université Grenoble Alpes, 12 septembre 2019
Title: Novel computational developments for protein structure analysis and prediction.
Thesis committee: Sergei Grudinin, Pablo Chacón, Česlovas Venclovas, Elodie Laine, Konrad Hinsen, Stéphane Redon, Arne Elofsson.
Summary: Proteins are ubiquitous for virtually all biological processes. Identifying their role helps to understand and potentially control these processes. However, even though protein sequence determination is now a routine procedure, it is often very difficult to use this information to extract relevant functional knowledge about system under study. Indeed, the function of a protein relies on a combination of its chemical and mechanical properties, which are defined by its structure. Thus, understanding, analysis and prediction of protein structure are the key challenges in molecular biology.

Prediction and analysis of individual protein folds is the central topic of this thesis. However, many proteins are organized in higher-level assemblies, which are symmetric in most of the cases, and also some proteins contain internal repetitions. In many cases, designing a fold with repetitions or designing a symmetric protein assembly is the simplest way for evolution to achieve a specific function. This is because the number of combinatorial possibilities in the interactions of designed folds reduces exponentially in the symmetric cases. This motivated us to develop specific methods for symmetric protein assemblies and also for individual proteins with internal repeats. Another motivation behind this thesis was to explore and advance the emerging deep neural network field in application to atomistic 3-dimensional (3D) data.

This thesis can be logically split into two parts. In the first part, we propose algorithms to analyse structures of protein assemblies, and more specifically putative structural symmetries. We start with a definition of a symmetry measure based on 3D Euclidean distance, and describe an algorithm to efficiently compute this measure, and to determine the axes of symmetry of protein assemblies. This algorithm is able to deal with all point groups, which include cyclic, dihedral, tetrahedral, octahedral and icosahedral symmetries, thanks to a robust heuristic that perceives correspondence between asymmetric subunits. We then extend the boundaries of the problem, and propose a method applicable to the atomistic structures without atom correspondence, internal symmetries, and repetitions in raw density maps. We tackle this problem using a deep neural network (DNN), and we propose a method that predicts the symmetry order and a 3D symmetry axis.

Then, we extend the DNN architecture to recognise folding quality of 3D protein models. We trained the DNN using as input the local geometry around each residue in a protein model represented as a density map, and we predicted the CAD-scores of these residues. The DNN was specifically conceived to be invariant with respect to the orientation of the input model. We also designed some parts of the network to automatically recognise atom properties and robustly select features. Finally, we provide an analysis of the features learned by the DNN. We show that our architecture correctly learns atomic, amino acid, and also higher-level molecular descriptors. Some of them are rather complex, but well understood from the biophysical point of view. These include atom partial charges, atom chemical elements, properties of amino acids, protein secondary structure and atom solvent exposure. We also demonstrate that our network learns novel structural features.

This study introduces novel tools for structural biology. Some of them are already used in the community, for example, by the PDBe database and CASP assessors. It also demonstrates the power of deep learning in the representation of protein structure and shows applicability of DNNs to computational tasks that involve 3D data.

PhD : Phd thesis defence of François Rousse, Université Grenoble Alpes, 2019
Title: Incremental Algorithm for Orbital-Free Density Functional Theory.
Thesis committee: Stéphane Redon, Jean Clérouin, Reinhold Schneider, Johannes Dieterich, Philippe Blaise, Florent Calvo, Xavier Bouju.
Summary: The ability to model molecular systems on a computer has become a crucial tool for chemists. Indeed molecular simulations have helped to understand and predict properties of nanoscopic world, and during the last decades have had large impact on domains like biology, electronic or materials development. Particle simulation is a classical method of molecular dynamic. In particle simulation, molecules are split into atoms, their inter-atomic interactions are computed, and their time trajectories are derived step by step. Unfortunately, inter-atomic interactions computation costs prevent large systems to be modeled in a reasonable time. In this context, our research team looks for new accurate and efficient molecular simulation models. One of our team's focus is the search and elimination of useless calculus in dynamical simulations. Hence has been proposed a new adaptively restrained dynamical model in which the slowest particles movement is frozen, computational time is saved if the interaction calculus method do not compute again interactions between static atoms. The team also developed several interaction models that benefit from a restrained dynamical model, they often updates interactions incrementally using the previous time step results and the knowledge of which particle have moved.

In the wake of our team's work, we propose in this thesis an incremental First-principles interaction models. Precisely, we have developed an incremental Orbital-Free Density Functional Theory method that benefits from an adaptively restrained dynamical model. The new OF-DFT model keeps computation in Real-Space, so can adaptively focus computations where they are necessary. The method is first proof-tested, then we show its ability to speed up computations when a majority of particle are static and with a restrained particle dynamic model. This work is a first step toward a combination of incremental First-principle interaction models and adaptively restrained particle dynamic models.

PhD in progress : Maria Kadukova, "Novel computational approaches for protein ligand interactions", Sep 2016-, supervisors: Sergei Grudinin (France) and Vladimir Chupin (MIPT, Russia).

8.2.3. Juries

Sergei Grudinin served as an opponent at the defence of David Menéndez Hurtado's PhD thesis entitled 'Structured Learning for Structural Bioinformatics'. The defence took place at the Department of Biochemistry and Biophysics, Stockholm University, Sweden on the 11th of October.

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] F. ROUSSE.*Incremental algorithms for Orbital-Free Density Functional Theory*, Université Grenoble - Alpes, October 2019, https://hal.archives-ouvertes.fr/tel-02435781

Articles in International Peer-Reviewed Journal

[2] J. E. FAJARDO, R. SHRESTHA, N. GIL, A. BELSOM, S. CRIVELLI, C. CZAPLEWSKI, K. FIDELIS, S. GRUDININ, M. KARASIKOV, A. KARCZYŃSKA, A. KRYSHTAFOVYCH, A. LEITNER, A. LIWO, E. LUBECKA, B. MONASTYRSKYY, G. PAGÈS, J. RAPPSILBER, A. SIERADZAN, C. SIKORSKA, E. TRABJERG, A. FISER.*Assessment of chemical-crosslink-assisted protein structure modeling in CASP13*, in "Proteins - Structure, Function and Bioinformatics", 2019, forthcoming [*DOI :* 10.1002/PROT.25816], https://hal.inria.fr/hal-02315542

[3] G. FONTI, M. MARCAIDA, L. BRYAN, S. TRÄGER, A. KALANTZI, P.-Y. J. HELLEBOID, D. DEMURTAS, M. TULLY, S. GRUDININ, D. TRONO, B. FIERZ, M. DAL PERARO.*KAP1 is an antiparallel dimer with a functional asymmetry*, in "Life Science Alliance", August 2019, vol. 2, n^o 4, e201900349 [*DOI :* 10.26508/LSA.201900349], https://hal.archives-ouvertes.fr/hal-02291553

[4] G. HURA, C. HODGE, D. ROSENBERG, D. GUZENKO, J. DUARTE, B. MONASTYRSKYY, S. GRUDININ, A. KRYSHTAFOVYCH, J. TAINER, K. FIDELIS, S. TSUTAKAWA.*Small angle X-ray scattering-assisted protein structure prediction in CASP13 and emergence of solution structure differences*, in "Proteins - Structure, Function and Bioinformatics", 2019, p. 1-17, forthcoming [*DOI :* 10.1002/PROT.25827], https://hal.inria.fr/hal-02315292

[5] M. KADUKOVA, V. CHUPIN, S. GRUDININ.*Docking rigid macrocycles using Convex-PL, AutoDock Vina, and RDKit in the D3R Grand Challenge 4*, in "Journal of Computer-Aided Molecular Design", November 2019, p. 1-10 [*DOI :* 10.1007/S10822-019-00263-3], https://hal.archives-ouvertes.fr/hal-02434514

[6] M. KARASIKOV, G. PAGÈS, S. GRUDININ.*Smooth orientation-dependent scoring function for coarse-grained protein quality assessment*, in "Bioinformatics", August 2019, vol. 35, n^o 16, p. 2801–2808 [*DOI :* 10.1093/BIOINFORMATICS/BTY1037], https://hal.inria.fr/hal-01971128

[7] M. LENSINK, G. BRYSBAERT, N. NADZIRIN, S. VELANKAR, R. A. CHALEIL, T. GERGURI, P. BATES, E. LAINE, A. CARBONE, S. GRUDININ, R. KONG, R. LIU, X. XU, H. SHI, S. CHANG, M. EISENSTEIN, A. KARCZYŃSKA, C. CZAPLEWSKI, E. LUBECKA, A. LIPSKA, P. KRUPA, M. MOZOLEWSKA, Ł. GOLON, S. SAMSONOV, A. LIWO, S. CRIVELLI, G. PAGÈS, M. KARASIKOV, M. KADUKOVA, Y. YAN, S. HUANG, M. ROSELL, L. A. RODRÍGUEZ-LUMBRERAS, M. ROMERO-DURANA, L. DÍAZ-BUENO, J. FERNANDEZ-RECIO, C. CHRISTOFFER, G. TERASHI, W. SHIN, T. ADERINWALE, S. RAGHAVENDRA MADDHURI VENKATA SUBRAM, D. KIHARA, D. KOZAKOV, S. VAJDA, K. PORTER, D. PADHORNY, I. DESTA, D. BEGLOV, M. IGNATOV, S. KOTELNIKOV, I. MOAL, D. RITCHIE, I. CHAUVOT DE BEAUCHÊNE, B. MAIGRET, M. E. R. ECHARTEA, D. BARRADAS-BAUTISTA, Z. CAO, L. CAVALLO, R. OLIVA, Y. CAO, Y. SHEN, M. BAEK, T. PARK, H. WOO, C. SEOK, M. BRAITBARD, L. BITTON, D. SCHEIDMAN-DUHOVNY, J. DAPKŪNAS, K. OLECHNOVIČ, Č. VENCLOVAS, P. J. KUNDROTAS, S. BELKIN, D. CHAKRAVARTY, V. BADAL, I. A. VAKSER, T. VREVEN, S. VANGAVETI, T. M. BORRMAN, Z. WENG, J. D. GUEST, R. GOWTHAMAN, B. G. PIERCE, X. XU, R. DUAN, L. QIU, J. HOU, B. RYAN MERIDETH, Z. MA, J. CHENG, X. ZOU, P. KOUKOS, J. ROEL-TOURIS, F. AMBROSETTI, C. GENG, J. SCHAARSCHMIDT, M. TRELLET, A. S. MELQUIOND, L. XUE, B. JIMÉNEZ-GARCÍA, C. NOORT, R. HONORATO, A. M. BONVIN, S. J. WODAK.*Blind prediction of homo- and hetero- protein complexes: The CASP13-CAPRI experiment*, in "Proteins - Structure, Function and Bioinformatics", October 2019, vol. 87, n^o 12, p. 1200-1221 [*DOI :* 10.1002/PROT.25838], https://hal.inria.fr/hal-02320974

[8] G. PAGÈS, B. CHARMETTANT, S. GRUDININ.*Protein model quality assessment using 3D oriented convolutional neural networks*, in "Bioinformatics", September 2019, vol. 35, n^o 18, p. 3313–3319 [*DOI :* 10.1093/BIOINFORMATICS/BTZ122], https://hal.inria.fr/hal-01899468

[9] G. PAGÈS, S. GRUDININ.*DeepSymmetry : Using 3D convolutional networks for identification of tandem repeats and internal symmetries in protein structures*, in "Bioinformatics", June 2019, p. 1-24, https://arxiv.org/abs/1810.12026 , forthcoming [*DOI :* 10.1093/BIOINFORMATICS/BTZ454], https://hal.inria.fr/hal-01903624

[10] P. POPOV, S. GRUDININ, A. KURDIUK, P. BUSLAEV, S. REDON.*Controlled-advancement rigid-body optimization of nanosystems*, in "Journal of Computational Chemistry", October 2019, vol. 40, n^{o} 27, p. 2391-2399 [*DOI :* 10.1002/JCC.26016], https://hal.inria.fr/hal-02315276

[11] F. ROUSSE, S. REDON.*Incremental solver for Orbital-Free Density Functional Theory*, in "Journal of Computational Chemistry", September 2019, vol. 40, n^{o} 23, p. 2013-2027 [*DOI :* 10.1002/JCC.25854], https://hal.inria.fr/hal-02135603

Other Publications

[12] S. GRUDININ, E. LAINE, A. HOFFMANN.*Predicting protein functional motions: an old recipe with a new twist*, September 2019, working paper or preprint [*DOI :* 10.1101/703652], https://hal.archives-ouvertes.fr/hal-02291552

[13] A. H. LARSEN, Y. WANG, A. BOTTARO, S. GRUDININ, L. ARLETH, K. LINDORFF-LARSEN.*Combining molecular dynamics simulations with small-angle X-ray and neutron scattering data to study multi-domain proteins in solution*, January 2020, working paper or preprint [*DOI :* 10.1101/2019.12.26.888834], https://hal.archives-ouvertes.fr/hal-02434585

References in notes

[14] 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, http://www.sciencedirect.com/science/article/pii/S0300944007001464

[15] 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

[16] 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

[17] 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, http://www.sciencedirect.com/science/article/pii/S003960280600700X

[18] 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

[19] 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

[20] 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

[21] 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

[22] 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

[23] 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, http://www.sciencedirect.com/science/article/pii/S0021999111007042

[24] 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

[25] 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, http://www.sciencedirect.com/science/article/pii/S1089315601000265

[26] 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

[27] 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

[28] 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

[29] 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

[30] 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

[31] E. DARVE.*The Fast Multipole Method: Numerical Implementation*, in "Journal of Computational Physics", 2000

[32] 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

[33] 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

[34] 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, http://www.sciencedirect.com/science/article/pii/S0257897210003178

[35] Z. GAIEB, S. LIU, S. GATHIAKA, M. CHIU, H. YANG, C. SHAO, V. A. FEHER, W. P. WALTERS, B. KUHN, M. G. RUDOLPH, S. K. BURLEY, M. K. GILSON, R. E. AMARO.*D3R Grand Challenge 2: blind prediction of protein–ligand poses, affinity rankings, and relative binding free energies*, in "Journal of Computer-Aided Molecular Design", 2018, vol. 32, n^{o} 1, p. 1–20, https://doi.org/10.1007/s10822-017-0088-4

[36] Z. GAIEB, C. D. PARKS, M. CHIU, H. YANG, C. SHAO, W. P. WALTERS, M. H. LAMBERT, N. NEVINS, S. D. BEMBENEK, M. K. AMERIKS, T. MIRZADEGAN, S. K. BURLEY, R. E. AMARO, M. K. GILSON.*D3R Grand Challenge 3: blind prediction of protein–ligand poses and affinity rankings*, in "Journal of Computer-Aided Molecular Design", 2019, vol. 33, n^{o} 1, p. 1–20, https://doi.org/10.1007/s10822-018-0180-4

[37] 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

[38] 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

[39] 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

[40] 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

[41] 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

[42] A. JAHN, G. HINSELMANN, N. FECHNER, A. ZELL.*Optimal assignment methods for ligand-based virtual screening*, in "Journal of Cheminformatics", 2009, vol. 1, n^{o} 1, p. 1–23, https://doi.org/10.1186/1758-2946-1-14

[43] 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

[44] 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", 2017, vol. 31, n^{o} 10, https://doi.org/10.1007/s10822-017-0068-8

[45] 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, http://www.sciencedirect.com/science/article/pii/S0021999199962010

[46] 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

[47] Y. LI, M. SU, Z. LIU, J. LI, J. LIU, L. HAN, R. WANG.*Assessing protein–ligand interaction scoring functions with the CASF-2013 benchmark*, in "Nature protocols", 2018, vol. 13, n^{o} 4, 666

[48] 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

[49] J. R. LOPEZ-BLANCO, P. CHACON.*KORP: Knowledge-based 6D potential for fast protein and loop modeling*, in "Bioinformatics", 2019, vol. 35, n^{o} 17, p. 3013–3019, https://doi.org/10.1093/bioinformatics/btz026

[50] W. LU, C. M. LIEBER.*Nanoelectronics from the bottom up*, in "Nature materials", 2007, vol. 6, n^{o} 11, p. 841-850

[51] 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

[52] 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

[53] 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

[54] 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

[55] 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

[56] 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

[57] 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° 24, p. 10464-10468, http://pubs.acs.org/doi/abs/10.1021/jp960295w

[58] 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

[59] D. W. RITCHIE, G. J. KEMP.*Protein docking using spherical polar Fourier correlations*, in "Proteins: Structure, Function, and Bioinformatics", 2000, vol. 39, n° 2, p. 178–194

[60] M. C. ROCO.*The long view of nanotechnology development: the National Nanotechnology Initiative at 10 years*, in "Journal of Nanoparticle Research", 2010

[61] B. ROOKS.*A shorter product development time with digital mock-up*, in "Assembly Automation", 1998, vol. 18, n° 1, p. 34-38 [*DOI :* DOI:10.1108/01445159810201405], http://www.ingentaconnect.com/content/mcb/033/1998/00000018/00000001/art00004

[62] 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° 13 [*DOI :* 10.1093/BIOINFORMATICS/BTM191], http://bioinformatics.oxfordjournals.org/content/23/13/i408.abstract

[63] R. E. RUDD.*Coarse-Grained Molecular Dynamics for Computer Modeling of Nanomechanical Systems*, in "International Journal for Numerical Methods in Engineering", 2004

[64] 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

[65] 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

[66] 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, http://www.sciencedirect.com/science/article/pii/S1090780796910277

[67] M. SU, Q. YANG, Y. DU, G. FENG, Z. LIU, Y. LI, R. WANG.*Comparative assessment of scoring functions: the casf-2016 update*, in "Journal of chemical information and modeling", 2018, vol. 59, n° 2, p. 895–913

[68] 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

[69] 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

[70] N. VAIDEHI, W. A. GODDARD.*Domain Motions in Phosphoglycerate Kinase using Hierarchical NEIMO Molecular Dynamics Simulations*, in "The Journal of Physical Chemistry A", 2000, vol. 104, n^{o} 11, p. 2375-2383, http://pubs.acs.org/doi/abs/10.1021/jp991985d

[71] 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^{o} 2, 025013, http://stacks.iop.org/1478-3975/6/i=2/a=025013

[72] 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

[73] 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^{o} 0, p. 461–479, http://www.sciencedirect.com/science/article/pii/016612809285024F

[74] A. C. T. VAN DUIN, S. DASGUPTA, F. LORANT, W. A. GODDARD.*ReaxFF: A Reactive Force Field for Hydrocarbons*, in "The Journal of Physical Chemistry A", 2001, vol. 105, n^{o} 41, p. 9396-9409, http://pubs.acs.org/doi/abs/10.1021/jp004368u

Team NECS

Networked Controlled Systems

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 NECS

Creation of the Team: 2019 January 01

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]
Federica Garin [Inria, Researcher]
Paolo Frasca [CNRS, Researcher]

Faculty Members

Hassen Fourati [Univ. Grenoble Alpes, Associate Professor]
Alain Kibangou [Univ. Grenoble Alpes, Associate Professor, HDR]

External Collaborator

Stéphane Mollier [CNRS, from Oct 2019]

Technical Staff

Vadim Bertrand [CNRS]

PhD Students

Maria Castaldo [CNRS, PhD Student, from Nov 2019]
Nicolas Martin [CNRS, PhD Student, until Nov 2019]
Stéphane Mollier [CNRS, PhD Student, until Sep 2019]
Muhammad Umar B Niazi [CNRS, PhD Student]
Denis Nikitin [CNRS, PhD Student]
Bassel Othman [IFPEN, PhD Student]
Ujjwal Pratap [CNRS, PhD Student, from Feb 2019]

Martin Rodriguez [CNRS, PhD Student]
Liudmila Tumash [CNRS, PhD Student]
Renato Vizuete Haro [École supérieure d'électricité, PhD Student, from Sep 2019]
Makia Zmitri [CNRS, PhD Student]

Post-Doctoral Fellows
Thibault Liard [Inria, Post-Doctoral Fellow, until Mar 2019]
Fadoua Taia Alaoui [Univ. Grenoble Alpes, Post-Doctoral Fellow, from Oct 2019]

Visiting Scientists
Diego Deplano [Univ. di Cagliari (Italy), until Mar 2019]
Natalia Lopez [Apr 2019]
Raphael Stern [Vanderbilt Univ. Nashville (USA), Mar 2019]

Administrative Assistant
Myriam Etienne [Inria, Administrative Assistant]

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).

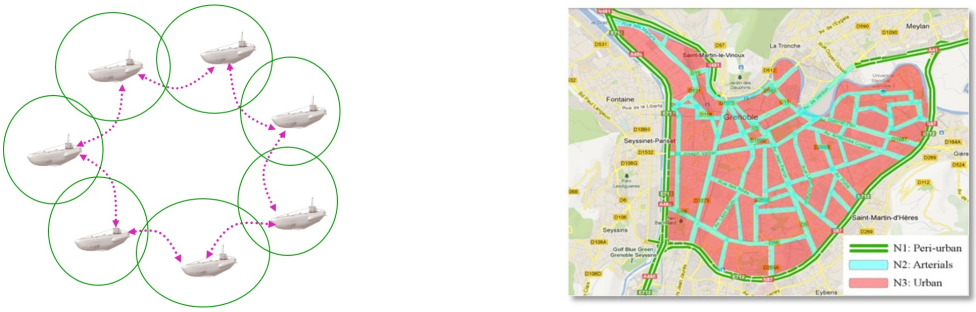


Figure 1. Left: a system of autonomous agents, where the network structure is created by the feedback, used to coordinate agents towards a common goal. Right: a system naturally having a network structure.

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, navigation 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.
- Inertial navigation and 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.
- Collaborative indoor and outdoor navigation of pedestrians.

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, transportation modes and human activities recognition. These units are composed of accelerometers, magnetometers and gyroscopes, sensors that we find usually in smartphones, tablets and smartwatches. 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

- C. Canudas-de-Wit was the General Chair of IEEE Conference on Decision and Control 2019 (CDC) in Nice (11-13 Dec. 2019).
- H. Fourati was elected as member of CNU61 (Conseil national des universités, Génie informatique, Automatique et Traitement du Signal), 2020-2023.
- H. Fourati has co-edited the book “Cooperative Localization and Navigation: Theory, Research and Practice”, by Taylor and Francis Group LLC.

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, Remi 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 [22] and in [48].

7.1.2. Boundary Control for Output Regulation in Scale-FreePositive Networks

Participants: D. Nikitin, C. Canudas-de-Wit [Contact person], P. Frasca.

This work addresses the problem of controlling aggregate quantities in large networks. More precisely, we deal with the problem of controlling a scalar output of a large-scale positive scale-free network to a constant reference value. We design an output-feedback controller such that no information about state vector or system matrices is needed. This controller can have arbitrary positivegains, and only one sufficient sign condition on system matricesshould be satisfied. This controller can be used to regulate the average state in a large-scale network with control applied to boundary nodes of the domain [51].

7.1.3. A functional approach to target controllability of networks

Participants: C. Commault, J. Van Der Woude [TU Delft], P. Frasca [Contact person].

In the control of networks, it is natural to consider the problem of controlling a limited number of *target nodes* of a network. Equivalently, we can see this problem as controlling the target variables of a structured system, where the state variables of the system are associated to the nodes of the network. We deal with this problem from a different point of view as compared to most recent literature. Indeed, instead of considering controllability in the Kalman sense, that is, as the ability to drive the target states to a desired value, we consider the stronger requirement of driving the target variables as time functions. The latter notion is called functional target controllability. We think that restricting the controllability requirement to a limited set of important variables justifies using a more accurate notion of controllability for these variables. Remarkably, the notion of functional controllability allows formulating very simple graphical conditions for target controllability in the spirit of the structural approach to controllability. The functional approach enables us, moreover, to determine the smallest set of steering nodes that need to be actuated to ensure target controllability, where these steering nodes are constrained to belong to a given set. We show that such a smallest set can be found in polynomial time. We are also able to classify the possible actuated variables in terms of their importance with respect to the functional target controllability problem. This research is reported in [16].

7.1.4. Cyber-Physical Systems: a control-theoretic approach to privacy and security

Participants: F. Garin [Contact person], A. Kibangou, S. Gracy [KTH Stockholm], S.m. Fosson [Politecnico di Torino].

Cyber-physical systems are composed of many simple components (agents) with interconnections giving rise to a global complex behaviour. One line or research on security of cyber-physical systems models an attack as an unknown input being maliciously injected in the system. 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: structural results, true for almost all interaction weights, and strongly structural results, true for all non-zero interaction weights. Our results in 2019 concern structural and strongly structural ISO for time-varying systems [19], strongly structural ISO for time-invariant systems [46]. Moreover in [44] we study delay-L left-invertibility, where the input reconstruction is allowed to take L time steps instead of requiring immediate reconstruction in a single step. We obtain preliminary results for structural delay-L left-invertibility, which include a full characterization for the case where the input is scalar, and for the cases where L is one and two, while the general case remains an open problem. When the conditions for ISO are satisfied, one can run well-known algorithms in the same vein as a Kalman filter, in order to reconstruct the state and the unknown input from noisy measurements. In [43], we consider cases where the system is not ISO, and we exploit compressive sensing techniques in order to obtain nevertheless a unique reconstruction of the input, under the assumption that the input is highly sparse (e.g., when only one or few states are under attack, albeit the attack position is unknown).

7.1.5. Collaborative monitoring of network structural robustness

Participants: A. Kibangou [Contact person], T.m.d. Tran [Univ. of Danang].

Interacting systems can be naturally viewed as networks modelled by graphs, whose vertices represent the components of the system while edges stand for the interactions between these components. The efficiency of a network of a network can be evaluated through its functional robustness and structural robustness. The former usually stands for robustness against noise while the latter is related to the network performance despite changes in network topology (node or edge failure). Structural robustness has been an important topic in various domains: in distribution networks (e.g. power or water distribution networks), breakdowns can prevent service to customers; in communication networks, equipment failures may disrupt the network and block users from communicating; in contact networks, removing nodes (persons) by means of vaccination can prevent epidemic propagation. In [31] we have considered the critical threshold of a network and the effective graph resistance (Kirchhoff index) of a sub-graph characterizing the interconnection of sub-networks, that are partitioned from the given network as robustness metric. In which, the critical threshold depends only on the two first moments of the degree distribution while the Kirchhoff index can be computed with Laplacian eigenvalues. Therefore, we show how to estimate jointly the Laplacian eigenvalues and the two first moments of the degree distribution in a distributed way.

7.1.6. Estimation of the average state in large scale networks

Participants: A. Kibangou [Contact person], C. Canudas-de-Wit, U. Niazi, D. Deplano [Univ. Cagliari].

State estimation for monitoring large-scale systems requires tremendous amounts of computational and sensing resources, which is impractical in most applications. However, knowledge of some aggregated quantity of the state suffices in several applications. Processes over physical networks such as traffic, epidemic spread, and thermal control are examples of large-scale systems. Due to the diffusive nature of these systems, the average state is usually sufficient for monitoring purposes. For instance, estimating the average traffic density in some sector of a traffic network helps to monitor the congestion effectively. In the event of an epidemic, estimating the average proportion of infected people over several towns, which are interconnected through people commuting for work or other purposes, helps to devise the preventive measures for controlling the epidemic spread. For the temperature regulation of a building, the thermistors can only be placed either on the walls or the roof, therefore, estimating the average temperature of the interior of a large corridor is crucial. Other examples include the averaging systems such as opinion networks and wireless sensor networks where the average state is of paramount importance. In [40] we address observability and detectability of the average state of a network system when few gateway nodes are available. To reduce the complexity of the problem, the system is transformed to a lower dimensional state space by aggregation. The notions of average observability and average detectability are then defined, and the respective necessary and sufficient conditions are provided. In [25] we provide a computationally tractable necessary and sufficient condition for the existence of an average state observer for large-scale linear time-invariant (LTI) systems. Two design procedures, each with its own significance, are proposed. When the necessary and sufficient condition is not satisfied, a methodology is devised to obtain an optimal asymptotic estimate of the average state. In particular, the estimation problem is addressed by aggregating the unmeasured states of the original system and obtaining a projected system of reduced dimension. This approach reduces the complexity of the estimation task and yields an observer of dimension one. Moreover, it turns out that the dimension of the system also does not affect the upper bound on the estimation error.

7.1.7. Structure-based Clustering Algorithm for Model Reduction of Large-scale Network Systems

Participants: C. Canudas-de-Wit [Contact person], U. Niazi, J. Scherpen [Univ. Groningen], X. Cheng [Univ. Groningen].

In [41], A model reduction technique is presented that identifies and aggregates clusters in a large-scale network system and yields a reduced model with tractable dimension. The network clustering problem is translated to a graph reduction problem, which is formulated as a minimization of distance from lumpability.

The problem is a non-convex, mixed-integer optimization problem and only depends on the graph structure of the system. We provide a heuristic algorithm to identify clusters that are not only suboptimal but are also connected, that is, each cluster forms a connected induced subgraph in the network system.

7.2. Control of multi-agent systems and opinion dynamics

7.2.1. Robust average consensus over unreliable networks

Participants: F. Acciani [Univ. Twente], P. Frasca [Contact person], G. Heijenk [Univ. Twente], A. Stoorvogel [Univ. Twente].

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 [11] provides a complete account of our results.

7.2.2. Message-passing computation of harmonic influence in social networks

Participants: W. S. Rossi [Univ. Groningen], 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 [29], we prove that the algorithm converges on general graphs.

7.2.3. Hybrid models of opinion dynamics

Participants: P. Frasca [Contact person], S. Tarbouriech [LAAS CNRS], L. Zaccarian [LAAS CNRS].

Hybrid dynamical systems are a promising framework to model social interactions. In this research line, we are beginning to use tools from the theory of hybrid systems to study opinion dynamics on networks with opinion-dependent connectivity. According to the hybrid framework, our dynamics are represented by the combination of continuous flow dynamics and discrete jump dynamics. The flow embodies the attractive forces between the agents and is defined by an ordinary differential equation whose right-hand side is a Laplacian, whereas the jumps describe the activation or deactivation of the pairwise interactions between agents. We first reformulate the classical Hegselmann–Krause model in this framework and then define a novel interaction model, which has the property of being scale-invariant. We study the stability and convergence properties of both models by a Lyapunov analysis, showing convergence and clusterization of opinions [18].

7.2.4. Stability of Metabolic Networks

Participants: F. Garin [Contact person], B. Piccoli [Rutgers Univ. Camden], N. Merrill [Rutgers Univ. Camden], Z. An [Rutgers Univ. Camden], S. Mc Quade [Rutgers Univ. Camden].

Quantitative Systems Pharmacology (QSP) aims to gain more information about a potential drug treatment on a human patient before the more expensive stages of development begin. QSP models allow us to perform insilico experiments on a simulated metabolic system that predicts the response ofperturbing a flux. The methodology named LIFE (Linear-in-Flux Expressions) was developed with the purpose of simulating and analyzing large metabolic systems. These systems can be associated to directed graphs: the edges represent the reaction rates (fluxes), and the vertices represent quantities of chemical compounds (metabolites). In [23], we study LIFE systems, addressing two main problems: 1. for fixed metabolite levels, find all fluxes for which the metabolite levels are an equilibrium, and 2. for fixed fluxes, find all metabolite levels which are equilibria for the system. We show how stability analysis from the fields of network flows, compartmental systems, control theory and Markov chains apply to LIFE systems.

7.3. Transportation networks and vehicular systems

7.3.1. Heterogeneity in synchronization: an adaptive control approach, with applications to vehicle platooning

Participants: S. Baldi [Univ. Delft], 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 [12] and of output-feedback [14], under the assumption that the topology of the interconnections has no circuits. Further investigation on relaxing this restrictive assumption is in progress. We also showed that agents need no leader to synchronise, even in presence of heterogeneity [13].

7.3.2. Stability of vehicle platoons with AVs

Participants: V. Giammarino [Univ. Delft], M. Lv [Univ. Delft], P. Frasca [Contact person], M.l. Delle Monache, S. Baldi [TU Delft].

A key notion to understand the impact of Autonomous Vehicles on traffic is the notion of *stability* of the vehicle collective motion. In this line of research, we have sought criteria to determine when stop-and-go waves form in platoons of human-driven vehicles, and when they can be dissipated by the presence of an autonomous vehicle. Our analysis takes the start from the observation that the standard notion of string/ring stability definition, which requires uniformity with respect to the number of vehicles in the platoon, is too demanding for a mixed traffic scenario. The setting under consideration is the following: the vehicles run along a ring road and the human-driven vehicles obey a combined follow-the-leader and optimal velocity model, while the autonomous vehicle obeys an appropriately designed model. The criteria are tested on a linearized version of the resulting platoon dynamics and simulation tests using nonlinear model are carried out [45].

7.3.3. Control and estimation using autonomous vehicles

Participants: R. Stern [Vanderbilt University], S. Cui [Temple University], M.l. Delle Monache [Contact person], T. Liard, Y. Chen [Vanderbilt University], R. Bhadani [University of Arizona], M. Bunting [University of Arizona], M. Churchill [UIUC], N. Hamilton [Vanderbilt University], R. Haulcy [Yale University], H. Pohlmann [Temple University], F. Wu [UC Berkeley], B. Piccoli [Rutgers University], B. Seibold [Temple University], J. Sprinkle [University of Arizona], D.b. Work [Vanderbilt University].

It is anticipated that in the near future, the penetration rate of vehicles with some autonomous capabilities will increase on roadways. In [30], we analyze the potential reduction of vehicular emissions caused by the whole traffic stream, when a small number of autonomous vehicles are designed to stabilize the traffic flow and dampen stop-and-go waves. To demonstrate this, vehicle velocity and acceleration data are collected form a series of field experiments that use a single autonomous-capable vehicle to dampen traffic waves on a circular ring road with 20 to 21 human-piloted vehicles. From the experimental data, vehicle emissions (hydrocarbons, carbon monoxide, carbon dioxide, and nitrogen oxides) are estimated using the MOVES emissions models. We find that vehicle emissions of the entire fleet may be reduced by between 15% (for carbon dioxide) and 73% (for nitrogen oxides) when stop-and-go waves are reduced or eleiminated by the dampening action of the autonomous vehicle in the flow of human drivers. This is possible if a small fraction (5%) of vehicles are autonomous and designed to actively dampen traffic waves. In [57], we look at the problem of traffic control in which an autonomous vehicle is used to regulate human piloted traffic to dissipate stop and go traffic waves. We investigate the controllability of well-known microscopic traffic flow models: i) the Bando model (also known as the optimal velocity model), ii) the follow-the-leader model and iii) a combined optimal velocity – follow the leader model. Based on the controllability results, we propose three control strategies for an autonomous vehicles to stabilize the human piloted traffic. After, we simulate the control effects on the microscopic models of human drivers in numerical experiments to quantify the potential benefits of the controllers. Based on the simulations, finally we conduct a field experiment with 22 human drivers and a fully autonomous-capable vehicle, to assess the feasibility of autonomous vehicle based traffic control on real human piloted traffic. We show that both in simulation and in the field test an autonomous vehicle is able to dampen waves generated by 22 cars, and that as a consequence, the total fuel consumption of all vehicles is reduced by up to 20%. In [17], we consider a partial differential equation – ordinary differential equation system to describe the dynamics of traffic with autonomous vehicles. In the model the bulk flow is represented by a scalar conservation law, while each autonomous vehicle is described by a car following model. The autonomous vehicles act as tracer vehicles in the flow and collect measurements along their trajectory to estimate the bulk flow. The main result is to prove theoretically and show numerically how to reconstruct the correct traffic density using only the measurements from the autonomous vehicles.

7.3.4. Two-dimensional traffic flow models

Participants: S. Mollier, M.l. Delle Monache, C. Canudas-de-Wit [Contact person], B. Seibold [Temple University].

In [24], we introduce a new traffic flow model for a dense urban area. We consider a two-dimensional conservation law in which the velocity magnitude is given by the fundamental diagram and the velocity direction is constructed following the network geometry. The model is validated using synthetic data from Aimsun and a reconstruction technique to recover the 2D density from the data of individual vehicles is proposed. In [50], [49], we introduce a two dimensional and multi-layer traffic model with a new planning and decision making method in large scale traffic networks for predicting how traffic evolves in special events, emergencies and changes in the city mobility demands. The proposed method is based on a 2-D aggregated traffic model for large scale traffic networks which describes traffic evolution as a fluid in two space dimensions extended with additional state density variables, each one associated to a particular layer describing vehicles evolving in different directions. The model is a 2D-PDE described by a system of conservation laws. For this specific case, the resulting PDE is not anymore hyperbolic as typically the LWR model but results in a hybrid hyperbolic-elliptic PDE depending on the density level. In this case, usual numerical schemes may be not valid and often lead to oscillation in the solution. Thus, we consider a high order numerical scheme to improve the numerical solution. Finally, the model is used to predict how the typical traffic evolution will be impacted in particular scenarios like special events or changes in demands.

7.3.5. High-fidelity vehicle trajectory data

Participants: F. Wu [UC Berkeley], R. Stern [Vanderbilt University], S. Cui [Temple University], M.l. Delle Monache [Contact person], R. Bhadani [University of Arizona], M. Bunting [University of Arizona], M. Churchill [UIUC], N. Hamilton [Vanderbilt University], R. Haulcy [Yale University], B. Piccoli [Rutgers University], J. Sprinkle [University of Arizona], D.b. Work [Vanderbilt University], B. Seibold [Temple University].

High fidelity-vehicle trajectory data is becoming increasingly important in traffic modeling, especially to capture dynamic features such as stop-and-go waves. In [34], we present data collected in a series of eight experiments on a circular track with human drivers. The data contains smooth flowing and stop-and-go traffic conditions. The vehicle trajectories are collected using a panoramic 360-degree camera, and fuel rate data is recorded via an on-board diagnostics scanner installed in each vehicle. The video data from the 360-degree camera is processed with an offline unsupervised algorithm to extract vehicle trajectories from experimental data. The trajectories are highly accurate, with a mean positional bias of less than 0.01 m and a standard deviation of 0.11 m. The velocities are also validated to be highly accurate with a bias of 0.02 m/s and standard deviation of 0.09 m/s.

7.3.6. Robust tracking control design for fluid traffic dynamics

Participants: L. Tumash, C. Canudas-de-Wit [contact person], M.l. Delle Monache.

In [53] we analyze the boundary control of the traffic system described by the LWR model with a triangular fundamental diagram and a space-dependent indomain unknown disturbance, which can be described as an inhomogeneous transport equation. The controller design strategy aims first at stabilizing the deviation from the desired time-dependent trajectory and then at minimizing the deviation in the sense of two possible space-norms.

7.3.7. Urban traffic control

Participants: C. Canudas-de-Wit [Contact person], F. Garin, P. Grandinetti.

In [20] we study near-optimal operation of traffic lights in an urban area, e.g., a town or a neighborhood. The goal is on-line optimization of traffic lights schedule in real time, so as to take into account variable traffic demands, with the objective of obtaining a better use of the road infrastructure. More precisely, we aim at maximizing total travel distance within the network, together with balancing densities across the network. The complexity of optimization over a large area is addressed both in the formulation of the optimization problem, with a suitable choice of the traffic model, and in a distributed solution, which not only parallelizes computations, but also respects the geometry of the town, i.e., it is suitable for an implementation in a smart infrastructure where each intersection can compute its optimal traffic lights by local computations combined with exchanges of information with neighbor intersections.

7.3.8. Modeling and control strategies for improving environmental sustainability of road transportation

Participants: B. Othman, G. de Nunzio [IFP Energies nouvelles], D. Di Domenico [IFP Energies nouvelles], C. Canudas-de-Wit [Contact person].

As road transportation energy use and environmental impact are globally rising at an alarming pace, authorities seek in research and technological advancement innovative solutions to increase road traffic sustainability. The unclear and partial correlation between road congestion and environmental impact is promoting new research directions in traffic management. We review the existing modeling approaches to accurately represent traffic behavior and the associated energy consumption and pollutant emissions [26]. The review then covers the transportation problems and control strategies that address directly environmental performance criteria, especially in urban networks. A discussion on the advantages of the different methods and on the future outlook for the eco-traffic management completes the proposed survey.

7.3.9. Data analysis for smart multi-modal transportation planning

Participants: A. Kibangou [Contact person], T. Moyo [Univ. of Johannesburg], W. Musakwa [Univ. of Johannesburg].

Modern cities have managed to balance the relationship between supply and demand of services through clear planning strategies which advocate smart solutions to the ever increasing demand for public tranportation services. The end goal is not to prohibit citizens to use their private cars, but to create an enabling smart system at a suitable scale which would lead to citizens not needing to own or drive a car. Having an efficiently and effectively run public transportation system is a crucial and indispensable factor for any developing city region. However as the provision of public transportation is a multifaceted process, with intertwining elements such as culture, politics,finance and shareholder interests, smart means of monitoring and mitigating the challenges faced in the provision of public transportation need to be developed continuously. The Gauteng city region is likewise faced with this challenge. With this region being the economic hub of South Africa, this has greatly affected the operation of the Gautrain system and the BRT systems within the region, as more and more people require a fast and reliable transportation means to move in and out the metropolitan cities. The study relied on a questionnaire-based survey that was administered to 60 respondents. The questionnaire had both closed and open-ended questions which were administered online through Google forms so as to obtain a good response rate from commuters who reside within the study area. The questions centred on identifying factors influencing the commuter's travelling patterns. Gautrain Management Agency reports and literature were also utilised to supplement infomation gleaned from the questionnaire. Besides the questionnaire, secondary data was collected from Twitter (tweets) concerning the Gaubus and Gautrain (between the period of August to November 2018). Posts from 380 users were analysed. This data was used to spatially identify POI of Gaubus users and also to identify the spatial relationship between land use activities, Gaubus routes, Gaubus stops, Gautrain stations and Guatrain routes. A neighboorhood analysis was run using a focal statistics based tool to map the spatial distribution of commuters of the Gaubus [56].

7.3.10. Location of turning ratio and flow sensors for flow reconstruction in large traffic networks

Participants: M. Rodriguez-Vega, C. Canudas-de-Wit [Contact person], H. Fourati.

We examine the problem of minimizing the number of sensors needed to completely recover the vehicular flow in a steady state traffic network [28]. We consider two possible sensor technologies: one that allows the measurement of turning ratios at a given intersection and the other that directly measures the flow in a road. We formulate an optimization problem that finds the optimal location of both types of sensors, such that a minimum number is required. To solve this problem, we propose a method that relies on the structure of the underlying graph, which has a quasi-linear computational complexity, resulting in less computing time when compared to other works in the literature. We evaluate our results using dynamical traffic simulations in synthetic networks.

7.4. Multisensor data fusion for navigation

7.4.1. Heterogeneity and uncertainty in distributed estimation from relative measurements

Participants: C. Ravazzi [Politecnico Torino], N.k. Chan [Univ. Groningen], P. Frasca [Contact person].

This work, presented in [27], 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.

7.4.2. Cooperative localization and navigation: Theory, research, and practice

Participants: C. Gao [Naval Aviation University, China], G. Zhao [Naval Aviation University, China], H. Fourati [Contact person].

The idea of the book [58] comes as a response to the immense interest and strong activities in the field of cooperative localization and navigation during the past few years, both in theoretical and practical aspects. This book is targeted toward researchers, academics, engineers, and graduate students working in the field of sensor fusion, filtering, and signal processing for localization and navigation. This book, entitled Cooperative Localization and Navigation: Theory, Research and Practice, captures the latest results and techniques for cooperative navigation drawn from a broad array of disciplines. It is intended to provide the reader with a generic and comprehensive view of contemporary state estimation methodologies for localization and navigation, as well as the most recent researches and novel advances on cooperative localization and navigation task exploring the design of algorithms and architectures, benefits, and challenging aspects, as well as a potential broad array of disciplines, including wireless communication, in-door localization, robotics, and emergency rescue. These issues arise from the imperfection and diversity of cooperative sources, the contention and collision of communication channels, the selection and fusion of cooperative data, and the nature of the application environment. The issues that make cooperative-based navigational state estimation a challenging task, and which will be discussed through the different chapters of the book, are related to (1) the nature and model of sensors and cooperative sources (e.g., range-based sensor, angle-based sensor, inertialsensor, and vision sensor); (2) the communication medium and cooperative strategies; (3) the theoretical developments of state estimation and data fusion; and (4) the applicable platforms.

7.4.3. Data fusion from multi-inertial and magnetic sensors

- **Attitude estimation from multi-sensor observations**
  Participants: J. Wu [Hong Kong University of Science and Technology], Z. Zhou [University of Electronic Science and Technology of China], H. Fourati [Contact person], R. Li [University of Electronic Science and Technology of China], M. Liu [Hong Kong University of Science and Technology], A. Kibangou, A. Makni.

  Focusing on generalized sensor combinations, we deal with attitude estimation problem using a linear complementary filter [36]. The quaternion observation model is obtained via a gradient descent algorithm (GDA). An additive measurement model is then established according to derived results. The filter is named as the generalized complementary filter (GCF) where the observation model is simplified to its limit as a linear one that is quite different from previous-reported brute-force computation results. Moreover, we prove that representative derivative-based optimization algorithms are essentially equivalent to each other. Derivations are given to establish the state model based on the quaternion kinematic equation. The proposed algorithm is validated under several experimental conditions involving free-living environment, harsh external field disturbances and aerial flight test aided by robotic vision. Using the specially designed experimental devices, data acquisition and algorithm computations are performed to give comparisons on accuracy, robustness, time-consumption and etc. with representative methods. The results show that not only the proposed filter can give fast, accurate and stable estimates in terms of various sensor combinations, but it also produces robust attitude estimation in the presence of harsh situations e.g. irregular magnetic distortion. In other recent work, related to the attitude estimation, we add some corrections to update that version [35]. In [21], we propose the design of an attitude estimation algorithm for a rigid body subject to accelerated maneuvers. Unlike the current literature where the process model is usually driven by triaxial gyroscope measurements, we investigate a new formulation of the state-space model where the process model is given by triaxial accelerometer measurements. The observation model is given by triaxial gyroscope and magnetometer measurements. The proposed model is written as a descriptor system and takes the external acceleration sensed by the accelerometer into account. Based on this model, a Quaternion Descriptor Filter (QDF) is developed and its performance is evaluated through simulations and experimental tests in pedestrian navigation.

- **Convexity analysis of optimization framework of attitude determination**
  Participants: J. Wu [Hong Kong University of Science and Technology], Z. Zhou [University of Electronic Science and Technology of China], H. Fourati [Contact person], M. Liu [Hong Kong University of Science and Technology].

  In the past several years, there have been several representative attitude determination methods developed using derivative-based optimization algorithms. Optimization techniques e.g. gradient-descent algorithm (GDA), Gauss-Newton algorithm (GNA), LevenbergMarquadt algorithm (LMA) suffer from local optimum in real engineering practices. A brief discussion on the convexity of this problem was presented recently, stating that the problem is neither convex nor concave. In our work, we give analytic proofs on this problem. The results reveal that the target loss function is convex in the common practice of quaternion normalization, which leads to non-existence of local optimum.

- **Behaviors classification based distance measuring system for pedestrians via a foot-mounted multi-inertial sensors**
  Participants: Z. Zhou [University of Electronic Science and Technology of China], S. Mo [University of Electronic Science and Technology of China], J. Wu [Hong Kong University of Science and Technology], H. Fourati [Contact person].

  We developed 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 [37]. Conventional methods for step detection and step length estimation cannot adapt well to the general pedestrian applications since the parameters in these methods may vary for different persons and motions. In this paper, 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. Finally, real experiments are carried out to verify our proposed step detection method and step length model. The results show that the proposed method outperforms the existing representative methods and it exhibits the merits of accuracy and adaptability for different persons in real time and significantly improves the accuracy of distance measuring.

- **Human activities and postures recognition: from inertial measurements to quaternion-based approaches**
  Participants: M. Zmitri, H. Fourati [contact person], N. Vuillerme [AGEIS, UGA].

  We present two approaches to assess the effect of the number of inertial sensors and their location placements on recognition of human postures and activities [38]. Inertial and Magnetic Measurement Units (IMMUs)—which consist of a triad of three-axis accelerometer, three-axis gyroscope, and three-axis magnetometer sensors—are used in this work. Five IMMUs are initially used and attached to different body segments. Placements of up to three IMMUs are then considered: back, left foot, and left thigh. The subspace k-nearest neighbors (KNN) classifier is used to achieve the supervised learning process and the recognition task. In a first approach, we feed raw data from three-axis accelerometer and three-axis gyroscope into the classifier without any filtering or pre-processing, unlike what is usually reported in the state-of-the-art where statistical features were computed instead. Results show the efficiency of this method for the recognition of the studied activities and postures. With the proposed algorithm, more than 80% of the activities and postures are correctly classified using one IMMU, placed on the lower back, left thigh, or left foot location, and more than 90% when combining all three placements. In a second approach, we extract attitude, in term of quaternion, from IMMUs in order to more precisely achieve the recognition process. The obtained accuracy results are compared to those obtained when only raw data is exploited. Results show that the use of attitude significantly improves the performance of the classifier, especially for certain specific activities. In that case, it was further shown that using a smaller number of features, with quaternion, in the recognition process leads to a lower computation time and better accuracy.

- **Improving inertial velocity estimation through magnetic field gradient-based extended kalman filter**
  Participants: M. Zmitri, H. Fourati [contact person], C. Prieur [GIPSA-Lab, UGA].

  We focused on the velocity estimation problem of a rigid body and how to improve it with magnetoinertial sensors-based theory [55]. We provide a continuous-time model that describes the motion of the body and we augment it after by introducing a new magnetic field gradient equation instead of using its value directly as an input for the model, as done usually in the corresponding literature. We investigate the advantage of moving to higher order spatial derivatives of the magnetic field in the estimation of velocity. These derivatives are computed thanks to a determined arrangement of magnetometers array. Within this framework, a specific set configuration of Extended Kalman Filters (EKFs) is proposed to focus mainly on the estimation of velocity and attitude of the body, but includes also an estimation of the magnetic field and its gradient. Some simulations for a specific scenario are proposed to show the improvements that we bring to the velocity estimation.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

TMI-V (Tachymètre Magnéto-Inertiel couplé Vision). Co-PI: H. Fourati (2018-2022)
The objective of the TMI-V project is the indoor localization without infrastructure, by developing an autonomous, precise, robust solution with no prior knowledge of the environment integrated in equipment worn on the upper body to be used in virtual reality and augmented reality applications. An array of magnetometers and inertial sensors will be used. The project is ongoing, in collaboration with SysNav company.

9. Partnerships and Cooperations

9.1. Regional Initiatives

DATASAFE (Understanding data accidents for traffic safety). PI: M.L. Delle Monache (2018-2019)
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.

MAVIT (Modeling autonomous vehicles in traffic flow). PI: M.L. Delle Monache (2018-2019)
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

SPACE (NanoSatellite Project: Advanced modelling and Control of attitude dynamics for quantum communication). PI: H. Fourati (2018-2019)
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).

CAPTIMOVE (CAPture et analyse d’acTivités humaInes par MOdules inertiels : vers une solution adaptée à la naVigation multimodalE urbaine intelligente). PI: H. Fourati (2018-2019)

Mobility is currently evolving in urban scenarios and multimodality today is the key tomore efficient transportation. It is important to analyze the ecological impact of the varioustransportation modes, to be able to detect the mode used by the commuter and the rule usedto switch from one mode to another. The ultimate goal is to suggest smarter itineraries tocommuters. To this purpose, detection and classification of activities in human mobility fromhis principal residence to his destination (for example, place of work, place of entertainment,etc.) is an important study to carry out. We aim to identify, with high precision, the natureof the transportation modes used during the day (walking, cycling, public transportation, car,etc.) as well as transitions from one mode to another. To reach this goal, we will use inertial and attitude modules, embedded in most inertial units, connected watches and smartphones.These technological tools constitute truly innovative and promising instrumentation for bothnon-invasive automatic capture information in situ, over extended periods, only for accurateand reliable analysis of activities of a person during his/her trip. In terms of research, we willexploit techniques from Machine Learning and state estimation to address this issue. A studyshall be conducted to determine the type, number and location of sensors to be used., Issuesrelated to the quality of data to be provided to algorithms and how to detect and discarderroneous ones from our computation process, will be also addressed. This research finds itsmajor future interest later in the development of a multimodal intelligent navigation systemfor indoor and outdoor environments. These results, once obtained, can also be used to studyand analyze the behavior (choice) of users regarding pedestrian navigation (walking) or theuse of modes of transport (convenience, cost, speed, safety and more and more frequentlyeffects on the environment) or respect for the privacy of individuals (dynamic anonymizationof data while retaining their usefulness).

9.2. National Initiatives

DOOM (Systems-theory for the Disorders Of Online Media). 80 PRIME from CNRS MITI (2019–2022). PI: P. Frasca
Online social media have a key role in contemporary society and the debates that take place on them are known to shape political and societal trends. For this reason, pathological phenomena like the formation of "filter bubbles" and the viral propagation of "fake news" are observed with concern. The scientific assumption of this proposal is that these information disorders are direct consequences of the inherent nature of these communication media, and more specifically of the collective dynamics of attention thereby. In order to capture these dynamics, this proposal advocates the mathematical modelling of the interplay between the medium (algorithmic component) and the users (human component). The resulting dynamics shall be explored by a system-theoretic approach, using notions such as feedback and stability. This quantitative and rigorous approach will not only unlock fundamental insights but also deliver suggestions on suitable policies to manage the media.

HANDY (Hybrid and Networked Dynamical Systems). ANR PRC (2019-2022). Co-PI: P. Frasca
Networked dynamical systems are ubiquitous in current and emerging technologies. From energy grids, fleets of connected autonomous vehicles to online social networks, the same scenario arises in each case: dynamical units interact locally to achieve a global behavior. When considering a networked system as a whole, very often continuous-time dynamics are affected by instantaneous changes, called jumps, leading to so-called hybrid dynamical systems. Hybrid phenomena thus play an essential role in these control applications, and call upon the development of novel adapted tools for stability and performance analysis and control design. In this context, the aim of HANDY project is to provide methodological control-oriented tools for realistic networked models, which account for hybrid phenomena. The project brings together researchers from LAAS in Toulouse, CRAN in Nancy, GIPSA in Grenoble and LSS in Gif-sur-Yvette, with expertise in various domains of automatic control, ranging from geometric control and optimization, switched systems, hybrid dynamics, nonlinear control, and multi-agent systems. See also: http://projects.laas.fr/handy

AgileWorld-MRSEI. PI: A. Kibangou 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. For this purpose a workshop was organized in November 2019 with the partners of the project in Grenoble.

9.3. European Initiatives

9.3.1. Collaborations in European Programs, Except FP7 & H2020

COST (Mathematical models for interacting dynamics on networks). Action no. 18232, 2019-2023, Management committee substitute member. PI: M.L. Delle Monache

Many physical, biological, chemical, financial or even social phenomena can be described by dynamical systems. It is quite common that the dynamics arises as a compound effect of the interaction between subsystems in which case we speak about coupled systems. This Action shall study such interactions in particular cases from three points of view: 1. the abstract approach to the theory behind these systems, 2. applications of the abstract theory to coupled structures like networks, neighbouring domains divided by permeable membranes, possibly non-homogeneous simplicial complexes, etc., 3. modelling real-life situations within this framework. The purpose of this Action is to bring together leading groups in Europe working on a range of issues connected with modelling and analysing mathematical models for dynamical systems on networks. It aims to develop a semigroup approach to various (non-)linear dynamical systems on networks as well as numerical methods based on modern variational methods and applying them to road traffic, biological systems, and further real-life models. The Action also explores the possibility of estimating solutions and long time behaviour of these systems by collecting basic combinatorial information about underlying networks

9.4. International Initiatives

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

MEMENTO (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

PI: M.L. Delle Monache

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.

9.4.2. Participation in Other International Programs

(Mean field game models for traffic application). Rutgers Global Grant - International collaborative research grant: International partner : Rutgers University - Camden (USA). PI: M.L. Delle Monache

This project focuses on the theoretical tools for traffic systems to mitigate traffic events that adversely affect. Specifically, the project will build algorithms to mitigate “phantom” traffic jams, which are instabilities caused by human driving behavior, lane changes, and other disturbances. This project is premised on the concept that connected and autonomous vehicles (CAVs) can act as instability pacifiers and enable a new era of freeway traffic management in which CAVs themselves are part of the traffic control system. The stabilizing Lagrangian (i.e., mobile) control signal will be fed directly to the vehicles, which will adjust their speed and lanes to match the requirements of the control.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Raphael Stern (University of Minnesota (USA)) visited the team in March 2019 to work with Maria Laura Delle Monache and Thibault Liard, in the framework of the associated team MEMENTO.

9.5.2. Visits to International Teams

- P. Frasca is a Visiting Scientist at the IEIIT-CNR Institute, National Research Council CNR, Turin, Italy. By this collaboration, he performs research on distributed estimation in sensor networks and distributed control of social networks. He visited Turin three times in 2019. He is also 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.
- Maria Laura Delle Monache visited Rutgers University – Camden in March and in November 2019 to work with Prof. Piccoli in the framework of the Rutgers collaborative grant.
- Maria Laura Delle Monache visited Vanderbilt University in November 2019 in the framework of the of the associated team MEMENTO.
- Stephane Mollier visited Temple University in January 2019 to discuss with Prof. Seibold concerning 2D traffic models.
- A. Kibangou visited the University of Johannesburg (South Africa) in March and November 2019. 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. He also attended the first French-South African Science and Innovation days (December 2-3, 2019).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- C. Canudas-de-Wit was the General Chair of IEEE Conference on Decision and Control 2019 (CDC) in Nice (11-13 Dec. 2019).

10.1.1.2. Member of the Organizing Committees

- Maria Laura Delle Monache organized the workshop on “Lagrangian Control for Traffic Flow Smoothing in Mixed Autonomy Settings ”, CDC December 2019 (with Alexandre Bayen*, George J. Pappas, Benedetto Piccoli, Daniel B. Work, Jonathan Sprinkle, Maria Laura Delle Monache, Benjamin Seibold, Cathy Wu, Abdul Rahman Kreidieh, Eugene Vinitsky, Yashar Zeiynali Farid).
- Maria Laura Delle Monache organized a tutorial session on “Autonomous Vehicles and Traffic Control in Mixed Autonomy Environments” at CDC, December 2019 (with Jonathan Sprinkle, Ram Vasudevan, Dan Work).
- Team members organized the following invited sessions at the CDC 2019:

  • “Novel Approaches to Traffic Estimation and Control Using Automated Vehicles” (M.L. Delle Monache with R.Stern (University of Minnesota))
  • “Control for Large Scale Traffic Networks” (M.L. Delle Monache, C. Canudas de Wit with N. bekiaris-Liberis (Univeristy of Crete))
  • “Models and Control Methods Tor Traffic Networks” (M. L. Delle Monache with S. Siri and C. Pasquale (University of Genoa))
  • “Multi-Sensor Fusion Techniques for State Estimation in Navigation” (H. Fourati with A. Barrau (Safran), J. Farrell (University of California Riverside), M. Liu (Hong Kong University of Science) and Z. Zhou (University of Electronic Science and Technology of China).

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

- F. Garin is Associate Editor in the IEEE Control System Society Conference Editorial Board (this year, she served for CDC 2019, ACC 2020) and Associate Editor in the European Control Association (EUCA) Conference Editorial Board (this year, she served for ECC 2020).
- H. Fourati is member of (1) the International Program Committee (IPC) of international conferences STA'19, ICCAD'19, ISAECT'19 (2) the Technical Program Committee (TCP) for the International Conference on Indoor Positioning and Indoor Navigation (IPIN'19), Pisa (Italy), Sep. 2019 (3) Member of the publication chairs of the International Conference on Control, Automation and Diagnosis (ICCAD'19), Grenoble (France), Jul. 2019.

10.1.2.2. 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.

10.1.3. Journal

10.1.3.1. 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. He has also been lead guest editor of the special issue "Recent Advances on Data Fusion, Estimation in Navigation and Control" for Asian Journal of Control (AJC), 2019.

10.1.3.2. 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.

10.1.4. Invited Talks

P. Frasca gave the following talks:

• "The closed loop between opinion formation and personalised recommendations", Workshop "Network Dynamics in the Social, Economic, and Financial Sciences", Turin, Italy, November 5-8, 2019.

• "Non-smooth opinion dynamics", Workshop "European Network for Nonsmooth Dynamical Systems, Grenoble, September 18, 2019.

• “The closed loop between opinion formation and personalised recommendations”, Workshop “Reti sociali e comportamenti emergenti”, Napoli, February 4, 2019.

M.L. Delle Monache gave the following talks:

• Traffic flow implications of autonomous and partially autonomous vehicles, Workshop on "Connected and automated vehicles for energy efficient and environmental impact", IFPEN, Rueil-Malmaison, France, September 2019.

• Modeling autonomous vehicles in traffic flow, International Congress on Industrial and Applied Mathematics (ICIAM) 2019, Valencia, Spain, July 2019.

• Micro - macro models for traffic with autonomous vehicles, IPAM workshop on Autonomous vehicles, IPAM (UCLA), USA, February 2019.

• Traffic control and estimation with autonomous vehicles, Journée du groupe de travail en automatique et transports terrestres, Université Grenoble Alpes, France, January 2019.

• Traffic reconstruction using autonomous vehicles, Sixth Chilean Workskshop on Numerical Analysis of Partial Differential Equations (WONAPDE), Concepción, Chile, January 2019.

10.1.5. 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 (IFAC):

  • 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 Technical Committee 7.1 Automotive Control (C. Canudas-de-Wit);

  • IFAC Technical Committee 7.4 Transportation systems (C. Canudas-de-Wit);

  • IFAC Technical Committee 9.2 on Social Impact of Automation (P. Frasca);
- 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.
- A. Kibangou reviewed project proposals for NRF (South-African research agency)

10.1.6. Research Administration

- From July 2019, F. Garin is ‘Présidente du CES du Centre (Comité des Emplois Scientifiques)’.
- In Nov. 2019, F. Garin has been elected as ‘responsable du pôle Automatique et Diagnostic (PAD)’ at GIPSA-Lab, a role to be started in Jan. 2020.
- Since Nov. 2019, F. Garin is co-head of the CSP (Cyber-Physical Systems) action of Persyval2.
- A. Kibangou has been elected member of the research department MSTIC (mathematics, information and communication sciences) of Univ. Grenoble Alpes.
- A. Kibangou is co-head of the PCS (Pervasive Computing Systems) action of Persyval-Lab (until November 2019).
- A. Kibangou is academic director (L2) IUT1 (GEII).
- A. Kibangou is 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)

- H. Fourati is member of CNU61 (Conseil national des universités, Génie informatique, Automatique et Traitement du Signal) since 2016.
- H. Fourati is in charge of communication mission and visits to high school within the Department of Electrical Engineering, IUT1 Grenoble, France (2017-present).
- M.L. Delle Monache is member of two local committees at Inria Rhône-Alpes: "Commission de développement technologique (research engineers) and Comité des Études Doctorales (PhD grants CORDI-S).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

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, 111h, L1, IUT 1 (GEII), Univ. Grenoble Alpes, France.

Licence: H. Fourati, Réseaux locaux industriels, 34h, L2, IUT1 (GEII), Univ. Grenoble Alpes, France.

Licence: H. Fourati, Automatique, 39h, L3, UFR physique, Univ. Grenoble Alpes, France.

Licence: H. Fourati, Automatique continue et discrete, 27h, L2, IUT1 (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.

10.2.2. Supervision

- PhD: 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: 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: Ujjwal Pratap, Resilient control in scale-free networks, from Feb. 2019, co-advised by C. Canudas-de-Wit, F. Garin, and H. Sandberg (KTH Stockholm).

10.2.3. Juries

- P. Frasca was committee member of the PhD defense of:

  • Wenjing Yang. Influence Maximization in Social Networks. Universit Aix-Marseille, France. PhD advisors: Alessandro Giua and Leonardo Brenner. November 2019.

  • Zhiyang Ju. Persistent Communication Connectivity of Multi-agent Systems. University of Melbourne, Australia. PhD advisors: Dragan Nesic and Iman Shames. February 2019.
- F. Garin was committee member of the PhD defence of Gustav Nilsson, Lund Univ., in Feb. 2019; Thesis: On robust distributed control of transportation networks, supervisor: Giacomo Como, co-supervisor: Anders Rantzer.
- F. Garin was committee member of the PhD defence of Han Zhang, KTH Stochkolm, in Feb. 2019; Thesis: Optimizing Networked Systems and Inverse Optimal Control, supervisor: Xiaoming Hu, co-supervisor: Elias Jarlebring.
- F. Garin was committee member of the PhD defence of Tommaso Borzone, Univ. Lorraine, Nancy, in Sept. 2019; Thesis: Decentralised control of multi-agent systems: a hybrid formalism, supervisor: Irinel-Constantin Morarescu, co-supervisor: Marc Jungers.

10.3. Popularization

10.3.1. Education

Maria Laura Delle Monache gave a talk to students of 8th grade (4ème) and 11th Grade (1ère) in the Cérémonie de remise de prix des Olympiades de Mathématiques, La mobilité et les vehicules autonomes, , Université Grenoble Alpes, France, May 2019.

11. 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, forthcoming [*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

Articles in International Peer-Reviewed Journal

[11] F. ACCIANI, P. FRASCA, G. HEIJENK, A. STOORVOGEL.*Achieving robust average consensus over lossy wireless networks*, in "IEEE Transactions on Control of Network Systems", March 2019, vol. 6, n^{o} 1, p. 127-137 [*DOI :* 10.1109/TCNS.2018.2800407], https://hal.archives-ouvertes.fr/hal-01699122

[12] S. BALDI, P. FRASCA.*Adaptive synchronization of unknown heterogeneous agents: an adaptive virtual model reference approach*, in "Journal of The Franklin Institute", 2019, vol. 356, n^{o} 2, p. 935-955 [*DOI :* 10.1016/J.JFRANKLIN.2018.01.022], https://hal.archives-ouvertes.fr/hal-01702085

[13] S. BALDI, P. FRASCA.*Leaderless synchronization of heterogeneous oscillators by adaptively learning the group model*, in "IEEE Transactions on Automatic Control", 2019, p. 1-7 [*DOI :* 10.1109/TAC.2019.2914664], https://hal.archives-ouvertes.fr/hal-02334955

[14] 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", 2019, vol. 6, n^{o} 2, p. 515–525 [*DOI :* 10.1109/TCNS.2018.2844260], https://hal.archives-ouvertes.fr/hal-01809667

[15] C. COMMAULT, A. KIBANGOU.*Generic controllability of networks with identical SISO dynamical nodes*, in "IEEE Transactions on Control of Network Systems", 2019, p. 1-1, forthcoming [*DOI :* 10.1109/TCNS.2019.2950587], http://hal.univ-grenoble-alpes.fr/hal-02356358

[16] C. COMMAULT, J. VAN DER WOUDE, P. FRASCA.*Functional target controllability of networks: structural properties and efficient algorithms*, in "IEEE Transactions on Network Science and Engineering", 2019, p. 1-10, forthcoming [*DOI :* 10.1109/TNSE.2019.2937404], https://hal.archives-ouvertes.fr/hal-02334957

[17] M. L. DELLE MONACHE, T. LIARD, B. PICCOLI, R. E. STERN, D. B. WORK.*Traffic reconstruction using autonomous vehicles*, in "SIAM Journal on Applied Mathematics", 2019, vol. 75, n^{o} 5, p. 1748-1767 [*DOI :* 10.1137/18M1217000], https://hal.inria.fr/hal-01882309

[18] 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

[19] 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", June 2019, vol. 6, n^{o} 2, p. 897-905 [*DOI :* 10.1109/TCNS.2018.2880304], https://hal.archives-ouvertes.fr/hal-01918497

[20] 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", May 2019, vol. 27, n^{o} 3, p. 950-963 [*DOI :* 10.1109/TCST.2018.2807792], https://hal.inria.fr/hal-01727937

[21] A. MAKNI, A. KIBANGOU, H. FOURATI. *Data Fusion-Based Descriptor Approach for Attitude Estimation underaccelerated maneuvers*, in "Asian Journal of Control", July 2019, vol. 21, n^{o} 4, p. 1433-1442 [*DOI :* 10.1002/ASJC.2084], https://hal.archives-ouvertes.fr/hal-01982463

[22] N. MARTIN, P. FRASCA, C. CANUDAS DE WIT. *Large-scale network reduction towards scale-free structure*, in "IEEE Transactions on Network Science and Engineering", 2019, vol. 6, n^{o} 4, p. 711-723 [*DOI :* 10.1109/TNSE.2018.2871348], https://hal.archives-ouvertes.fr/hal-01885140

[23] N. J. MERRILL, Z. AN, S. T. MCQUADE, F. GARIN, K. AZER, R. E. ABRAMS, B. PICCOLI. *Stability of Metabolic Networks via Linear-In-Flux-Expressions*, in "Networks and Heterogeneous Media", March 2019, vol. 4, n^{o} 1, p. 101-130 [*DOI :* 10.3934/NHM.2019006], https://hal.inria.fr/hal-02088450

[24] S. MOLLIER, M. L. DELLE MONACHE, C. CANUDAS DE WIT, B. SEIBOLD. *Two-dimensional macroscopic model for large scale traffic networks*, in "Transportation Research Part B: Methodological", April 2019, vol. 122, p. 309-326 [*DOI :* 10.1016/J.TRB.2019.02.016], https://hal.archives-ouvertes.fr/hal-01819013

[25] M. U. B. NIAZI, D. DEPLANO, C. CANUDAS DE WIT, A. KIBANGOU. *Scale-free estimation of the average state in large-scale systems*, in "IEEE Control Systems Letters", June 2019, p. 1-6 [*DOI :* 10.1109/LCSYS.2019.2923086], https://hal.archives-ouvertes.fr/hal-02158678

[26] B. OTHMAN, G. DE NUNZIO, D. DI DOMENICO, C. CANUDAS DE WIT. *Ecological traffic management: A review of the modeling and control strategies for improving environmental sustainability of road transportation*, in "Annual Reviews in Control", September 2019, vol. 48, p. 292-311 [*DOI :* 10.1016/J.ARCONTROL.2019.09.003], https://hal.archives-ouvertes.fr/hal-02302722

[27] 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", 2019, vol. 5, n^{o} 2, p. 203-217 [*DOI :* 10.1109/TSIPN.2018.2869117], https://hal.archives-ouvertes.fr/hal-01877623

[28] 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 Part B: Methodological", January 2019, vol. 121, p. 21-40 [*DOI :* 10.1016/J.TRB.2018.12.005], https://hal.archives-ouvertes.fr/hal-01958601

[29] 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", 2019, vol. 6, n^{o} 2, p. 116-129 [*DOI :* 10.1109/TNSE.2018.2792401], https://hal.archives-ouvertes.fr/hal-01683634

[30] R. E. STERN, Y. CHEN, M. CHURCHILL, F. WU, M. L. DELLE MONACHE, B. PICCOLI, B. SEIBOLD, J. SPRINKLE, D. B. WORK. *Quantifying air quality benefits resulting from few autonomous vehicles stabilizing traffic*, in "Transportation Research Part D: Transport and Environment", February 2019, vol. 67, p. 351-365 [*DOI :* 10.1016/J.TRD.2018.12.008], https://hal.inria.fr/hal-02022692

[31] T. M. D. TRAN, A. KIBANGOU. *Collaborative Network Monitoring by Means of Laplacian Spectrum Estimation and Average Consensus*, in "International Journal of Control, Automation and Systems", July 2019, vol. 17, n^{o} 7, p. 1826-1837 [*DOI :* 10.1007/S12555-018-0638-0], https://hal.archives-ouvertes.fr/hal-02166871

[32] L. TUMASH, S. OLMI, E. SCHÖLL.*Stability and control of power grids with diluted network topology*, in "Chaos", December 2019, vol. 29, 123105, https://arxiv.org/abs/1905.13664 - 10 pages, 5 figures [*DOI :* 10.1063/1.5111686], https://hal.inria.fr/hal-02374725

[33] L. TUMASH, E. PANTELEY, A. ZAKHAROVA, E. SCHÖLL.*Synchronization patterns in Stuart–Landau networks: a reduced system approach*, in "European Physical Journal B: Condensed Matter and Complex Systems", May 2019, vol. 92, n° 5 [*DOI :* 10.1140/EPJB/E2019-90483-5], https://hal-centralesupelec.archives-ouvertes.fr/hal-02414119

[34] F. WU, R. E. STERN, S. CUI, M. L. DELLE MONACHE, R. BHADANI, M. BUNTING, M. CHURCHILL, N. HAMILTON, R. HAULCY, B. PICCOLI, B. SEIBOLD, J. SPRINKLE, D. B. WORK.*Tracking vehicle trajectories and fuel rates in phantom traffic jams: Methodology and data*, in "Transportation research. Part C, Emerging technologies", February 2019, vol. 99, p. 82-109 [*DOI :* 10.1016/J.TRC.2018.12.012], https://hal.inria.fr/hal-01614665

[35] J. WU, Z. ZHOU, J. CHEN, H. FOURATI, R. LI.*Corrections to Fast Complementary Filter for Attitude Estimation Using Low-Cost MARG Sensors*, in "IEEE Sensors Journal", December 2019, vol. 19, n° 24, 12511 [*DOI :* 10.1109/JSEN.2019.2939646], https://hal.archives-ouvertes.fr/hal-02297692

[36] 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 [*DOI :* 10.1109/TASE.2018.2888908], https://hal.archives-ouvertes.fr/hal-01958652

[37] 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", July 2019, vol. 21, n° 4, p. 1483-1495 [*DOI :* 10.1002/ASJC.2005], https://hal.inria.fr/hal-01929624

[38] M. ZMITRI, H. FOURATI, N. VUILLERME.*Human Activities and Postures Recognition: From Inertial Measurements to Quaternion-Based Approaches*, in "Sensors", September 2019, vol. 19, n° 19, 4058 [*DOI :* 10.3390/S19194058], https://hal.archives-ouvertes.fr/hal-02294094

Invited Conferences

[39] C. CANUDAS DE WIT.*An Average Study of the Signalized Link-Transmission-Model*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, IEEE, June 2019, p. 3185-3191 [*DOI :* 10.23919/ECC.2019.8795830], https://hal.archives-ouvertes.fr/hal-02268168

[40] M. U. B. NIAZI, C. CANUDAS DE WIT, A. KIBANGOU.*Average observability of large-scale network systems*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, June 2019, p. 1506-1511 [*DOI :* 10.23919/ECC.2019.8795929], https://hal.archives-ouvertes.fr/hal-02073668

[41] M. U. B. NIAZI, X. CHENG, C. CANUDAS DE WIT, J. M. A. SCHERPEN.*Structure-based Clustering Algorithm for Model Reduction of Large-scale Network Systems*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, IEEE, December 2019, https://hal.archives-ouvertes.fr/hal-02278884

International Conferences with Proceedings

[42] M. L. DELLE MONACHE, J. SPRINKLE, R. VASUDEVAN, D. B. WORK.*Autonomous vehicles: From vehicular control to traffic control*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, December 2019, p. 1-28, https://hal.inria.fr/hal-02335658

[43] S. M. FOSSON, F. GARIN, S. GRACY, A. KIBANGOU, D. SWART.*Input and state estimation exploiting input sparsity*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, IEEE, June 2019, p. 2344-2349 [*DOI :* 10.23919/ECC.2019.8795699], https://hal.inria.fr/hal-02094213

[44] F. GARIN, A. KIBANGOU.*Generic Delay-L Left Invertibility of Structured Systems with Scalar Unknown Input*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://hal.inria.fr/hal-02307596

[45] V. GIAMMARINO, M. LV, S. BALDI, P. FRASCA, M. L. DELLE MONACHE.*On a weaker notion of ring stability for mixed traffic with human-driven and autonomous vehicles*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://hal.inria.fr/hal-02335169

[46] S. GRACY, F. GARIN, A. KIBANGOU.*Strongly Structural Input and State Observability for Linear Time Invariant Network Systems*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, IEEE, June 2019, p. 2516-2521 [*DOI :* 10.23919/ECC.2019.8796307], https://hal.inria.fr/hal-02094223

[47] G. GUNTER, Y. WANG, D. GLOUDEMANS, R. E. STERN, D. B. WORK, M. L. DELLE MONACHE, R. BHADANI, M. BUNTING, R. LYSECKY, J. SPRINKLE, B. SEIBOLD, B. PICCOLI.*WiP Abstract: String stability of commercial adaptive cruise control vehicles*, in "ICCPS 2019 - 10th ACM/IEEE International Conference on Cyber-Physical Systems", Montreal, Canada, ICCPS 2019 - Proceedings of the 2019 ACM/IEEE International Conference on Cyber-Physical System, ACM Press, April 2019, p. 328-329 [*DOI :* 10.1145/3302509.3313325], https://hal.inria.fr/hal-02335590

[48] N. MARTIN, P. FRASCA, T. ISHIZAKI, J.-I. IMURA, C. CANUDAS DE WIT.*The price of connectedness in graph partitioning problems*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, European Control Conference, IEEE, June 2019, p. 2313-2318 [*DOI :* 10.23919/ECC.2019.8795723], https://hal.archives-ouvertes.fr/hal-02171099

[49] S. MOLLIER, M. L. DELLE MONACHE, C. CANUDAS DE WIT.*A decision support and planning mobility method for large scale traffic networks*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, IEEE, June 2019, p. 3668-3673 [*DOI :* 10.23919/ECC.2019.8795881], https://hal.archives-ouvertes.fr/hal-02157899

[50] S. MOLLIER, M. L. DELLE MONACHE, C. CANUDAS DE WIT.*A step towards a multidirectional 2D model for large scale traffic networks*, in "TRB 2019 - 98th Annual Meeting Transportation Research Board", Washington D.C., United States, January 2019, p. 1-7, https://hal.archives-ouvertes.fr/hal-01948466

[51] D. NIKITIN, C. CANUDAS DE WIT, P. FRASCA.*Boundary Control for Output Regulation in Scale-Free Positive Networks*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, IEEE Conference on Decision and Control, December 2019, p. 1-6, https://hal.archives-ouvertes.fr/hal-02335142

[52] N. TOKTASSYNOVA, H. FOURATI, B. SULEIMENOV.*Grey Forecasting Model and Particle Swarm based Control of a Phosphorite Sinter Process*, in "ICCAD 2019 - 3rd International Conference on Control, Automation and Diagnosis", Grenoble, France, July 2019, https://hal.archives-ouvertes.fr/hal-02297891

[53] L. TUMASH, C. CANUDAS DE WIT, M. L. DELLE MONACHE.*Robust tracking control design for fluid traffic dynamics*, in "58th IEEE Conference on Decision and Control (CDC 2019)", Nice, France, December 2019, https://hal.archives-ouvertes.fr/hal-02331493

[54] J. WU, Z. ZHOU, M. SONG, H. FOURATI, M. LIU.*Convexity Analysis of Optimization Framework of Attitude Determination from Vector Observations*, in "CoDIT 2019 - 6th International Conference on Control, Decision and Information Technologies", Paris, France, April 2019, https://hal.inria.fr/hal-02432733

[55] M. ZMITRI, H. FOURATI, C. PRIEUR.*Improving Inertial Velocity Estimation Through Magnetic Field Gradient-based Extended Kalman Filter*, in "IPIN 2019 - International Conference on Indoor Positioning and Indoor Navigation", Pisa, Italy, September 2019, p. 1-7, https://hal.archives-ouvertes.fr/hal-02297850

Conferences without Proceedings

[56] T. MOYO, W. MUSAKWA, A. KIBANGOU, T. GUMBO, E. INGWANI.*The potential of using volunteered locational data in planning for smart multi-mobility systems*, in "Real CORP 2019 - 24th International Conference on Urban Planning and Regional Development in the Information Society GeoMultimedia", Karlsruhe, Germany, April 2019, p. 771-779, https://hal.archives-ouvertes.fr/hal-02167817

Scientific Books (or Scientific Book chapters)

[57] M. L. DELLE MONACHE, T. LIARD, A. RAT, R. E. STERN, R. BHADANI, B. SEIBOLD, J. SPRINKLE, D. B. WORK, B. PICCOLI.*Feedback control algorithms for the dissipation of traffic waves with autonomous vehicles*, in "Computational Intelligence and Optimization Methods for Control Engineering", Springer, 2019, p. 275-299 [*DOI :* 10.1007/978-3-030-25446-9_12], https://hal.inria.fr/hal-01930724

Books or Proceedings Editing

[58] C. GAO, G. ZHAO, H. FOURATI (editors). *Cooperative Localization and Navigation: Theory, Research, and Practice*, CRC Press, August 2019, p. 1-660 [*DOI :* 10.1201/9780429507229], https://hal.archives-ouvertes.fr/hal-02295010

Other Publications

[59] G. GUNTER, D. GLOUDEMANS, R. E. STERN, S. T. MCQUADE, R. BHADANI, M. BUNTING, M. L. DELLE MONACHE, R. LYSECKY, B. SEIBOLD, J. SPRINKLE, B. PICCOLI, D. B. WORK.*Are commercially implemented adaptive cruise control systems string stable?*, October 2019, https://arxiv.org/abs/1905.02108 - working paper or preprint, https://hal.inria.fr/hal-02332052

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]

Faculty Members
Emmanuel Grenier [Team leader, École Normale Supérieure de Lyon, Professor, HDR]
Arthur Marly [École Normale Supérieure de Lyon, Associate Professor]
Paul Vigneaux [École Normale Supérieure de Lyon, Associate Professor, HDR]

Technical Staff
David Coulette [CNRS, Engineer]

Administrative Assistant
Sylvie Boyer [Inria, Administrative Assistant]

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 one main axe of applications, namely cancer modeling 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 bacterias.

  Finally, H. Leman studied the large time behavior of continuous state branching processes with competition and Lévy environment. These kind of stochastic processes are used to represent the fluctuations of the size of a population. In particular, she studied the extinction time of such a process.

- 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 Herschel-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 (see figure 1), 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 [cf joint Figure and Fernandez-Nieto et al. JCP 2018]. Currently, through a collaboration with IRSTEA Grenoble, we also revisit the theory of viscoplastic boundary layers (see figure (2) by extending the Oldroyd's asymptotic scaling (1947) to the cases of moderate Bingham numbers (or Herschel-Bulkley numbers). Also with IRSTEA, we are developping a joint study (numerical and experimental) of viscoplastic avalanches in the lab, to challenge various yield stress models.

Figure 1. An example of avalanche simulation

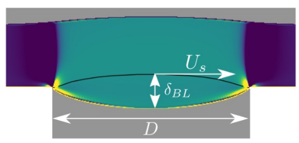


Figure 2. An example of boundary layer for complex flows

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).
- Comparison between numerical simulations and physical experiments for the dam-break of viscoplastic materials: collaboration with IRSTEA (now INRAE, since Jan. 2020).

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

INSERM / Plan Cancer 2019 - 2022: Evolutionary Mechanisms of Metabolic Adaptation and Scheduling of Therapy in ONcology (250 kE).

Project: This project combines mathematical models integrating heterogeneous phenotypic and genetic data with multiple in vitro models of cancer evolution. Triple Negative Breast Cancers (TNBC) are unsuited to targeted therapy and display high diversity and resistance. We will thus use 3 existing TNBC models, of common origin but subjected to different tumor initiating oncogenic insults, treated over several generations with two drugs targeting antagonist receptors involved in metabolism. By following phenotypic and genetic properties over time, we aim to uncover and quantify how distinct tumor initiation contexts shape evolutionary trajectories and the emergence of resistance. Using mathematical models and simulations, we will investigate how to optimise therapeutic regimens based on the intrinsic evolutionary properties of each model, before validating our predictions in vivo via murine xenografts. Results: The results of this project will help better characterize the influence of the initiating genetic alterations on the ensuing dynamics of development and resistance in TNBC. It will also pave the way to optimise novel therapeutic strategies aiming to leverage cell metabolism to control tumor evolution in the clinic.

5.2. International Research Visitors

5.2.1. Visits to International Teams

5.2.1.1. Research Stays Abroad

Paul Vigneaux spend one year at UCB (University British Columbia)

6. Dissemination

6.1. Promoting Scientific Activities

6.1.1. Scientific Events: Organisation

6.1.1.1. General Chair, Scientific Chair

- E. Grenier: member of the Scientific Board of the CLARA (Regional Cancer organization)
- H. Leman: Inria Bio (october 2019)
- H. Leman: "mating preferencies" symposium in the international conference MMEE (june 2019)
- P. Vigneaux: Scientific co-head of the national CNRS research group GdR EGRIN (with Emmanuel Audusse).

6.2. Teaching - Supervision - Juries

6.2.1. Teaching

- E. Grenier: L3 (integration theory), M1 (PDEs) and "agregation" (modeling).
- Paul Vigneaux: "agregation" (modeling), "computational sciences, an introduction to modelling" (for L3 - M2 students from physics, computer science and biology)
- Paul Vigneaux: Member of the Board of MILYON, the Laboratory of Excellence (Labex) in Mathematics of Lyon (Since September 2011). This Labex aims at federating international research, higher education and society activities. In charge of evaluation and grant attribution for foreign students for M1, M2 and PhD in Lyon, since 2011.

6.3. Popularization

Emmanuel Grenier: "Mathematics and vaccines" (Université de Chambéry, for first year students).

6.3.1. Interventions

- H. Leman: Mentoring of two high school students
- D. Coulette: Scientific speed dating at "Musée des Confluences" (Lyon) for the "Fête de la Science" and $80th$ birthday of CNRS.

7. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] V. CALVEZ.*Chemotactic waves of bacteria at the mesoscale*, in "Journal of the European Mathematical Society", 2019, vol. 22, n^{o} 2, p. 593–668 [*DOI :* 10.4171/JEMS/929], https://hal.archives-ouvertes.fr/hal-01340375

Invited Conferences

[2] P. VIGNEAUX.*Flow of a Yield-stress Fluid over a Cavity: Experimental and Numerical Investigation of an Oldroyd's Boundary Layer*, in "CSE 2019 - SIAM Conference on Computational Science and Engineering", Spokane, United States, February 2019, https://hal.archives-ouvertes.fr/hal-02087288

Conferences without Proceedings

[3] P. VIGNEAUX, G. CHAMBON, L.-H. LUU, A. MARLY, P. PHILIPPE.*Flow of a yield stress-fluid over cavity and viscoplastic boundary layers: following the quest*, in "VPF8 2019 - 8th Viscoplastic Fluids Workshop: From Theory to Applications", Cambridge, United Kingdom, September 2019, https://hal.archives-ouvertes.fr/hal-02296110

[4] P. VIGNEAUX.*Numerical methods for geophysical yield stress fluid flows*, in "ICIAM 2019 - 9th International Congress on Industrial and Applied Mathematics", Valencia, Spain, July 2019, https://hal.archives-ouvertes.fr/hal-02202001

Other Publications

[5] E. BOUIN, V. CALVEZ, E. GRENIER, G. NADIN.*Large deviations for velocity-jump processes and non-local Hamilton-Jacobi equations*, September 2019, https://arxiv.org/abs/1607.03676 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01344939

[6] C. CORON, M. COSTA, F. LAROCHE, H. LEMAN, C. SMADI.*Emergence of homogamy in a two-loci stochastic population model*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02043536

[7] H. LEMAN, J. C. PARDO MILLAN.*Extinction and coming down from infinity of CB-processes with competition in a Lévy environment*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01694581

[8] H. LEMAN, J. C. PARDO MILLAN.*Extinction time of logistic branching processes in a Brownian environment*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02306075

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]

Faculty Member

Laurent Girin [Institut polytechnique de Grenoble, Professor, HDR]

Technical Staff

Soraya Arias [Inria, Engineer]
Alex Auternaud [Inria, Engineer, from Nov 2019]
Bastien Mourgue [Inria, Engineer, until May 2019]
Guillaume Sarrazin [Inria, Engineer, until May 2019]

PhD Students

Anand Ballou [Univ Grenoble Alpes, PhD Student, from Nov 2019]
Yutong Ban [Inria, PhD Student, until May 2019]
Xiaoyu Bie [Univ Grenoble Alpes, PhD Student, from Dec 2019]
Guillaume Delorme [Inria, PhD Student]
Wen Guo [Univ Grenoble Alpes, PhD Student, from Oct 2019]
Louis Airale [Univ Grenoble Alpes, PhD Student, from Oct 2019]
Sylvain Guy [Univ Grenoble Alpes, PhD Student]
Yihong Xu [Inria, PhD Student]

Post-Doctoral Fellows

Simon Leglaive [Inria, Post-Doctoral Fellow, until Aug 2019]
Mostafa Sadeghi [Inria, Post-Doctoral Fellow]

Administrative Assistant

Nathalie Gillot [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Audio-Visual Machine Perception

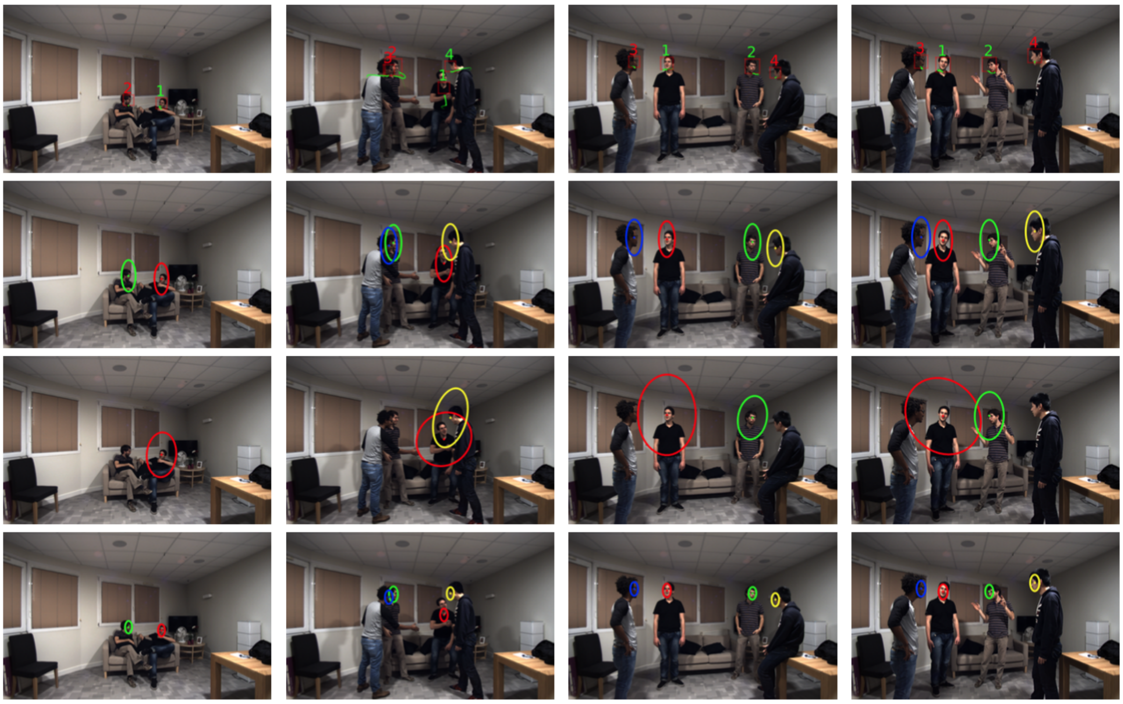


Figure 1. This figure illustrates the audio-visual multiple-person tracking that has been developed by the team [38], [41]. The tracker is based on variational inference [5] and on supervised sound-source localization [10], [29]. Each person is identified with a digit. Green digits denote active speakers 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 [25], [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 [9]. 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 [8]. 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 [19], [32]. 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 [20]. 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 [12].

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 [8] and audio-visual learning [10].

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 [35]. Subsequently, we proposed to combine 3D representations of shape (meshes) with photometric data [33]. 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 [34]. 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 [21], [17],[16]. 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 [12].

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 [26]. 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 [24], [23]. 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 [7]. 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

4.1.1. IEEE Senior Member.

Xavier Alameda-Pineda has become an IEEE Senior Member on February 1st, 2019. The grade of Senior Member requires experience reflecting professional maturity as an engineer, scientist, educator, technical executive, or originator in IEEE-designated fields for a total of 10 years and have demonstrated 5 years of significant performance.

4.1.2. H2020 Project SPRING

(1 Januray 2020 – 31 December 2023) is a research and innovation action (RIA) with eight partners: Inria Grenoble (coordinator), Università degli Studi di Trento, Czech Technical University Prague, Heriot-Watt University Edinburgh, Bar-Ilan University Tel Aviv, ERM Automatismes Industriels Carpentras, PAL Robotics Barcelona, and Hôpital Broca Paris.. The main objective of SPRING (Socially Pertinent Robots in Gerontological Healthcare) is the development of socially assistive robots with the capacity of performing multimodal multiple-person interaction and open-domain dialogue. In more detail:

- The scientific objective of SPRING is to develop a novel paradigm and novel concept of socially-aware robots, and to conceive innovative methods and algorithms for computer vision, audio processing, sensor-based control, and spoken dialog systems based on modern statistical- and deep-learning to ground the required social robot skills.
- The technological objective of SPRING is to create and launch a brand new generation of robots that are flexible enough to adapt to the needs of the users, and not the other way around.
- The experimental objective of SPRING is twofold: to validate the technology based on HRI experiments in a gerontology hospital, and to assess its acceptability by patients and medical staff.

Website: https://spring-h2020.eu/

4.1.3. ANR JCJC Project ML3RI

(1 March 2020 – 28 February 2024) has been awarded to Xavier Alameda-Pineda. Multi-person robot interaction in the wild (i.e. unconstrained and using only the robot's resources) is nowadays unachievable because of the lack of suitable machine perception and decision-taking models. *Multi-Modal Multi-person Low-Level Learning models for Robot Interaction* (ML3RI) has the ambition to develop the capacity to understand and react to low-level behavioral cues, which is crucial for autonomous robot communication. The main scientific impact of ML3RI is to develop new learning methods and algorithms, thus opening the door to study multi-party conversations with robots. In addition, the project supports open and reproducible research.

4.1.4. MIAI Chair.

The Multidisciplinary Institue in Artificial Intelligence (MIAI) is one of the four AI French institutes launched in 2019 by the French government. MIAI is structured around several chairs, each chair gathering 3-6 researchers as well as postdocs and PhD students. Team members Radu Horaud and Xavier Alameda-Pineda are co-chairs of the *Audio-visual machine perception and interaction for companion robots* chair. The development of methods and algorithms for enabling socially-aware robot behavior with the specific goal of interacting with humans is the core topic. The emphasis is put on unsupervised and weakly supervised learning with audio and visual data, based on Bayesian methods, deep learning and reinforcement learning. It is planned to develop challenging proof-of-concept implementations and demonstrators.

5. New Software and Platforms

5.1. 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 Radu Horaud
- Contact: Radu Horaud
- URL: https://team.inria.fr/perception/research/naolab/

5.2. Associations of Audio Cues with 3D locations library

FUNCTIONAL DESCRIPTION: Library to associate some auditory cues with 3D locations (points). It provides an estimation of the emitting state of each of the input locations. There are two main assumptions : 1 - The 3D locations are valid during the acquisition interval related to the audio cues 2 - The 3D locations are the only possible locations for the sound sources, no new locations will be created in this module

The software provides also a multimodal fusion library

- Participants: Antoine Deleforge, Jordi Sanchez-Riera, Radu Horaud and Xavier Alameda-pineda
- Contact: Radu Horaud

5.3. Audio Cue Extractor Library

FUNCTIONAL DESCRIPTION: This module extracts auditory cues from the raw audio streams. The interaural time difference (ITD) is estimated using cross-correlation methods.

- Participants: Antoine Deleforge, Radu Horaud and Soraya Arias
- Contact: Soraya Arias

5.4. Audiovisual Robots and Heads

FUNCTIONAL DESCRIPTION: The team has developed two audiovisual (AV) robot heads: the POPEYE head and the NAO stereo head. Both are equipped with a binocular vision system and with four microphones. The software modules comprise stereo matching and reconstruction, sound-source localization and audio-visual fusion. POPEYE has been developed within the European project POP in collaboration with the project-team MISTIS and with two other POP partners: the Speech and Hearing group of the University of Sheffield and the Institute for Systems and Robotics of the University of Coimbra. The NAO stereo head was developed under the European project HUMAVIPS in collaboration with Aldebaran Robotics (which manufactures the humanoid robot NAO) and with the University of Bielefeld, the Czech Technical Institute, and IDIAP. The software modules that we develop are compatible with both these robot heads.

- Contact: Radu Horaud
- URL: https://team.inria.fr/perception/popeye/

5.5. GLLiM

Gaussian Locally Linear Mapping

KEYWORDS: Regression - Machine learning - Gaussian mixture

SCIENTIFIC DESCRIPTION: GLLiM is a flexible tool for probabilistic non-linear regression using Gaussian mixtures. Using an inverse regression strategy with a reduced number of parameters, it is particularly suited for high- to low-dimensional regression tasks. It also enables the modeling of additional unobserved non-linear effects on input data. The method was published in [Deleforge et al., IJNS 2015]. The toolbox include an example of application to head pose estimation from synthetic images.

- Participant: Antoine Deleforge
- Contact: Antoine Deleforge
- Publication: hal-00863468, version 3
- URL: https://team.inria.fr/perception/gllim_toolbox/

5.6. Litbot

Live together with robots

KEYWORDS: Speaker Localization - Audio tracking - Visual tracking - NAO Robot - Computer vision - Signal processing

SCIENTIFIC DESCRIPTION: Litbot stands for "Live together with robots". This library aims to provide algorithms and associated software packages to perform audio-visual speaker localization and tracking with a consumer robot (in particular a NAO robot). The scope of this project is two-fold. The first is to develop the robust speaker localization and tracking algorithm in the presence of other audio-visual sources like TV. The second is to modify or optimize the original algorithm to be fit into real-time system. This library benefits from the work done with Online Multiple Sound-Source Localization package developed by X. Li.

FUNCTIONAL DESCRIPTION: This project develops algorithms and associated software packages to perform audio-visual speaker localization and tracking with a consumer robot. This version of the litbot library provides new functions to integrate the Samsung robotic platform to handle ROS middleware (robotic defacto standard) and modifies and optimizes tracking and audio localization processes (better handling of the residual noise signals, performance improved to match real time).

- Participants: Xiaofei Li, Yutong Ban, Soraya Arias, Radu Horaud, Guillaume Sarrazin and Bastien Mourgue
- Contact: Radu Horaud

5.7. Online Multiple Sound-Source Localization

KEYWORDS: Audio signal processing - Multiple sound-source localization - Matlab - Direct-path RTF

FUNCTIONAL DESCRIPTION: This project tackles multiple sound-source localization in noisy and reverberant environments, using binaural recordings of an acoustic scene. It provides Matlab routines to estimate multiple sound source (such as speakers) locations based on direct-path relative transfer function (DP-RTF) estimation.

- Participants: Xiaofei Li and Radu Horaud
- Contact: Radu Horaud
- Publications: Multiple-Speaker Localization Based on Direct-Path Features and Likelihood Maximization with Spatial Sparsity Regularization - Estimation of the Direct-Path Relative Transfer Function for Supervised Sound-Source Localization

5.8. RMP

RoMPers

KEYWORDS: Middleware - Robotics - NAO Robot

SCIENTIFIC DESCRIPTION: Robot Middleware developed by Perception. It follows the development done on RobotHandler and NAOLab. Its goal is to provide an abstraction which allows an easy access to robot sensors. In the same time, this high level access is independant of the robot.

FUNCTIONAL DESCRIPTION: Robot Middleware developed by Perception. It follows the development done on RobotHandler and NAOLab. Its goal is to provide an abstraction which allows an easy access to robot sensors. In the same time, this high level access is independant of the robot. And it also provides tools for sensor calibration (audio, video), video annotation, etc

- Participant: Guillaume Sarrazin
- Contact: Soraya Arias

5.9. SE-VAE-alpha-stable

KEYWORDS: Audio signal processing - Speech processing - Deep learning - Neural networks

FUNCTIONAL DESCRIPTION: This software provides an iterative algorithm for enhancing a speech signal in a noisy monophonic recording. The algorithm is detailed in the following paper: "Speech enhancement with variational autoencoders and alpha-stable distributions" Simon Leglaive, Umut Simsekli, Antoine Liutkus, Laurent Girin, Radu Horaud IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), Brighton, UK, May 2019

- Contact: Simon Leglaive

5.10. Sound recognition library

FUNCTIONAL DESCRIPTION: This recognition module is based on supervised learning.

- Participants: Maxime Janvier and Radu Horaud
- Contact: Radu Horaud

5.11. SE-VAE-NMF

KEYWORDS: Audio signal processing - Speech processing - Deep learning - Neural networks

FUNCTIONAL DESCRIPTION: This software provides an iterative algorithm for enhancing a speech signal in a noisy monophonic recording. The algorithm is detailed in the following paper: "A variance modeling framework based on variational autoencoders for speech enhancement" Simon Leglaive, Laurent Girin, Radu Horaud Proc. of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP), Aalborg, Denmark, September 2018

- Contact: Simon Leglaive
- Publication: hal-01832826v1

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 [43], [44]. 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 Denoising and Enhancement with LTSMs

We have started to address the problems of multichannel speech denoising [45] and enhancement [51] in the short-time Fourier transform (STFT) domain and in the framework of sequence-to-sequence deep learning. In the case of denoising, the magnitude of noisy speech is mapped onto the noise power spectral density. In the case of speech enhancement, the noisy speech is mapped onto clean speech. A long short-time memory (LSTM) network takes as input a sequence of STFT coefficients associated with a frequency bin of multichannel noisy-speech signals. The network's output is a sequence of single-channel cleaned speech at the same frequency bin. We propose several clean-speech network targets, namely, the magnitude ratio mask, the complex ideal ratio mask, the STFT coefficients and spatial filtering [54]. A prominent feature of the proposed model is that the same LSTM architecture, with identical parameters, is trained across frequency bins. The proposed method is referred to as narrow-band deep filtering. This choice stays in contrast with traditional wide-band speech enhancement methods. The proposed deep filter is able to discriminate between speech and noise by exploiting their different temporal and spatial characteristics: speech is non-stationary and spatially coherent while noise is relatively stationary and weakly correlated across channels. This is similar in spirit with unsupervised techniques, such as spectral subtraction and beamforming. We describe extensive experiments with both mixed signals (noise is added to clean speech) and real signals (live recordings). We empirically evaluate the proposed architecture variants using speech enhancement and speech recognition metrics, and we compare our results with the results obtained with several state of the art methods. In the light of these experiments we conclude that narrow-band deep filtering has very good performance, and excellent generalization capabilities in terms of speaker variability and noise type, e.g. Figure 2.

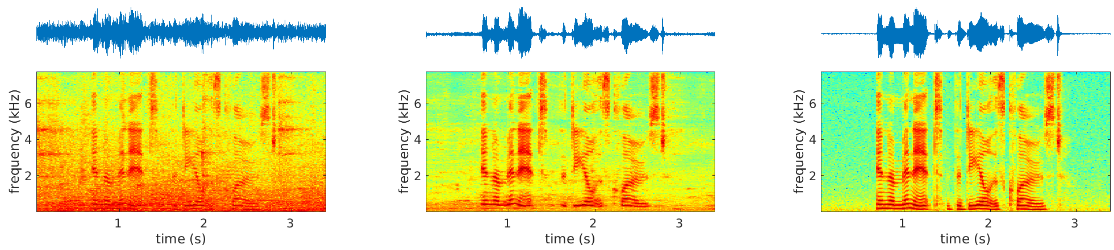


Figure 2. An example of narrow-band deep filtering for speech enhancement [54]. Waveforms and spectrograms of the noisy (unprocessed) input, the filtered output and the ground-truth clean-speech. Four microphones were used in this example. The signal-to-noise ratio in this example is 0 dB.

Website: https://team.inria.fr/perception/research/mse-lstm/.

6.3. Multichannel Speech Enhancement with Variational Auto-Encoder

We addressed speaker-independent multichannel speech enhancement in unknown noisy environments. Our work is based on a well-established multichannel local Gaussian modeling framework. We propose to use a neural network for modeling the speech spectro-temporal content. The parameters of this supervised model are learned using the framework of variational autoencoders. The noisy recording environment is supposed to be unknown, so the noise spectro-temporal modeling remains unsupervised and is based on non-negative matrix factorization (NMF). We develop a Monte Carlo expectation-maximization algorithm and we experimentally show that the proposed approach outperforms its NMF-based counterpart, where speech is modeled using supervised NMF [49].
Website: https://team.inria.fr/perception/research/icassp-2019-mvae/

6.4. Audio-visual Speech Enhancement with Conditional Variational Auto-Encoder

Variational auto-encoders (VAEs) are deep generative latent variable models that can be used for learning the distribution of complex data. VAEs have been successfully used to learn a probabilistic prior over speech signals, which is then used to perform speech enhancement. One advantage of this generative approach is that it does not require pairs of clean and noisy speech signals at training. In this work, we propose audio-visual variants of VAEs for single-channel and speaker-independent speech enhancement. We developed a conditional VAE (CVAE) where the audio speech generative process is conditioned on visual information of the lip region, e.g. Figure 3. At test time, the audio-visual speech generative model is combined with a noise model, based on nonnegative matrix factorization, and speech enhancement relies on a Monte Carlo expectation-maximization algorithm. Experiments were conducted with the recently published NTCD-TIMIT dataset. The results confirm that the proposed audio-visual CVAE effectively fuse audio and visual information, and it improves the speech enhancement performance compared with the audio-only VAE model, especially when the speech signal is highly corrupted by noise. We also showed that the proposed unsupervised audio-visual speech enhancement approach outperforms a state-of-the-art supervised deep learning method [55].
Website: https://team.inria.fr/perception/research/av-vae-se/

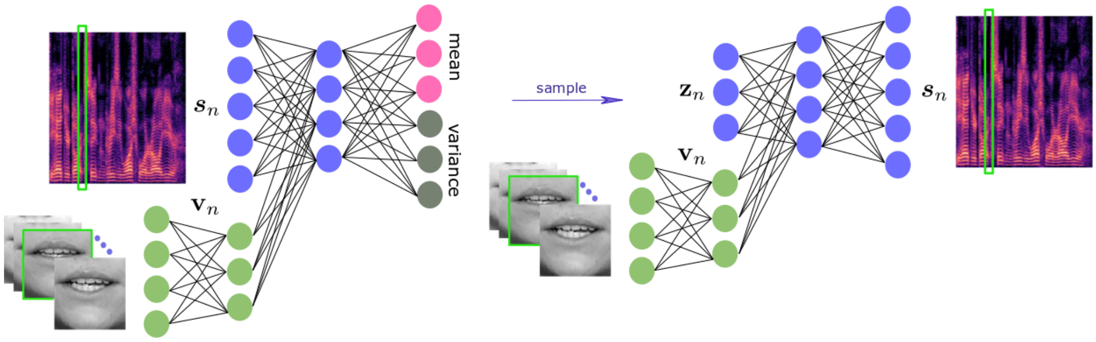


Figure 3. We proposed a conditional variational auto-encoder architecture for fusing audio and visual data for speech enhancement [55].

6.5. Variational Bayesian Inference of Audio-visual Speaker Tracking

We addressed the problem of tracking multiple speakers via the fusion of visual and auditory information [36]. 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 proposed a variational inference model which amounts to approximate the joint distribution with a factorized distribution. The solution takes the form of closed-form expectation maximization procedures using Gaussian distributions [38]. We described in detail the inference algorithm, we evaluated its performance and we compared the results with several baseline methods. These experiments show that the proposed audio-visual tracker performs well in informal meetings involving a time-varying number of people. Real-time versions of the algorithm were implemented on our robotic platform [47].
Website: https://team.inria.fr/perception/research/var-av-track/.

6.6. Detection, Localization and Tracking of Multiple Audio Sources

We addressed the problem of online detection, localization and tracking of multiple moving speakers in reverberant environments [36]. The work has the following contributions. We used the direct-path relative transfer function (DP-RTF), an inter-channel feature that encodes acoustic information robust against reverberation, and we proposed an online algorithm well suited for estimating DP-RTFs associated with moving audio sources. Another crucial ingredient of the proposed method is its ability to properly assign DP-RTFs to audio-source directions. Towards this goal, we adopted a maximum-likelihood formulation and we proposed to use the exponentiated gradient (EG) to efficiently update source-direction estimates starting from their currently available values. The problem of multiple-speaker tracking is computationally intractable because the number of possible associations between observed source directions and physical speakers grows exponentially with time. We adopt a Bayesian framework and we proposed two variational approximations of the posterior filtering distributions associated with multiple speaker tracking, as well as two efficient variational expectation maximization (VEM) solvers [41], [37]. The proposed online localization and tracking methods were thoroughly evaluated using two datasets that contain recordings performed in real environments.
Websites:

https://team.inria.fr/perception/research/audiotrack-vonm/
https://team.inria.fr/perception/research/multi-speaker-tracking/.

6.7. The Kinovis Multiple-Speaker Tracking Datasets

The Kinovis multiple speaker tracking (Kinovis-MST) datasets contain live acoustic recordings of multiple moving speakers in a reverberant environment. The data were recorded in the Kinovis multiple-camera laboratory at Inria Grenoble Rhône-Alpes. The room size is $10.2 \times 9.9 \times 5.6$ meters with T60 = 0.53 seconds. The data were recorded with four microphones embedded into the head of a NAO robot. Because there is a fan located inside the robot head nearby the microphones, there is a fair amount of stationary and spatially correlated microphone noise. The signal-to-noise ratio of the microphone signals is of approximately 2.7 dB. The recordings contain between one and three moving participants that speak naturally, hence the number of active speech sources varies over time. The robot-to-speaker distance ranges between 1.5 and 3.5 meters. Ground-truth trajectories and speech activity information were obtained in the following way. Participants were wearing optical markers placed on their heads such that the Kinovis motion capture system provides accurate 3D trajectories for each participant. Moreover, an infrared marker is placed on the participants' foreheads. This enables the identification of each participant over time. Whenever time a participant is silent, he/she hides his/her infrared marker, thus allowing speaking/silent annotations of the recordings.
Website: https://team.inria.fr/perception/the-kinovis-mst-dataset/.

6.8. Deep Regression

Deep learning revolutionized data science, and recently its popularity has grown exponentially, as did the amount of papers employing deep networks. Vision tasks, such as human pose estimation, did not escape from this trend. There is a large number of deep models, where small changes in the network architecture, or in the data pre-processing, together with the stochastic nature of the optimization procedures, produce notably different results, making extremely difficult to sift methods that significantly outperform others. This situation motivates the current study, in which we perform a systematic evaluation and statistical analysis of vanilla deep regression, i.e. convolutional neural networks with a linear regression top layer. This is the first comprehensive analysis of deep regression techniques. We perform experiments on four vision problems, and report confidence intervals for the median performance as well as the statistical significance of the results, if any. Surprisingly, the variability due to different data pre-processing procedures generally eclipses the variability due to modifications in the network architecture. Our results reinforce the hypothesis according to which, in general, a general-purpose network (e.g. VGG-16 or ResNet-50) adequately tuned can yield results close to the state-of-the-art without having to resort to more complex and ad-hoc regression models, [40].
Website: https://team.inria.fr/perception/research/deep-regression/.

6.9. Deep Reinforcement Learning for Audio-Visual Robot Control

More recently, we investigated the use of reinforcement learning (RL) as an alternative to sensor-based robot control. 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. 4. 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 [39] we proposed the use of a simulated environment for pre-training the RL parameters, thus avoiding spending hours of tedious interaction.
Website: https://team.inria.fr/perception/research/deep-rl-for-gaze-control/.

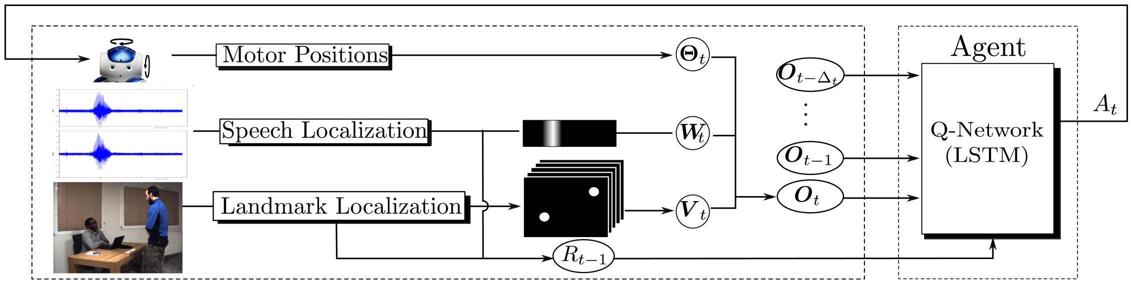


Figure 4. Overview of the proposed deep RL method for controlling the gaze of a robot. At each time index t, audio and visual data are represented as binary maps which, together with motor positions, form the set of observations $\mathbf{O}_t$. A motor action A_t (rotate the head left, right, up, down, 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 [39] for further details.

7. Partnerships and Cooperations

7.1. European Initiatives

7.1.1. Collaborations with Major European Organizations

Universitat Politècnica de Catalunya (UPC), Spain

Physical complex Interactions and Multi-person Pose Estimation (PIMPE) is three year project financed by IDEX. The scientific challenges of PIMPE are the followings: (i) Modeling multi-person interactions in full-body pose estimation, (ii) Estimating human poses in complex multi-person physical interactions, and (iii) Generating controlled and realistic multi-person complex pose images.

7.2. International Research Visitors

7.2.1. Research Stays Abroad

Xavier Alameda-Pineda spent three months at the University of Verona, Italy.

Yihong Xu (Ph.D. student) spent three months at the Technical University Munich, Germany.

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

- 10th International Workshop on Human Behavior Understanding, in conjunction with the 2019 International Conference on Computer Vision (ICCV).
- 1st Workshop on Fairness, Accountability and Transparency in Multimedia, in conjunction with the 2019 ACM International Conference on Multimedia (ACM MM).
- 3rd Workshop on Understanding Subjective Properties of Data, Focus on Fashion and Subjective Search, in conjunction with the 2019 International Conference on Computer Vision and Pattern Recognition (CVPR).

8.1.2. Scientific Events: Selection

8.1.2.1. Member of the Conference Program Committees

- Area Chair of the 2019 ACM International Conference on Multimedia (ACM MM).
- Area Chair of the 2019 International Conference on Image Analysis and Processing (ICIAP).

8.1.2.2. Reviewer

Xavier Alameda-Pineda: CVPR 2019, NeurIPS 2019, ICCV 2019, ACII 2019, ICLR 2019, ICML 2019, ICIP 2019.

8.1.3. Journal

8.1.3.1. Member of the Editorial Boards

Xavier Alameda-Pineda is Associated Editor of the ACM Transactions on Multimedia Computing Communications and Applications

8.1.3.2. Reviewer - Reviewing Activities

Xavier Alameda-Pineda: TPAMI, TASLP, TMM.

8.1.4. Invited Talks

Xavier Alameda-Pineda:

- Significancy & Robustness in Deep Regression (July 2019) at University of Trento
- Probabilistic and deep methods for human behavior understanding (July 2019) at Media Integration and Communication Center

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master : Xavier Alameda-Pineda, Fundamentals of Probabilistic Data Mining, 18h, M2, UGA, France.

Master : Xavier Alameda-Pineda, Category Learning and Object Recognition, 11h, M2, UGA, France.

Master : Xavier Alameda-Pineda, Advanced Learning Models, 13.5h, M2, UGA, France.

8.2.2. Supervision

PhD: Yutong Ban, Audio-visual multiple-speaker tracking for robot perception [36], Université Grenoble Alpes, May 2019, Xavier Alameda-Pineda and Radu Horaud,

PhD in progress: Guillaume Delorme, Deep Person Re-identification, October 2017, Xavier Alameda-Pineda and Radu Horaud,

PhD in progress: Yihong Xu, Deep Multiple-person Tracking, October 2018, Xavier Alameda-Pineda and Radu Horaud,

PhD in progress: Wen Guo, Deep Human Pose, October 2019, Xavier Alameda-Pineda and Radu Horaud,

PhD in progress: Anand Ballou, Deep Reinforcement Learning for Robot Control, November 2019, Xavier Alameda-Pineda and Radu Horaud,

PhD in progress: Louis Airale, Data Generation for Deep Multimodal Interaction Algorithms, October 2019, Xavier Alameda-Pineda and Dominique Vaufreydaz,

PhD in progress: Xiaoyu Bie, Deep Generative Methods for Audio and Vision, December 2019, Xavier Alameda-Pineda and Laurent Girin.

8.2.3. Juries

Xavier Alameda-Pineda belonged to the following PhD Juries as “examinateur”:

- Yutong Ban (University Grenoble-Alpes)
- Irtiza Hasan (University of Verona)
- Theodoros Tsesmelis (University of Verona)

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] 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

[4] 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

[5] 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

[6] 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

[7] 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

[8] 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

[9] 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

[10] 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

[11] 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

[12] 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

[13] 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^o 6, p. 1397–1410, https://arxiv.org/abs/1609.01466 [*DOI :* 10.1109/TPAMI.2017.2717829], https://hal.inria.fr/hal-01413414

[14] 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

[15] 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^o 5, p. 1086–1099, https://arxiv.org/abs/1603.09725 [*DOI :* 10.1109/TPAMI.2017.2648793], https://hal.inria.fr/hal-01413403

[16] 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

[17] 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

[18] 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

[19] 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

[20] 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

[21] 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

[22] 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

[23] 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

[24] 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

[25] 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

[26] 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

[27] 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

[28] 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

[29] 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

[30] 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

[31] 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

[32] 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

[33] 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

[34] 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

[35] 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

[36] Y. Ban.*Audio-visual multiple-speaker tracking for robot perception*, Université Grenoble Alpes, May 2019, https://tel.archives-ouvertes.fr/tel-02163418

Articles in International Peer-Reviewed Journal

[37] Y. Ban, X. Alameda-Pineda, C. Evers, R. Horaud.*Tracking Multiple Audio Sources with the Von Mises Distribution and Variational EM*, in "IEEE Signal Processing Letters", June 2019, vol. 26, n° 6, p. 798 - 802, https://arxiv.org/abs/1812.08246 [*DOI :* 10.1109/LSP.2019.2908376], https://hal.inria.fr/hal-01969050

[38] Y. Ban, X. Alameda-Pineda, L. Girin, R. Horaud.*Variational Bayesian Inference for Audio-Visual Tracking of Multiple Speakers*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", November 2019, vol. 42, p. 1-17, https://arxiv.org/abs/1809.10961 [*DOI :* 10.1109/TPAMI.2019.2953020], https://hal.inria.fr/hal-01950866

[39] 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", February 2019, vol. 118, p. 61-71, https://arxiv.org/abs/1711.06834 [*DOI :* 10.1016/j.patrec.2018.05.023], https://hal.inria.fr/hal-01643775

[40] S. Lathuilière, P. Mesejo, X. Alameda-Pineda, R. Horaud.*A Comprehensive Analysis of Deep Regression*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", April 2019, vol. 41, p. 1-17, https://arxiv.org/abs/1803.08450 [*DOI :* 10.1109/TPAMI.2019.2910523], https://hal.inria.fr/hal-01754839

[41] X. Li, Y. Ban, L. Girin, X. Alameda-Pineda, R. Horaud.*Online Localization and Tracking of Multiple Moving Speakers in Reverberant Environments*, in "IEEE Journal of Selected Topics in Signal Processing", March 2019, vol. 13, n° 1, p. 88-103, https://arxiv.org/abs/1809.10936 [*DOI :* 10.1109/JSTSP.2019.2903472], https://hal.inria.fr/hal-01851985

[42] X. Li, L. Girin, S. Gannot, R. Horaud.*Multichannel Online Dereverberation based on Spectral Magnitude Inverse Filtering*, in "IEEE/ACM Transactions on Audio, Speech and Language Processing", May 2019, vol. 27, n° 9, p. 1365-1377, https://arxiv.org/abs/1812.08471 [*DOI :* 10.1109/TASLP.2019.2919183], https://hal.inria.fr/hal-01969041

[43] 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", March 2019, vol. 27, n° 3, p. 645-659, https://arxiv.org/abs/1711.07911 [*DOI :* 10.1109/TASLP.2019.2892412], https://hal.inria.fr/hal-01799809

[44] X. Li, L. Girin, R. Horaud.*Expectation-Maximization for Speech Source Separation using Convolutive Transfer Function*, in "CAAI Transactions on Intelligent Technologies", March 2019, vol. 4, n° 1, p. 47 - 53 [*DOI :* 10.1049/trit.2018.1061], https://hal.inria.fr/hal-01982250

[45] X. LI, S. LEGLAIVE, L. GIRIN, R. HORAUD.*Audio-noise Power Spectral Density Estimation Using Long Short-term Memory*, in "IEEE Signal Processing Letters", June 2019, vol. 26, n^o 6, p. 918-922, https://arxiv.org/abs/1904.05166 [*DOI :* 10.1109/LSP.2019.2911879], https://hal.inria.fr/hal-02100059

Invited Conferences

[46] F. FORBES, A. DELEFORGE, R. HORAUD, E. PERTHAME.*Robust non-linear regression approach for generalized inverse problems in a high dimensional setting*, in "AIP 2019 - Applied Inverse Problem conference", Grenoble, France, July 2019, https://hal.archives-ouvertes.fr/hal-02415115

International Conferences with Proceedings

[47] X. ALAMEDA-PINEDA, S. ARIAS, Y. BAN, G. DELORME, L. GIRIN, R. HORAUD, X. LI, B. MOURGUE, G. SARRAZIN.*Audio-Visual Variational Fusion for Multi-Person Tracking with Robots*, in "ACMMM 2019 - 27th ACM International Conference on Multimedia", Nice, France, ACM Press, October 2019, p. 1059-1061 [*DOI :* 10.1145/3343031.3350590], https://hal.inria.fr/hal-02354514

[48] L. GIRIN, F. ROCHE, T. HUEBER, S. LEGLAIVE.*Notes on the use of variational autoencoders for speech and audio spectrogram modeling*, in "DAFx 2019 - 22nd International Conference on Digital Audio Effects", Birmingham, United Kingdom, February 2019, p. 1-8, https://hal.archives-ouvertes.fr/hal-02349385

[49] S. LEGLAIVE, L. GIRIN, R. HORAUD.*Semi-supervised multichannel speech enhancement with variational autoencoders and non-negative matrix factorization*, in "ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)", Brighton, United Kingdom, IEEE, May 2019, p. 101-105 [*DOI :* 10.1109/ICASSP.2019.8683704], https://hal.inria.fr/hal-02005102

[50] S. LEGLAIVE, U. SIMSEKLI, A. LIUTKUS, L. GIRIN, R. HORAUD.*Speech enhancement with variational autoencoders and alpha-stable distributions*, in "ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing", Brighton, United Kingdom, IEEE, 2019, p. 541-545, https://arxiv.org/abs/1902.03926 [*DOI :* 10.1109/ICASSP.2019.8682546], https://hal.inria.fr/hal-02005106

[51] X. LI, R. HORAUD.*Multichannel Speech Enhancement Based on Time-frequency Masking Using Subband Long Short-Term Memory*, in "IEEE Workshop on Applications of Signal Processing to Audio and Acoustics", New Paltz, NY, United States, October 2019, p. 1-5, https://hal.inria.fr/hal-02264247

[52] B. MASSÉ, S. LATHUILIÈRE, P. MESEJO, R. HORAUD.*Extended Gaze Following: Detecting Objects in Videos Beyond the Camera Field of View*, in "FG 2019 - 14th IEEE International Conference on Automatic Face and Gesture Recognition", Lille, France, IEEE, May 2019, p. 1-8 [*DOI :* 10.1109/FG.2019.8756555], https://hal.inria.fr/hal-02054236

Other Publications

[53] S. LEGLAIVE, X. ALAMEDA-PINEDA, L. GIRIN, R. HORAUD.*A Recurrent Variational Autoencoder for Speech Enhancement*, October 2019, https://arxiv.org/abs/1910.10942 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02329000

[54] X. LI, R. HORAUD.*Narrow-band Deep Filtering for Multichannel Speech Enhancement*, November 2019, https://arxiv.org/abs/1911.10791 - working paper or preprint, https://hal.inria.fr/hal-02378413

[55] M. SADEGHI, S. LEGLAIVE, X. ALAMEDA-PINEDA, L. GIRIN, R. HORAUD.*Audio-visual Speech Enhancement Using Conditional Variational Auto-Encoder*, November 2019, https://arxiv.org/abs/1908.02590 - Submitted to IEEE/ACM Transactions on Audio, Speech, and Language Processing, https://hal.inria.fr/hal-02364900

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, HDR]
Sabine Coquillart [Inria, Senior Researcher, HDR]
Thierry Fraichard [Inria, Researcher, HDR]
Emmanuel Mazer [CNRS, Senior Researcher, until Oct 2019, HDR]

Faculty Members

James Crowley [Team leader, Institut polytechnique de Grenoble, Professor, HDR]
Philippe Dessus [Univ Grenoble Alpes, Professor, until Aug 2019, HDR]

Patrick Reignier [Institut polytechnique de Grenoble, Professor, HDR]
Dominique Vaufreydaz [Univ Grenoble Alpes, Associate Professor, HDR]

External Collaborators

Jeremy Belot [Univ Grenoble Alpes, until Aug 2019]
Joelle Coutaz [Univ Grenoble Alpes]
Philippe Dessus [Univ Grenoble Alpes, from Sep 2019, HDR]
Didier Piau [Univ Grenoble Alpes]

Technical Staff

Raffaella Balzarini [Inria, Engineer]
Maxime Belgodere [Inria, Engineer, until Sep 2019]
Eric Boucher [Inria, Engineer, from Mar 2019]
Anthony Chavoutier [Inria, Engineer, until Jun 2019]
Benjamin Girard [Inria, Engineer]
Samuel Heidmann [Inria, Engineer, until Aug 2019]

PhD Students

Nachwa Abou Bakr [Univ Grenoble Alpes, PhD Student]
Louis Airale [Univ Grenoble Alpes, PhD Student, from Oct 2019]
Matteo Ciocca [Univ Grenoble Alpes, PhD Student]
Jose Da Silva Filho [Inria, PhD Student]
Amr Al-Zouhri Al-Yafi [Univ Grenoble Alpes, PhD Student]
Raphael Lambert Frisch [Univ Grenoble Alpes, PhD Student, until Oct 2019]
Thomas Guntz [Inria, PhD Student]
Romain Laurent [Univ Grenoble Alpes, PhD Student, from Feb 2019]
Rottana Ly [Ministre des Affaires Etrangeres, PhD Student, from Nov 2019]

Post-Doctoral Fellows

Amr Al-Zouhri Al-Yafi [Univ Grenoble Alpes, Post-Doctoral Fellow, from Jun 2019]
Marvin Faix [Univ Grenoble Alpes, Post-Doctoral Fellow, until Oct 2019]
Ivana Schenk [Inria, Post-Doctoral Fellow, until Oct 2019]

Administrative Assistants

Alexandra Fitzgerald [Inria, Administrative Assistant, until Sep 2019]
Julia Di Toro [Inria, Administrative Assistant, from Oct 2019]

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. Modelling Human Awareness and Undestanding

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 has addressed 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.

The challenges in this research area have been 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 concerns 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 have developed 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 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 modelling 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 have explored 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 explored 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].

We have adapted this work to construct narrative descriptions of cooking activities from ego-centric vision, 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, pp. 1-17, 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), pp. 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, pp. 612-619, 2014.

[Edwards 14] C. Edwards. 2014, "Decoding the language of human movement". Commun. ACM 57, 12, pp. 12-14, November 2014.

[Kahneman 66] D. Kahneman and J. Beatty, Pupil diameter and load on memory. Science, 154(3756), pp. 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, pp. 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, pp. 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. 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.

3.4.2. 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.

3.4.3. 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), 1-20, 2010

4. Application Domains

4.1. Modelling of awareness and expertise from Eye Gaze and Emotion

Humans display awareness and expertise through a variety of non-verbal channels. It is increasingly possible to record and interpret such information with available technology. In the ANR CEEGE project, 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.

An initial experiment with multi-modal observation of human experts engaged in solving problems in Chess revealed an unexpected observation of rapid changes in emotion as players attempt to solve challenging problems. In a scientific collaboration with the NeuroCognition group at the Univ Bielefeld, we have constructed to explain for chess experts that explains these unexpected results. This model has recently been tested in a second experiment with 22 chess players. Our results indicate that chess players associate emotions to chess chunks, and reactively use these associations to guide search in chunks for planning and problem solving. These results have recently been reported in a paper at the International Conference on Multimodal Interaction, and is the subject of the nearly completed doctoral dissertation of Thomas Guntz.

The results are currently being used in the construction of a student aware driver training device to be commercialized by SME company Sym2B financed by the SATT Linksium Project MAT: Monitoring Attention of Trainees, starting in Sept 2019. IN this project we will construct a training simulator for operation of busses and tramways.

4.2. Narrative Description of Kitchen Activities from Egocentric Video

We have developed and evaluated a system to construct situated, narrative descriptions of cooking activities including food preparation, place setting, cleaning and placing objects in storage areas. We are specifically interested in real-time, on-line techniques that recognize and interpret food types, food states and manipulation actions for transformation preparation of food. We are exploring techniques for detecting, modelling, and recognising a large vocabulary of actions and activities under different observational conditions, and describing these activities in a larger context.

A full understanding of human actions requires: recognising what action has been performed, predicting how it will affect the surrounding environment, explaining why this action has been performed, and who is performing it . Classic approaches to action recognition interpret a spatio-temporal pattern in a video sequence to tell what action has been performed, and perhaps how and where it was performed. A more complete understanding requires information about why the action was performed, and how it affects the environment. This face of understanding can be provided by explaining the action as part of a narrative.

Most work on recognition of cooking activities has concentrated on recognizing actions from the spatio-temporal patterns of hand motions. While some cooking activities may be directly recognized from motion, the resulting description is incomplete, as it does not describe the state of the ingredients, and how these have been transformed by cooking actions. A fuller description requires a description of how food ingredients have been transformed during the food preparation process.

We have addressed the automatic construction of cooking narratives by first detecting and locating ingredients and tools used in food preparation. We then recognize actions that involve transformations of ingredients, such as "slicing", and use these transformations to segment the video stream into visual events. We can then interpret detected events as a causal sequence of voluntary actions, confirmed by spatio-temporal transformation patterns, in order to automatically provide a narrative.

Our method is inspired by the intuition that object states are visually more apparent than actions from a still frame and thus provide information that is complementary to spatio-temporal action recognition. We define a state transition matrix that maps action labels into a pre-state and a post-state. We identify key frames, and use these to learn appearance models of objects and their states. For recognition, we use a modified form of VGG neural network trained via transfer learning with a specially constructed data set of images of food types and food sates. Manipulation actions are hypothesized from the state transition matrix and provide complementary evidence to spatio temporal action recognition.

4.3. Embedded Computer Vision for low-power Bolometric Imaging

In cooperation with Schneider Electric, we have developed techniques for embedded real time image analysis and tracking algorithms using Bolometric Imaging Sensors. Such sensors capture light in the far infrared and return an image where each pixel is a measurement of temperature in degree centigrade. We have designed an integrated low-cost sensor that combines an 80x80 pixel Bolometric imager with a low power micro-processor. The device provides real-time, embedded image processing for target detection and activity analysis where all sensing and interpretation is local. No images are recorded and only relevant information is about activity is communicated. The design of this system is under consideration for a patent, and thus has not been published. The software system has been registered with French APP , and is to be licensed to Schneider Electric for use a line of sensors for detecting falls and other activities for assisted living for seniors, as well as monitoring customer activities in commercial environments.

4.4. Recognizing and predicting routine activities in smart homes

Most research on smart home systems has concentrated on techniques for recognizing context. However, many categories of service require information about likely future context. We have developed and approach that uses dynamic Bayesian networks to predict future activity and context in a smart home. Our approach extends a state-of-the-art prediction model with three contributions. First, we include sensor data through aggregation nodes, instead of limiting ourselves only to higher level context data. Second, our method uses higher order relations between activities, so that past activities can have a more meaningful impact on prediction of future activities. Third, we use a latent node that estimates the cognitive state of the occupant.

4.5. User centred energy management

Participants: Amr Al-Zouhri Al-Yafi, 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.

4.6. 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.

4.7. 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.

5. Highlights of the Year

5.1. Highlights of the Year

James Crowley was named to the Chair on Intelligent Collaborative Systems.

6. New Results

6.1. Observing and Modelling Expertise and Awareness from Eye-gaze and Emotion

Participants: Thomas Guntz, James Crowley, Dominique Vaufreydaz, Philippe Dessus, Raffaella Balzarini.

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]. 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.

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.

6.2. Recognition, Modelling and Description of Manipulation Actions

Participants: Nachwa Abou Bakr, James Crowley.

A full understanding of human actions requires: recognizing what action has been performed, predicting how it will affect the surrounding environment, explaining why this action has been performed, and who is performing it. Classic approaches to action recognition interpret a spatio-temporal pattern in a video sequence to tell what action has been performed, and perhaps how and where it was performed. A more complete understanding requires information about why the action was performed, and how it affects the environment. This face of understanding can be provided by explaining the action as part of a narrative.

We have addressed the problem of recognition, modelling and description of human activities, with results on three problems: (1) the use of transfer learning for simultaneous visual recognition of objects and object states, (2) the recognition of manipulation actions from state transitions, and (3) the interpretation of a series of actions and states as events in a predefined story to construct a narrative description.

These results have been developed using food preparation activities as an experimental domain. We start by recognizing food classes such as tomatoes and lettuce and food states, such as sliced and diced, during meal preparation. We adapt the VGG network architecture to jointly learn the representations of food items and food states using transfer learning. We model actions as the transformation of object states. We use recognised object properties (state and type) to detect corresponding manipulation actions by tracking object transformations in the video. Experimental performance evaluation for this approach is provided using the 50 salads and EPIC-Kitchen datasets. We use the resulting action descriptions to construct narrative descriptions for complex activities observed in videos of 50 salads dataset.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

7.1.1. CIFRE Doctoral Contract with eLichens

The main topic of the project is to develop cloud-based services for Building Management System (BMS) framework. The aim is to develop predictive algorithms to control Heat Ventilation and Air Conditioning (HVAC) systems addressing two main goals:

1. Improve the well-being of the occupants keeping different variables as temperature, humidity, CO2, air quality measures inside a pre established optimal range;
2. Saving costs optimizing energy consumption

This research is supervised by Patrick Reignier.

8. Partnerships and Cooperations

8.1. National Initiatives

8.1.1. LabEx Persyval, Project RHUM, Robots in Human Environments

Participants: Thierry Fraichard, 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.

8.1.2. ExpeSigno

Participants: Patrick Reignier, Amr Al-Zouhri Al-Yafi, Amine Awada.

Projet Région Pack Ambition Recherche EXPESIGNO : Expérimentation de la réactivité des ménages aux signaux des opérateurs de systèmes énergétiques
Other Partners : Laboratoire Gaël, Laboratoire G2ELAb, laboratoire G-Scop
Dates : 2018 - 2022

Buildings represent 66% of electricity consumption and they can act as nodes in a network of consumption, storage and energy production. In this case, it can be understood that buildings and their inhabitants will change from a passive consumer to an active consumer (the so called “pro-sumer”) who can respond quickly to price changes on the network and / or signals from operators, or even other pro-sumers offering energy production and storage solutions using solar panels or electric cars. To achieve this goal, energy systems must send consumers the right signal to induce appropriate local and global behavior. The introduction of equipment such as Smart Meters or interactive consumption management devices is decisive because they are considered as the solution to turn residential consumers into active users of their electricity or energy consumption. Nudges are an interesting way to induce lasting changes in consumer behavior. The idea of nudges is to set up environments of choice that help people make the choices that are best for them. During this project, we are going to deploy sensors within 4 volunteer families in order to study the impact of nudges on electricity consumption through a detailed analysis of the practices carried out. The objective is to establish the links between the sensor data and the activities declared by each household and to measure how nudges influence their activities.

8.1.3. ANR Project CEEGE: Chess Expertise from Eye Gaze and Emotion

Participants: Thomas Guntz, James Crowley, Dominique Vaufreydaz, Raffaella Balzarini.

Other Partners : Dept of NeuroCognition, CITEN, Bielefeld University
Dates : Jan 2016 to Dec 2019

The ANR CEEGE project 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 have observed 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 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 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 has been used to determine and model the degree to which a player understands the game situation in terms of abstract configurations of chess pieces. This provides a structured environment that use for experimental evaluation of current theories of mental modeling and emotional response during problem solving and social interaction.

The project have been organized in two phases. During the first phase, we will observed individual players of different levels of chess expertise solving known chess problems. We correlated scan-path from eye tracking and other information about visual attention to established configurations of pieces and known solutions to chess problems. We constructed 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 have observed the attention and face expressions of pairs of players of different levels of chess ability solving problems followed by verbal self reports. We have used 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.

8.1.4. CDP EcoSesa - Cross Disciplinary Project of the ComUE UGA

Participants: Patrick Reignier, James Crowley, Raffaella Balzarini, Amr Al-Zouhri Al-Yafi.

Funding : UGA Idex Cross disciplinary project
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

8.1.5. 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.

8.1.6. 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.

8.1.7. LabEx Persyval - Project MicroBayes: Probabilistic Machines for Low-level Sensor Interpretation

Participants: Emmanuel Mazer, Raphael Frisch.

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.

8.1.8. 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.

8.2. European Initiatives

8.2.1. FP7 & H2020 Projects

8.2.1.1. AI4EU - A European AI On-Demand Platform and Ecosystem

Call: H2020 ICT-26-2018-2020
Coordinateur: Thales Systems
Partners: 79 European institutions
Dates: Jan 2019 through Dec 2021

AI4EU will build a comprehensive European AI-on-demand Platform that provides innovators in all areas of society with access to expertise, knowledge, algorithms and tools for developing, deploying and funding innovations based on Artificial Intelligence.

The aim is to empower actors across a broad spectrum of commercial, industrial and societal sectors in Europe with tools for innovation through AI Technologies. By bringing together a whole ecosystem of researchers, innovators, SMEs, large corporations, students and many others, around a single access point to AI resources, we will lower the barriers to education, research and innovation. Moreover, the AI4EU Platform will embrace on European values, respect European laws and support a human-centric approach providing a competitive advantage for European players.

8.2.1.2. H2020 FET Human AI

Call: H2020 FETFLAG-01-2018
Coordinateur: DFKI
Partners: 49 European institutions
Dates:1 March 2019 to 31 May 2020.

Humane AI has been funded to create a European network of centers of excellence for Artificial Intelligence technologies that synergistically work with humans, seamlessly fit in with our complex social settings and dynamically adapt to changes in our environment. The project seeks to develop world-leading insights and AI technologies, from fundamental algorithms, through methods specific to concrete applied AI domains such as Computer Vision, Robotics, IoT, Language Technologies and multi Agent Systems all the way up to disruptive AI applications and broadly usable platforms. Core innovations include (1) tools for enhancing human cognitive capabilities, channeling human creativity, inventiveness and intuition and empowering humans to make important decisions in a more informed way, (2) AI systems that can intelligently interact with and within complex social settings and seamlessly adapt to changing, open-ended environments, (3) explainable, transparent, validated and thus trustworthy AI systems that will help us more effectively deal with the complexity of a networked globalized world and (4) ways to embed values, ethics, privacy and security as core design considerations in all AI systems and applications.

To ensure broad and lasting socio-economic impact in areas which are important to Europe and its citizens on top of the basic research we will implemented dedicated impact-oriented work packages in domains such as Society and Policy, Industry 4.0, Sustainability and Energy, Finance, Science and Education, Health and Mobility/Automotive. To realize the Humane AI vision the consortium has lined up key European players and brought the relevant community on board to mobilize the critical mass needed for success. Many of the partners have strong interdisciplinary research track records, and several PIs on this project hold ERC grants, documenting scientific excellence. With their capability, networks and experience, we have a solid plan to bring the remaining players into the flagship activity during the preparatory action phase.

8.3. International Research Visitors

8.3.1. Visits of International Scientists

8.3.1.1. Sethserey Sam, Vice-Président NIPTICT, Phnom Penh

Position: Vice-Président en charge de la recherche et des relations internationales du NIPTICT, Phnom Penh, Cambodge (et son assistante)
Date: Du 14 au 17 Avril 2019

8.3.1.2. Dr. Dao Trung Kien

Position: Directeur adjoint de l'Institut MICA, HUST, Hanoi, Vietnam
Date: novembre et décembre 2019
Travail sur la thématique de la localisation indoor de personnes grâce aux technologies sans fil et à la fusion intelligente de données hétérogènes.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. Member of the Organizing Committees

Eric Castelli Co-organised the workshop BREVES (Vietnam, France Belgique), in Hanoi the 5 and 6 november 2019

Sabine Coquilart served as co-chair of the Paper Award committee for IEEE VR 2019 - IEEE Virtual Reality Conference

Sabine Coquilart served as co-chair of the Paper Award committee for SUI 2019 - ACM Spatial User Interaction

James Crowley served as Chairman of the Awards committee for ACM ICMI 2019 - International Conference on Multimodal Interaction.

Eric Castelli served as “board member” of the international conference series SLTU’xx (Spoken Language Technologies for Under-resourced languages)

9.1.1.2. Member of the Conference Program Committees

Sabine Coquilart served on the conference program committee for

- AIVR 2019 - International Conference on Artificial Intelligence and Virtual Reality
- CENTRIC 2019 - The Twelfth International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services
- GRAPP 2019 - International Conference on Computer Graphics Theory and Applications
- ICGI 2019 - International Conference on Graphics and Interaction
- ISVC 2019 - International Symposium on Visual Computing
- VRST 2019 - ACM Symposium on Virtual Reality Software and Technology
- WSCG 2019 - International Conferences in Central Europe on Computer Graphics, Visualization and Computer Vision

9.1.2. Scientific Events: Selection

9.1.2.1. Chair of Conference Program Committees

Sabine Coquilart served as co-chair for the paper selection panel for IEEE VR 2019

Sabine Coquilart served as co-chair for the paper selection panel for SUI 2019

9.1.2.2. Reviewer

Sabine Coquilart reviewed papers IEEE World Haptics 2019

Thierry Fraichard reviewed papers IEEE ICRA 2019

James Crowley served as a reviewer for

- ICCV 2019 - International Conference on Computer Vision
- CVPR 2019 - IEEE Conference on Computer Vision and Pattern Recognition
- AAAI 2020 - Association for the Advancement of Artificial Intelligence
- iCMI 2019 - International Conference on Multimodal Interaction.
- IHM 2019 - Interaction Homme Machine

Eric Castelli served as a reviewer for

- SigTelCom 2019 (Hanoi, Vietnam)
- ICVES 2019 (Cairo, Egypt)
- ISEE2019 (Ho Chi Minh City, Vietnam)
- SPECOM 2019 (Istanbul, Turkey)
- SCS'19 (Kingdom of Bahrain)
- ATC2019 (Hanoi, Vietnam)
- ICMSAO2019 (Bahrain)
- IALP2019 (Taipei, Taiwan)
- SoICT 2019 (Hanoi & HaLong Bay, Vietnam)
- 3ICT'19 (Kingdom of Bahrain)
- NICS 2019 (Hanoi, Vietnam)
- FLAIRS'33 2020 (North Miami Beach, USA)
- ICWMC 2020 (Athens, Greece)
- LREC 2020 (Marseille, France)
- MAPR 2020 (Hanoi, Vietnam)

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

Patrick Reignier is a member of the editorial board of the Modeling and Using Context Journal.
Sabine Coquilart is a member of the Scientific Committee of the Journal of Virtual Reality and Broadcasting.

9.1.3.2. Reviewer - Reviewing Activities

Sabine Coquilart reviewed papers for IEEE Transactions on Haptics
Thierry Fraichard reviewed papers for IEEE Trans. Robotics and Automation (TRO)
James Crowley reviewed papers for Communications of the ACM

9.1.4. Invited Talks

James Crowley gave the following invited talks:

- **Artificial Intelligence: a Rupture Technology for Innovation**, invited lecture at Club des DIS, Dirigeants Innovation et Stratégie, Montbonnot, 14 March 2019.
- **Collaborative Intelligent Systems**, invited lecture at Simula Research Institute, Oslo Norway, Montbonnot, 27 September 2019.
- **Artificial Intelligence: a Rupture Technology for Innovation**, AI4EU Web Cafe (webinar), 13 November 2019
- **Artificial Intelligence for Human Computer Interaction**, Keynote lecture at IHM 2019, Grenoble 12 Dec 2019.

Eric Castelli gave the invited talk

- **Multimodal perception and Pervasive spaces for Sociable interaction** International workshop BREVES, USTH, Hanoi, Vietnam, 5 and 6 November 2019.

9.1.5. Leadership within the Scientific Community

Sabine Coquilart is elected member of the EUROGRAPHICS Executive Committee.
Sabine Coquilart is co-chair of the EUROGRAPHICS Working Group and Workshop board.
Sabine Coquilart is co-chair of the Steering Committee for EGVE - Eurographics Working Group on Virtual Environments
James Crowley is a member of the Steering Committee for ICMI - International Conference on Multimodal Interaction.

9.1.6. Scientific Expertise

Patrick Reignier is a member of the Scientific Council of the Amiqual4Home EquipEx.
Thierry Fraichard served as an expert reviewer for the European Commission, the IDEX of Aix-Marseille Université, and the IDEX of Université Grenoble Alpes

9.1.7. Research Administration

Thierry Fraichard is a member of the following Inria committees:

- Research Center Committee (CC), IT Service Committee (CUMI), and Technological Development Committee (CDT).

James Crowley is an elected member of the Conseil d'Administration of the Comue Univ Grenoble Alpes
James Crowley Served on the Bureau (executive committee) of the laboratoire LIG
Patrick Reignier is head of the engineering support group of the Laboratoire d'Informatique de Grenoble (13 members).
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)

9.2. Teaching - Supervision - Juries

Patrick Reignier teaches

- elected member of the Conseil des Etudes et de la Vie Universitaire of Grenoble INP
- nominated as a member of the Conseil de la Formation Continue de Grenoble INP
- co-director of the "formation en apprentissage" of Ensimag (3 years program : 1 year for the Licence and 2 years for the Master)
- Supervises the industrial part of the “formation en apprentissage” of the Ensimag engineering school.

James Crowley and **Dominique Vaufreydaz** co-direct the Graphics - Vision - Robotics Specialisation of the MoSIG M2 Masters.

9.2.1. Teaching

James Crowley teaches

- M2 MoSIG Computer Vision
- ENSIMAG 3 Machine Learning.
- M1 MoSIG Intelligent Systems.
- ENSIMAG 2 Intelligent Systems.

Patrick Reignier teaches

- Projet Genie Logiciel, 55h eqTD, M1, Ensimag/Grenoble INP, France.
- Développement d'applications communicantes, 18h eqTD, M2, Ensimag/Grenoble-INP, France
- Introduction aux applications reparties, 18h eqTD, M2, Ensimag/Grenoble- INP, France
- Applications Web et Mobiles , 27h eqTD, M1, Ensimag/Grenoble-INP, France
- Projet Systeme, 12h eq TD, M1, Ensimag/Grenoble-INP, France
- Projet C, 20h eqTD, L3, Ensimag/Grenoble-INP, France.

Thierry Fraichard, taught Autonomous Robotics, 22.5h eqTD, M2 MOSIG, Univ. Grenoble Alpes.

9.2.2. Supervision

Thierry Fraichard directed the following Doctoral students

- Jose Grimaldo Da Silva Filho, "Human-Robot Motion, a Shared Effort Approach", Octobre 2015, Thierry Fraichard and James Crowley.
- Matteo Ciocca, "Safe Robot Motion", Octobre 2016, Thierry Fraichard and Pierre-Brice Wieber.

James Crowley directed the following Doctoral students

- Nachwa Abou Bakr, "Recognition, Modeling and Description of Manipulation actions "
- Thomas Guntz, "Estimating Expertise from Eye Gaze and Emotions", co-directed with Dominique Vaufreydaz

9.2.3. Juries

Patrick Reignier served as a president of the Doctoral Jury for

- Van Bao Nguyen
- Alessandro Fenicio

James Crowley served as a president of the Doctoral Jury for

- Mathieu Barbier
- David Sierre Gonzalez

9.3. Interventions

Thierry Fraichard participated to the event “A robot on trial”, mock trial of a futuristic robot as part of the Transfo digital festival, Jan. 19. This successful event attracted around 200 participants

10. 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

Articles in International Peer-Reviewed Journal

[6] A. ALZOUHRI ALYAFI, M. PAL, 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, vol. 238, p. 1452-1470 [*DOI :* 10.1016/J.APENERGY.2019.01.118], https://hal.archives-ouvertes.fr/hal-01984833

[7] M. BOUGUERRA, T. FRAICHARD, M. FEZARI.*Viability-Based Guaranteed Safe Robot Navigation*, in "Journal of Intelligent and Robotic Systems", August 2019, vol. 95, n^{o} 2, p. 459–471 [*DOI :* 10.1007/s10846-018-0955-9], https://hal.inria.fr/hal-01924855

[8] T. FRAICHARD, V. LEVESY.*From Crowd Simulation to Robot Navigation in Crowds*, in "IEEE Robotics and Automation Letters", April 2020, vol. 5, n^{o} 2, p. 729-735 [*DOI :* 10.1109/LRA.2020.2965032], https://hal.inria.fr/hal-02461493

[9] G. NIETO, F. DEVERNAY, J. L. CROWLEY.*Rendu basé image avec contraintes sur les gradients*, in "Traitement du Signal", 2019, p. 1-26, forthcoming [*DOI :* 10.3166/HSP.X.1-26], https://hal.archives-ouvertes.fr/hal-01900200

[10] R. PAULIN, T. FRAICHARD, P. REIGNIER.*Using Human Attention to Address Human-Robot Motion*, in "IEEE Robotics and Automation Letters", February 2019, vol. 4, n^{o} 2, p. 2038-2045, The content of this paper was also selected by the Program Committee for presentation at the IEEE Int. Conf. On Robotics and Automation (ICRA), May 2019, Montréal (CA) [*DOI :* 10.1109/LRA.2019.2899429], https://hal.inria.fr/hal-02013578

[11] V.-C. TA, T.-K. DAO, D. VAUFREYDAZ, E. CASTELLI.*Collaborative Smartphone-Based User Positioning in a Multiple-User Context Using Wireless Technologies*, in "Sensors", January 2020, 25 [*DOI :* 10.3390/s20020405], https://hal.inria.fr/hal-02435610

International Conferences with Proceedings

[12] M. CIOCCA, P.-B. WIEBER, T. FRAICHARD.*Effect of Planning Period on MPC-based Navigation for a Biped Robot in a Crowd*, in "IROS 2019 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Macau, China, IEEE, November 2019, p. 1-8 [*DOI :* 10.1109/IROS40897.2019.8968070], https://hal.inria.fr/hal-02267426

[13] J. G. DA SILVA FILHO, A.-H. OLIVIER, A. CRÉTUAL, J. PETTRÉ, T. FRAICHARD.*Effective Human-Robot Collaboration in near symmetry collision scenarios*, in "RO-MAN 2019 - 28th IEEE International Conference on Robot & Human Interactive Communication", New Dehli, India, IEEE, October 2019, p. 1-8, https://hal.inria.fr/hal-02267705

[14] R. FRISCH, M. FAIX, J. DROULEZ, L. GIRIN, E. MAZER.*Bayesian time-domain multiple sound source localization for a stochastic machine*, in "EUSIPCO 2019 - 27th European Signal Processing Conference", A Coruna, Spain, IEEE, September 2019, p. 1-5 [*DOI :* 10.23919/EUSIPCO.2019.8902666], https://hal.archives-ouvertes.fr/hal-02377220

[15] J. LE LOUEDEC, T. GUNTZ, J. L. CROWLEY, D. VAUFREYDAZ.*Deep learning investigation for chess player attention prediction using eye-tracking and game data*, in "ETRA 2019 - ACM Symposium On Eye Tracking

Research & Applications", Denver, United States, ACM, June 2019, p. 1-15, https://arxiv.org/abs/1904.08155 [*DOI :* 10.1145/3314111.3319827], https://hal.inria.fr/hal-02100389

National Conferences with Proceeding

[16] M. TROEL-MADEC, J. ALAIMO, L. BOISSIEUX, S. CHATAGNON, S. BORKOSWKI, A. SPALANZANI, D. VAUFREYDAZ.*eHMI positioning for autonomous vehicle/pedestrians interaction*, in "IHM 2019 - 31e conférence francophone sur l'Interaction Homme-Machine", Grenoble, France, ACM, 2019, p. 1-8 [*DOI :* 10.1145/3366551.3370340], https://hal.archives-ouvertes.fr/hal-02388847

Conferences without Proceedings

[17] N. ABOUBAKR, J. L. CROWLEY, R. RONFARD.*Recognizing Manipulation Actions from State-Transformations*, in "EPIC@CVPR 2019 - Fourth International Workshop on Egocentric Perception Interaction and Computing", Long Beach, United States, June 2019, p. 1-4, https://arxiv.org/abs/1906.05147 - Accepted for presentation at EPIC@CVPR2019 workshop, https://hal.archives-ouvertes.fr/hal-02197549

[18] F. PORTET, S. CAFFIAU, F. RINGEVAL, M. VACHER, N. BONNEFOND, S. ROSSATO, B. LECOUTEUX, T. DESOT.*Context-Aware Voice-based Interaction in Smart Home -VocADom@A4H Corpus Collection and Empirical Assessment of its Usefulness*, in "PICom 2019 - 17th IEEE International Conference on Pervasive Intelligence and Computing", Fukuoka, Japan, 2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress, IEEE, August 2019, p. 811–818 [*DOI :* 10.1109/DASC/PICOM/CBDCOM/CYBERSCITECH.2019.00149], https://hal.archives-ouvertes.fr/hal-02165532

Research Reports

[19] N. ABOUBAKR, J. CROWLEY, R. RONFARD.*Recognizing Manipulation Actions from State-Transformations (Technical Report)*, Univ. Grenoble Alps, CNRS, Inria, Grenoble INP, LIG, 38000 Grenoble, France, June 2019, https://hal.archives-ouvertes.fr/hal-02197595

[20] R. BALZARINI.*Rapport d'étude campagne expérimentale -exploratoire -pour l'étude de l'attention visuelle en eye-tracking projet COM2SICA*, Inria GRA, Laboratoire GéoAzur CNRS, UMR Espace, November 2019, https://hal.inria.fr/hal-02388358

Scientific Popularization

[21] R. LAURENT, D. VAUFREYDAZ, P. DESSUS.*Ethical Teaching Analytics in a Context-Aware Classroom: A Manifesto*, in "ERCIM News", January 2020, n^{o} 120, p. 39–40, https://hal.archives-ouvertes.fr/hal-02438020

Other Publications

[22] M. TROEL-MADEC.*Concilier les comportements intentionnels du véhicule autonome et du piéton*, Pôle supérieur de Design Léonard de Vinci, 38090 Villefontaine, France, June 2019, https://hal.inria.fr/hal-02274535

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, Researcher, HDR]
Jonatha Anselmi [Inria, Researcher, from Apr 2019]

Nicolas Gast [Inria, Researcher]
Bruno Gaujal [Inria, Senior Researcher, HDR]
Patrick Loiseau [Inria, Researcher, HDR]
Panayotis Mertikopoulos [CNRS, Researcher, HDR]
Bary Pradelski [CNRS, Researcher]

Faculty Members
Vincent Danjean [Univ Grenoble Alpes, Associate Professor]
Guillaume Huard [Univ Grenoble Alpes, Associate Professor]
Florence Perronnin [Univ Grenoble Alpes, Associate Professor, HDR]
Jean-Marc Vincent [Univ Grenoble Alpes, Associate Professor]
Philippe Waille [Univ Grenoble Alpes, Associate Professor]

Technical Staff
Eleni Gkiouzepi [CNRS, Engineer, from Dec 2019]
Mouhcine Mendil [INP SA, Engineer, until Jul 2019]

PhD Students
Kimon Antonakopoulos [Inria, PhD Student]
Thomas Barzola [Univ Grenoble Alpes, PhD Student, from Oct 2019]
Tom Cornebize [Univ Grenoble Alpes, PhD Student]
Bruno de Moura Donassolo [Orange Labs, PhD Student, granted by CIFRE]
Vitalii Emelianov [Inria, PhD Student]
Franz Christian Heinrich [Inria, PhD Student, until Feb 2019]
Alexis Janon [Univ Grenoble Alpes, PhD Student]
Baptiste Jonglez [Univ Grenoble Alpes, PhD Student]
Kimang Khun [Inria, PhD Student, from Oct 2019]
Alexandre Marcastel [Univ de Cergy Pontoise, until Feb 2019]
Dimitrios Moustakas [Inria, PhD Student, from Dec 2019]
Umar Ozeer [Orange Labs, PhD Student, until Oct 2019, granted by CIFRE]
Stephan Plassart [Univ Grenoble Alpes, PhD Student]
Pedro Rocha Bruel [UFRGS, PhD Student]
Benjamin Roussillon [Univ Grenoble Alpes, PhD Student]
Vera Sosnovik [Univ Grenoble Alpes, PhD Student, from Oct 2019]
Dong Quan Vu [Nokia Bell Labs, PhD Student, granted by CIFRE]
Chen Yan [Univ Grenoble Alpes, PhD Student, from Oct 2019]

Post-Doctoral Fellows
Olivier Bilenne [CNRS, Post-Doctoral Fellow]
Takai Eddine Kennouche [Univ Grenoble Alpes, Post-Doctoral Fellow]
Mouhcine Mendil [INP SA, Post-Doctoral Fellow, from Jul 2019]

Administrative Assistant
Annie Simon [Inria, Administrative Assistant]

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 [91], Jülich and TU Dresden [90], [61], UIUC [79], [94], [82] and ANL [107], Inria Bordeaux [67] and Grenoble [109], ...) but most of these tools build on the classical information visualization mantra [99] 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 [86]. 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 [89] 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 [98], [85], [78] to identify and characterize (in term of performance and resource usage) application phases and present aggregated representations of the trace [97]. Researchers from Jülich and TU Darmstadt have proposed techniques to identify specific communication patterns that incur wait states [104], [54]

3.3. Fast and Faithful Performance Prediction of Very Large Systems

Participants: Jonatha Anselmi, 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. [53] 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 [65], 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 [57] and the validity of the underlying network models [55], [102]. Many simulators of MPI applications have been developed by renowned HPC groups (e.g., at SDSC [100], BSC [51], UIUC [108], Sandia Nat. Lab. [103], ORNL [64] or ETH Zürich [80] 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) [69], [62], 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 [92]) or simply to generate trajectories with a forward Monte-Carlo simulation leveraging time parallel simulation (pioneered by Fujimoto [73], Lin and Lazowska [84]). One of the strength of the PSI framework is its expressiveness that allows us to easily study networks with finite and infinite capacity queues [63]. 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: Jonatha Anselmi, 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 [93] or to perform model-checking [52]. 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 [74], [83], [106], and has been successfully applied in various domains: wireless networks [56], computer-based systems [77], [88], [101], epidemic or rumour propagation [66], [81] and bike-sharing systems [70]. It is also used to develop distributed control strategies [105], [87] or to construct approximate solutions of stochastic model checking problems [58], [59], [60].

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 [74], we would like to be able to handle heterogeneous or uncertain dynamics. Typical applications are caching mechanisms [77] or bike-sharing systems [71]. A second point of interest is the use of mean field or large deviation asymptotics to compute the time between two regimes [96] 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 [75] to build distributed control algorithms [68] and optimal pricing mechanisms [76].

3.5. Distributed Learning in Games and Online Optimization

Participants: Nicolas Gast, Bruno Gaujal, Arnaud Legrand, Patrick Loiseau, Panayotis Mertikopoulos, Bary Pradelski.

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 [72], [95]. 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

- N. Gast received an ANR JCJC grant.
- The team was highly involved in the 3IA institute MIAI @ Grenoble Alpes: P. Loiseau is co-holder of the chair “Explainable and Responsible AI” of which N. Gast and B. Pradelski are members; and P. Mertikopoulos is a member of the chair “Optimization & Learning”.
- Arnaud Legrand participated in the writing of a book [39] on Reproducible Research, which aims at helping students and engineers and researchers to find efficient and accessible ways leading them to improve their reproducible research practices.

5.1.1. Awards

- The paper “Privacy Risks with Facebook’s PII-based Targeting: Auditing a Data Broker’s Advertising Interface” by P. Loiseau and co-authors (IEEE S&P ’18) was runner up for the 2019 Caspar Bowden Award for Outstanding Research in Privacy Enhancing Technologies.
- The paper “Investigating ad transparency mechanisms in social media: A case study of Facebook’s explanations” by P. Loiseau and co-authors (NDSS ’18) was runner up for the 2019 CNIL-Inria Award for Privacy Protection.

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/

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 2019: Python bindings were introduced, SMPI now partially supports some of the MPI/IO functions, a new model for Wifi networks was proposed, and the API for the simulation of storage resources was completely revisited. We also pursued our efforts to improve the documentation of the software, simplified the web site, and made a lot of bug fixing and code refactoring.

- 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

Performance engineering of scientific HPC applications requires to measure repeatedly the performance of applications or of computation kernels, which consume a large amount of time and resources. It is essential to design experiments so as to reduce this cost as much as possible. Our contribution along this axis is twofold: (1) the investigation sound exploration techniques and (2) the control of experiments to ensure the measurements are as representative as possible of real workload.

Writing, porting, and optimizing scientific applications 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 large budgets of costly experimental measurements to provide good results, while rarely providing exploitable knowledge after optimization. In [16], we investigate the use of *Design of Experiments* to propose a user-transparent autotuning technique 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 $10\times$ fewer measurements than random sampling. Although the results are very encouraging, our approach relies on assumptions on the geometry of the search space that are difficult to test in very large dimension. We are thus currently pursuing this line of research using non parametric approaches based on gaussian process regression, space filling designs and iteratively selecting configurations that yield the best expected improvement.

Our second contribution is related to the control of measurements. In [40], we relate a surprising observation on the performance of the highly optimized and regular DGEMM function on modern processors. The DGEMM function is a widely used implementation of the matrix product. While the asymptotic complexity of the algorithm only depends on the sizes of the matrices, we show that the performance is significantly impacted by the matrices content. Although it would be expected that special values like 1 or 0 may yield to specific behavior, we show that arbitrary constant values are no different and that random values incur a significant performance drop. Our experiments show that this may be due to bit flips in the CPU causing an energy consumption overhead. Such phenomenon reminds the importance of thoroughly randomizing every single parameter of experiments to avoid bias toward specific behavior.

7.2. Predictive Simulation of HPC Applications

Finely tuning MPI applications (number of processes, granularity, collective operation algorithms, topology and process placement) is critical to obtain good performance on supercomputers. With a rising cost of modern supercomputers, running parallel applications at scale solely to optimize their performance is extremely expensive. Using SimGrid, we work toward providing a methodology allowing to provide inexpensive but faithful predictions of expected performance.

The methodology we propose relies on SimGrid/SMPI and captures the complexity of adaptive applications by emulating the MPI code while skipping insignificant parts. In [18] we demonstrate its capability with High Performance Linpack (HPL), the benchmark used to rank supercomputers in the TOP500 and which requires a careful tuning. We explain (1) how we both extended the SimGrid's SMPI simulator and slightly modified the open-source version of HPL to allow a fast emulation on a single commodity server at the scale of a supercomputer and (2) how to model the different components (network, BLAS, ...) of the system. We show that a careful modeling of both spatial and temporal node variability allows us to obtain predictions within a few percents of real experiments. The modeling of BLAS operations is particularly important and we have thus started investigating in the context of simulating a sparse direct solver how to automatically performance models for commonly used BLAS kernels [33]. A key difficulty remains the acquisition of faithful performance measurements as modern processors are often quite unstable. This effort is therefore particularly related to the aforementioned "Design of Experiments" line of research.

7.3. Simulation of Smart Grids

In [35], we present ASGriDS, an asynchronous Smart Grid simulation framework. ASGriDS is multi-domain, it simultaneously models the power network along with its physical loads/generators, controllers, and communication infrastructure. ASGriDS provides a unified workflow in a pythonic environment, to describe, run and control complex SmartGrid deployment scenarios. ASGriDS is an event-driven simulator that can run in either real-time or accelerated real-time. As it is modular and its components interact asynchronously, it can run either locally on a distributed infrastructure, also in hardware-in-the- loop setups, and on top of emulated/physical communication links. In this paper, we present the design of our simulator and we demonstrate its use with a generation control problem on a low voltage network. We use ASGriDS to deploy a real-time controller based on optimal power flow, on top of TCP and UDP based communication network, under various packet loss conditions.

7.4. Batch Scheduling

Despite the impressive growth and size of super-computers, the computational power they provide still cannot match the demand. Efficient and fair resource allocation is a critical task. Super-computers use Resource and Job Management Systems to schedule applications, which is generally done by relying on generic index policies such as First Come First Served and Shortest Processing time First in combination with Backfilling strategies. Unfortunately, such generic policies often fail to exploit specific characteristics of real workloads.

In [36], we focus on improving the performance of online schedulers by studying mixed policies, which are created by combining multiple job characteristics in a weighted linear expression, as opposed to classical pure policies which use only a single characteristic. This larger class of scheduling policies aims at providing more flexibility and adaptability. We use space coverage and black-box optimization techniques to explore this new space of mixed policies and we study how can they adapt to the changes in the workload. We perform an extensive experimental campaign through which we show that (1) the best pure policy is far from optimal and that (2) using a carefully tuned mixed policy would allow to significantly improve the performance of the system. (3) We also provide empirical evidence that there is no one size fits all policy, by showing that the rapid workload evolution seems to prevent classical online learning algorithms from being effective.

A careful investigation of why such mixed strategy fail to globally exploit weekly workload features reveal that some users sometimes provide widely inaccurate information, which dramatically fools the batch scheduling heuristic. Indeed, users typically provide a loose upper bound estimate for job execution times that are hardly useful. Previous studies attempted to improve these estimates using regression techniques. Although these attempts provide reasonable predictions, they require a long period of training data. Furthermore, aiming for perfect prediction may be of limited use for scheduling purposes. In [50], we propose a simpler approach by classifying jobs as small or large and prioritizing the execution of small jobs over large ones. Indeed, small jobs are the most impacted by queuing delays but they typically represent a light load and incur a small burden on the other jobs. The classifier operates online and learns by using data collected over the previous weeks, facilitating its deployment and enabling fast adaptations to changes in workload characteristics. We evaluate our approach using four scheduling policies on six HPC platform workload traces. We show that: (i) incorporating such classification significantly reduces the average bounded slowdown of jobs in all scenarios, and (ii) the obtained improvements are comparable, in most scenarios, to the ideal hypothetical situation where the scheduler would know the exact running time of jobs in advance.

7.5. Load Balancing

In distributed systems, load balancing is a powerful concept to improve the distribution of jobs across multiple computing resources and to control performance metrics such as delays and throughputs while avoiding the overload of any single resource. This section describes three contributions:

- In multi-server distributed queueing systems, the access of stochastically arriving jobs to resources is often regulated by a dispatcher, also known as load balancer. A fundamental problem consists in designing a load balancing algorithm that minimizes the delays experienced by jobs. During the last two decades, the power-of-d-choice algorithm, based on the idea of dispatching each job to the least loaded server out of d servers randomly sampled at the arrival of the job itself, has emerged as a breakthrough in the foundations of this area due to its versatility and appealing asymptotic properties. In [8], we consider the power-of-d-choice algorithm with the addition of a local memory that keeps track of the latest observations collected over time on the sampled servers. Then, each job is sent to a server with the lowest observation. We show that this algorithm is asymptotically optimal in the sense that the load balancer can always assign each job to an idle server in the large-system limit. This holds true if and only if the system load λ is less than $1 - 1/d$. If this condition is not satisfied, we show that queue lengths are tightly bounded by $\lceil \frac{-\log(1-\lambda)}{\log(\lambda d+1)} \rceil$. This is in contrast with the classic version of the power-of-d-choice algorithm, where at the fluid scale a strictly positive proportion of servers containing i jobs exists for all $i \geq 0$, in equilibrium. Our results quantify and highlight the importance of using memory as a means to enhance performance in randomized load balancing.
- When dispatching jobs to parallel servers, or queues, the highly scalable round-robin (RR) scheme reduces the variance of interarrival times at all queues to a great extent but has no impact on the variances of service processes. Contrariwise, size-interval task assignment (SITA) routing has little impact on the variances of interarrival times but makes the service processes as deterministic as possible. In [6], we unify both 'static' approaches to design a scalable load balancing framework able to control the variances of the arrival and service processes jointly. It turns out that the resulting

combination significantly improves performance and is able to drive the mean job delay to zero in the large-system limit; it is known that this property is not achieved when both approaches are considered separately. Within realistic parameters, we show that the optimal number of size intervals that partition the support of the job size distribution is small with respect to the system size. This enhances the applicability of the proposed load balancing scheme at a large scale. In fact, we find that adding a little bit of information about job sizes to a dispatcher operating under RR improves performance a lot. Under the optimal scaling of size intervals and assuming highly variable job sizes, numerical simulations indicate that the proposed algorithm is competitive with the (less scalable) join-the-shortest-workload algorithm even when the system size grows large.

- Size-based routing provides robust strategies to improve the performance of computer and communication systems with highly variable workloads because it is able to isolate small jobs from large ones in a static manner. The basic idea is that each server is assigned all jobs whose sizes belong to a distinct and continuous interval. In the literature, dispatching rules of this type are referred to as SITA (Size Interval Task Assignment) policies. Though their evident benefits, the problem of finding a SITA policy that minimizes the overall mean (steady-state) waiting time is known to be intractable. In particular it is not clear when it is preferable to balance or unbalance server loads and, in the latter case, how. In [7], we provide an answer to these questions in the celebrated limiting regime where the system capacity grows linearly with the system demand to infinity. Within this framework, we prove that the minimum mean waiting time achievable by a SITA policy necessarily converges to the mean waiting time achieved by SITA-E, the SITA policy that equalizes server loads, provided that servers are homogeneous. However, within the set of SITA policies we also show that SITA-E can perform arbitrarily bad if servers are heterogeneous. In this case we prove that there exist exactly C! asymptotically optimal policies, where C denotes the number of server types, and all of them are linked to the solution of a single strictly convex optimization problem. It turns out that the mean waiting time achieved by any of such asymptotically optimal policies does not depend on how job-size intervals are mapped to servers. Our theoretical results are validated by numerical simulations with respect to realistic parameters and suggest that the above insights are also accurate in small systems composed of a few servers, i.e., ten.

7.6. FoG Computing

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. It offers nearby computational, networking and storage resources to respond to the stringent requirements of these applications. However, despite its several advantages, Fog computing raises new challenges which slow its adoption down. Hence, there is a lack of practical solutions to enable the exploitation of this novel concept.

In [19], we propose FITOR, an orchestration system for IoT applications in the Fog environment. This solution builds a realistic Fog environment 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. In [46], we propose a novel strategy, which we call GO-FSP and which optimizes the placement of IoT application components while coping with their strict performance requirements. To do so, we first propose an Integer Linear Programming (ILP) formulation for the IoT application provisioning problem. The latter targets to minimize the deployment cost while ensuring a load balancing between heterogeneous devices. Then, a GRASP-based approach is proposed to achieve the aforementioned objectives. Finally, we make use of the FITOR orchestration system to evaluate the performance of our solution under real conditions. Obtained results show that our scheme outperforms the related strategies. We are currently comparing such strategy with other strategies based on online learning mechanisms under various information scenarios (delayed and noisy feedback, inaccurate application load information, etc.).

Last, fog computing also extends the capacities of the cloud to the edge of the network, near the physical world, so that Internet of Things (IoT) applications can benefit from properties such as short delays, real-time and privacy. Unfortunately, devices in the Fog-IoT environment are usually unstable and prone to failures. In this context, the consequences of failures may impact the physical world and can, therefore, be critical. In [28], we present a framework for end-to-end resilience of Fog-IoT applications. The framework was implemented and experimented on a smart home testbed.

7.7. Research Management: Research Reproducibility and Credit

We are actively promoting better research practices, in particular in term of research reproducibility and contribution recognition. Our contribution this year is threefold

First, we have participated to the writing of a book introducing reproducible research [39]. For a researcher, there is nothing more frustrating than the failure to reproduce major results obtained a few months back. The causes of such disappointments can be multiple and insidious. This phenomenon plays an important role in the so-called "research reproducibility crisis". This book takes a current perspective onto a number of potentially dangerous situations and practices, to examplify and highlight the symptoms of non-reproducibility in research. Each time, it provides efficient solutions ranging from good-practices that are easily and immediately implementable to more technical tools, all of which are free and have been put to the test by the authors themselves. Students and engineers and researchers should find efficient and accessible ways leading them to improve their reproducible research practices.

Second, to allow students and engineers and researchers to receive proper training in reproducible research, we have run the second session of the Mooc "Reproducible research: Methodological principles for a transparent science" on the FUN platform from April, 1 to June, 13 2019. This MOOC allows scientists to learn modern and reliable tools such as Markdown for taking structured notes, Desktop search applications, GitLab for version control and collaborative working, and Computational notebooks (Jupyter, RStudio, and Org-Mode) for efficiently combining the computation, presentation, and analysis of data. More than 2,100 persons registered to this session and we are currently working on a third session which is expected to start in the beginning of the year 2020.

Third, software is a fundamental pillar of modern scientific research, not only in computer science, but actually across all fields and disciplines. However, there is a lack of adequate means to cite and reference software, for many reasons. An obvious first reason is software authorship, which can range from a single developer to a whole team, and can even vary in time. The panorama is even more complex than that, because many roles can be involved in software development: software architect, coder, debugger, tester, team manager, and so on. Arguably, the researchers who have invented the key algorithms underlying the software can also claim a part of the authorship. And there are many other reasons that make this issue complex. We provide in [5] a contribution to the ongoing efforts to develop proper guidelines and recommendations for software citation, building upon the internal experience of Inria, the French research institute for digital sciences. As a central contribution, we make three key recommendations. (1) We propose a richer taxonomy for software contributions with a qualitative scale. (2) We claim that it is essential to put the human at the heart of the evaluation. And (3) we propose to distinguish citation from reference which is particularly important in the context of reproducible research.

7.8. Mean Field Games and Control

In [10], we consider mean field games with discrete state spaces (called discrete mean field games in the following) and we analyze these games in continuous and discrete time, over finite as well as infinite time horizons. We prove the existence of a mean field equilibrium assuming continuity of the cost and of the drift. These conditions are more general than the existing papers studying finite state space mean field games. Besides, we also study the convergence of the equilibria of N -player games to mean field equilibria in our four settings. On the one hand, we define a class of strategies in which any sequence of equilibria of the finite games converges weakly to a mean field equilibrium when the number of players goes to infinity. On the other hand, we exhibit equilibria outside this class that do not converge to mean field equilibria and for which the value of the game does not converge. In discrete time this non-convergence phenomenon implies that the Folk theorem does not scale to the mean field limit.

In [20], we consider a class of nonlinear systems of differential equations with uncertainties, i.e., with lack of knowledge in some of the parameters that is represented by a time-varying unknown bounded functions. An under-approximation of such systems consists of a subset of its reachable set, for any value of the unknown parameters. By relying on optimal control theory through Pontryagin's principle, we provide an algorithm for the under-approximation of a linear combination of the state variables in terms of a fully automated tool-chain named UTOPIC. This allows to establish tight under-approximations of common benchmarks models with dimensions as large as sixty-five.

7.9. Energy and Network Optimization

This section describes four contributions on energy and network optimization.

- One of the key challenges in Internet of Things (IoT) networks is to connect many different types of autonomous devices while reducing their individual power consumption. This problem is exacerbated by two main factors: first, the fact that these devices operate in and give rise to a highly dynamic and unpredictable environment where existing solutions (e.g., water-filling algorithms) are no longer relevant; and second, the lack of sufficient information at the device end. To address these issues, we propose a regret-based formulation that accounts for arbitrary network dynamics: this allows us to derive an online power control scheme that is provably capable of adapting to such changes, while relying solely on strictly causal feedback. In so doing, we identify an important tradeoff between the amount of feedback available at the transmitter side and the resulting system performance: if the device has access to unbiased gradient observations, the algorithm's regret after T stages is $O(T^{-1/2})$ (up to logarithmic factors); on the other hand, if the device only has access to scalar, utility-based information, this decay rate drops to $O(T^{-1/4})$. The above is validated by an extensive suite of numerical simulations in realistic channel conditions, which clearly exhibit the gains of the proposed online approach over traditional water-filling methods. This contribution appeared in [11].
- Many businesses possess a small infrastructure that they can use for their computing tasks, but also often buy extra computing resources from clouds. Cloud vendors such as Amazon EC2 offer two types of purchase options: on-demand and spot instances. As tenants have limited budgets to satisfy their computing needs, it is crucial for them to determine how to purchase different options and utilize them (in addition to possible self-owned instances) in a cost-effective manner while respecting their response-time targets. In this paper, we propose a framework to design policies to allocate self-owned, on-demand and spot instances to arriving jobs. In particular, we propose a near-optimal policy to determine the number of self-owned instances and an optimal policy to determine the number of on-demand instances to buy and the number of spot instances to bid for at each time unit. Our policies rely on a small number of parameters and we use an online learning technique to infer their optimal values. Through numerical simulations, we show the effectiveness of our proposed policies, in particular that they achieve a cost reduction of up to 64.51% when spot and on-demand instances are considered and of up to 43.74% when self-owned instances are considered, compared to previously proposed or intuitive policies. This contribution appeared in [13].
- In [22], we consider the classical problem of minimizing offline the total energy consumption required to execute a set of n real-time jobs on a single processor with varying speed. Each real-time job is defined by its release time, size, and deadline (all integers). The goal is to find a sequence of processor speeds, chosen among a finite set of available speeds, such that no job misses its deadline and the energy consumption is minimal. Such a sequence is called an optimal speed schedule. We propose a linear time algorithm that checks the schedulability of the given set of n jobs and computes an optimal speed schedule. The time complexity of our algorithm is in $O(n)$, to be compared with $O(n \log(n))$ for the best known solutions. Besides the complexity gain, the main interest of

our algorithm is that it is based on a completely different idea: instead of computing the critical intervals, it sweeps the set of jobs and uses a dynamic programming approach to compute an optimal speed schedule. Our linear time algorithm is still valid (with some changes) with an arbitrary power function (not necessarily convex) and arbitrary switching times

- Network utility maximization (NUM) is an iconic problem in network traffic management which is at the core of many current and emerging network design paradigms - and, in particular, software-defined networks (SDNs). Thus, given the exponential growth of modern-day networks (in both size and complexity), it is crucial to develop scalable algorithmic tools that are capable of providing efficient solutions in time which is dimension-free, i.e., independent-or nearly-independent-on the size of the system. To do so, we leverage a suite of modified gradient methods known as “mirror descent” and we derive a scalable and efficient algorithm for the NUM problem based on gradient exponentiation. We show that the convergence speed of the proposed algorithm only carries a logarithmic dependence on the size of the network, so it can be implemented reliably and efficiently in massively large networks where traditional gradient methods are prohibitively slow. These theoretical results are sub-sequently validated by extensive numerical simulations showing an improvement of several order of magnitudes over standard gradient methods in large-scale networks. This contribution appeared in [31].
- In the DNS resolution process, packet losses and ensuing retransmission timeouts induce marked latencies: the current UDP-based resolution process takes up to 5 seconds to detect a loss event. In [24], [24], we find that persistent DNS connections based on TCP or TLS can provide an elegant solution to this problem. With controlled experiments on a testbed, we show that persistent DNS connections significantly reduces worst-case latency. We then leverage a large-scale platform to study the performance impact of TCP/TLS on recursive resolvers. We find that off-the-shelf software and reasonably powerful hardware can effectively provide recursive DNS service over TCP and TLS, with a manageable performance hit compared to UDP.

7.10. Privacy, Fairness, and Transparency in Online Social Medias

This section describes four contributions on privacy, fairness and transparency in online social medias

- The Facebook advertising platform has been subject to a number of controversies in the past years regarding privacy violations, lack of transparency, as well as its capacity to be used by dishonest actors for discrimination or propaganda. In this study, we aim to provide a better understanding of the Facebook advertising ecosystem, focusing on how it is being used by advertisers. We first analyze the set of advertisers and then investigate how those advertisers are targeting users and customizing ads via the platform. Our analysis is based on the data we collected from over 600 real-world users via a browser extension that collects the ads our users receive when they browse their Facebook timeline, as well as the explanations for why users received these ads. Our results reveal that users are targeted by a wide range of advertisers (e.g., from popular to niche advertisers); that a non-negligible fraction of advertisers are part of potentially sensitive categories such as news and politics, health or religion; that a significant number of advertisers employ targeting strategies that could be either invasive or opaque; and that many advertisers use a variety of targeting parameters and ad texts. Overall, our work emphasizes the need for better mechanisms to audit ads and advertisers in social media and provides an overview of the platform usage that can help move towards such mechanisms.

  This contribution appeared in [14].
- To help their users to discover important items at a particular time, major websites like Twitter, Yelp, TripAdvisor or NYTimes provide Top-K recommendations (e.g., 10 Trending Topics, Top 5 Hotels in Paris or 10 Most Viewed News Stories), which rely on crowd-sourced popularity signals to select the items. However, diferent sections of a crowd may have diferent preferences, and there is a large silent majority who do not explicitly express their opinion. Also, the crowd often consists of actors like bots, spammers, or people running orchestrated campaigns. Recommendation algorithms today largely do not consider such nuances, hence are vulnerable to strategic manipulation by small

but hyper-active user groups. To fairly aggregate the preferences of all users while recommending top-K items, we borrow ideas from prior research on social choice theory, and identify a voting mechanism called Single Trans-ferable Vote (STV) as having many of the fairness properties we desire in top-K item (s)elections. We develop an innovative mechanism to attribute preferences of silent majority which also make STV completely operational. We show the generalizability of our approach by implementing it on two diferent real-world datasets. Through extensive experimentation and comparison with state-of-the-art techniques, we show that our proposed approach provides maximum user satisfaction, and cuts down drastically on items disliked by most but hyper-actively promoted by a few users.

This contribution appeared in [17].

- The rise of algorithmic decision making led to active researches on how to define and guarantee fairness, mostly focusing on one-shot decision making. In several important applications such as hiring, however, decisions are made in multiple stage with additional information at each stage. In such cases, fairness issues remain poorly understood. In this paper we study fairness in k-stage selection problems where additional features are observed at every stage. We first introduce two fairness notions, local (per stage) and global (final stage) fairness, that extend the classical fairness notions to the k-stage setting. We propose a simple model based on a probabilistic formulation and show that the locally and globally fair selections that maximize precision can be computed via a linear program. We then define the price of local fairness to measure the loss of precision induced by local constraints; and investigate theoretically and empirically this quantity. In particular, our experiments show that the price of local fairness is generally smaller when the sensitive attribute is observed at the first stage; but globally fair selections are more locally fair when the sensitive attribute is observed at the second stage—hence in both cases it is often possible to have a selection that has a small price of local fairness and is close to locally fair.

  This contribution appeared in [21].

- Most social platforms offer mechanisms allowing users to delete their posts, and a significant fraction of users exercise this right to be forgotten. However, ironically, users' attempt to reduce attention to sensitive posts via deletion, in practice, attracts unwanted attention from stalkers specifically to those (deleted) posts. Thus, deletions may leave users more vulnerable to attacks on their privacy in general. Users hoping to make their posts forgotten face a "damned if I do, damned if I don't" dilemma. Many are shifting towards ephemeral social platform like Snapchat, which will deprive us of important user-data archival. In the form of intermittent withdrawals, we present, Lethe, a novel solution to this problem of (really) forgetting the forgotten. If the next-generation social platforms are willing to give up the uninterrupted availability of non-deleted posts by a very small fraction, Lethe provides privacy to the deleted posts over long durations. In presence of Lethe, an adversarial observer becomes unsure if some posts are permanently deleted or just temporarily withdrawn by Lethe; at the same time, the adversarial observer is overwhelmed by a large number of falsely flagged un-deleted posts. To demonstrate the feasibility and performance of Lethe, we analyze large-scale real data about users' deletion over Twitter and thoroughly investigate how to choose time duration distributions for alternating between temporary withdrawals and resurrections of non-deleted posts. We find a favorable trade-off between privacy, availability and adversarial overhead in different settings for users exercising their right to delete. We show that, even against an ultimate adversary with an uninterrupted access to the entire platform, Lethe offers deletion privacy for up to 3 months from the time of deletion, while maintaining content availability as high as 95% and keeping the adversarial precision to 20%.

  This contribution appeared in [27],

7.11. Optimization Methods

This section describes six contributions on optimization.

- In [9], we propose an interior-point method for linearly constrained – and possibly nonconvex – optimization problems. The proposed method – which we call the Hessian barrier algorithm (HBA) – combines a forward Euler discretization of Hessian Riemannian gradient flows with an Armijo backtracking step-size policy. In this way, HBA can be seen as an alternative to mirror descent (MD), and contains as special cases the affine scaling algorithm, regularized Newton processes, and several other iterative solution methods. Our main result is that, modulo a non-degeneracy condition, the algorithm converges to the problem's critical set; hence, in the convex case, the algorithm converges globally to the problem's minimum set. In the case of linearly constrained quadratic programs (not necessarily convex), we also show that the method's convergence rate is $O(1/k^{\rho})$ for some $\rho \in (0, 1]$ that depends only on the choice of kernel function (i.e., not on the problem's primitives). These theoretical results are validated by numerical experiments in standard non-convex test functions and large-scale traffic assignment problems.
- In[15], Lipschitz continuity is a central requirement for achieving the optimal $O(1/T)$ rate of convergence in monotone, deterministic variational inequalities (a setting that includes convex minimization, convex-concave optimization, nonatomic games, and many other problems). However, in many cases of practical interest, the operator defining the variational inequality may exhibit singularities at the boundary of the feasible region, precluding in this way the use of fast gradient methods that attain this optimal rate (such as Nemirovski's mirror-prox algorithm and its variants). To address this issue, we propose a novel regularity condition which we call Bregman continuity, and which relates the variation of the operator to that of a suitably chosen Bregman function. Leveraging this condition, we derive an adaptive mirror-prox algorithm which attains the optimal $O(1/T)$ rate of convergence in problems with possibly singular operators, without any prior knowledge of the degree of smoothness (the Bregman analogue of the Lipschitz constant). We also show that, under Bregman continuity, the mirror-prox algorithm achieves a $O(1/\sqrt{T})$ convergence rate in stochastic variational inequalities.
- In [23] Variational inequalities have recently attracted considerable interest in machine learning as a flexible paradigm for models that go beyond ordinary loss function minimization (such as generative adversarial networks and related deep learning systems). In this setting, the optimal O(1/t) convergence rate for solving smooth monotone variational inequalities is achieved by the Extra-Gradient (EG) algorithm and its variants. Aiming to alleviate the cost of an extra gradient step per iteration (which can become quite substantial in deep learning applications), several algorithms have been proposed as surrogates to Extra-Gradient with a *single* oracle call per iteration. In this paper, we develop a synthetic view of such algorithms, and we complement the existing literature by showing that they retain a O(1/t) ergodic convergence rate in smooth, deterministic problems. Subsequently, beyond the monotone deterministic case, we also show that the last iterate of single-call, *stochastic* extra-gradient methods still enjoys a O(1/t) local convergence rate to solutions of non-monotone variational inequalities that satisfy a second-order sufficient condition.
- In [25], we study a class of online convex optimization problems with long-term budget constraints that arise naturally as reliability guarantees or total consumption constraints. In this general setting, prior work by Mannor et al. (2009) has shown that achieving no regret is impossible if the functions defining the agent's budget are chosen by an adversary. To overcome this obstacle, we refine the agent's regret metric by introducing the notion of a "K-benchmark", i.e., a comparator which meets the problem's allotted budget over any window of length K. The impossibility analysis of Mannor et al. (2009) is recovered when $K = T$; however, for $K = o(T)$, we show that it is possible to minimize regret while still meeting the problem's long-term budget constraints. We achieve this via an online learning policy based on Cautious Online Lagrangiant Descent (COLD) for which we derive explicit bounds, in terms of both the incurred regret and the residual budget violations.
- In [26], owing to their connection with generative adversarial networks (GANs), saddle-point problems have recently attracted considerable interest in machine learning and beyond. By necessity, most theoretical guarantees revolve around convex-concave (or even linear) problems; however, making theoretical inroads towards efficient GAN training depends crucially on moving beyond

this classic framework. To make piecemeal progress along these lines, we analyze the behavior of mirror descent (MD) in a class of non-monotone problems whose solutions coincide with those of a naturally associated variational inequality - a property which we call coherence. We first show that ordinary, "vanilla" MD converges under a strict version of this condition, but not otherwise; in particular, it may fail to converge even in bilinear models with a unique solution. We then show that this deficiency is mitigated by optimism: by taking an "extra-gradient" step, optimistic mirror descent (OMD) converges in all coherent problems. Our analysis generalizes and extends the results of Daskalakis et al. (2018) for optimistic gradient descent (OGD) in bilinear problems, and makes concrete headway for establishing convergence beyond convex-concave games. We also provide stochastic analogues of these results, and we validate our analysis by numerical experiments in a wide array of GAN models (including Gaussian mixture models, as well as the CelebA and CIFAR-10 datasets).

- In [30], we develop 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.

7.12. Learning

This section describes three contributions on machine learning.

- In [12], we examine the convergence of no-regret learning in games with continuous action sets. For concreteness, we focus on learning via "dual averaging", a widely used class of no-regret learning schemes where players take small steps along their individual payoff gradients and then "mirror" the output back to their action sets. In terms of feedback, we assume that players can only estimate their payoff gradients up to a zero-mean error with bounded variance. To study the convergence of the induced sequence of play, we introduce the notion of variational stability, and we show that stable equilibria are locally attracting with high probability whereas globally stable equilibria are globally attracting with probability 1. We also discuss some applications to mixed-strategy learning in finite games, and we provide explicit estimates of the method's convergence speed.
- Resource allocation games such as the famous Colonel Blotto (CB) and Hide-and-Seek (HS) games are often used to model a large variety of practical problems, but only in their one-shot versions. Indeed, due to their extremely large strategy space, it remains an open question how one can efficiently learn in these games. In this work, we show that the online CB and HS games can be cast as path planning problems with side-observations (SOPPP): at each stage, a learner chooses a path on a directed acyclic graph and suffers the sum of losses that are adversarially assigned to the corresponding edges; and she then receives semi-bandit feedback with side-observations (i.e., she observes the losses on the chosen edges plus some others). We propose a novel algorithm, EXP3-OE, the first-of-its-kind with guaranteed efficient running time for SOPPP without requiring any auxiliary oracle. We provide an expected-regret bound of EXP3-OE in SOPPP matching the order of the best benchmark in the literature. Moreover, we introduce additional assumptions on the observability model under which we can further improve the regret bounds of EXP3-OE. We illustrate the benefit of using EXP3-OE in SOPPP by applying it to the online CB and HS games.

  This contribution appeared in [29], [49]. In an earlier article [38], we also looked at the sequential Colonel Blotto game under bandit feedback and we proposed a blackbox optimization based method to optimize the exploration distribution of the classical COMBAND algorithm.

- In [32], we study nonzero-sum hypothesis testing games that arise in the context of adversarial classification, in both the Bayesian as well as the Neyman-Pearson frameworks. We first show that these games admit mixed strategy Nash equilibria, and then we examine some interesting concentration phenomena of these equilibria. Our main results are on the exponential rates of convergence of classification errors at equilibrium, which are analogous to the well-known Chernoff-Stein lemma and Chernoff information that describe the error exponents in the classical binary hypothesis testing problem, but with parameters derived from the adversarial model. The results are validated through numerical experiments.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

Nicolas Gast obtained funding Enedis for a study on the PLC-G3 protocol (≈50k euros).

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. Grenoble INP grant

Patrick Loiseau and Bary Pradelski received a grant from the Presidence of Grenoble INP that covers half of the funding of PhD student Dimitrios Moustakas to work on dynamic matching. This PhD is done in collaboration with Univ. Zurich (Heinrich Nax), which covers the rest.

9.2.3. 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.

9.2.4. 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.

9.2.5. PEPS

Panayotis Mertikopoulos is co-PI of a PEPS I3A project: MixedGAN ("Mixed-strategy generative adversarial networks") (PI: R. Laraki, U. Dauphine).

9.2.6. Fondation Blaise Pascal

Project IAM (Informatique à la Main) funded by fondation Blaise Pascal (Jean-Marc Vincent).

9.2.7. MIAI @ Grenoble Alpes

MIAI @ Grenoble Alpes (Multidisciplinary Institute in Artificial Intelligence) is the 3IA institute of Grenoble that was selected by the government in 2019. With the MIAI institute, Patrick Loiseau is the co-holder of a chair on “Explainable and Responsible AI” of which Nicolas Gast and Bary Pradelski are also members; and Panayotis Mertikopoulos is a member of the “Optimization and Learning” chair.

9.2.8. ANR

- Nicolas Gast obtained funding from the ANR JCJC for the project REFINO. 250k euros. Duration: 4 years
- Bary Pradelski (PI), P. Mertikopoulos and P. Loiseau obtained funding from the ANR for the project ALIAS (Adaptive Learning for Interactive Agents and Systems). This is a bilateral PRCI (collaboration internationale) project joint with Singapore University of Technology and Design (SUTD). The Singapore team consists of G. Piliouras and G. Panageas.
- *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

9.3.1.1. ReDaS

Title: Reproducible Data Science

International Partner (Institution - Laboratory - Researcher):

Universidade Federal do Rio Grande do Sul (Brazil) - Industrial Engineering and Operations Research Departments - Lucas Mello Schnorr

Start year: 2019

See also: https://associatedteam.gitlabpages.inria.fr/redas

Data science builds on a variety of technique and tools that makes analysis often difficult to follow and reproduce. The goal of this project is to develop interactive, reproducible and scalable analysis workflows that provide uncertainty and quality estimators about the analysis.

9.3.1.2. International Initiatives

GENE

Title: Stochastic dynamics of large games and networks

International Partners (Institution - Laboratory - Researcher):

Universidad de Buenos Aires (Argentina) - Matthieu Jonckheere

Universidad de la Republica Uruguay (Uruguay) - Federico La Rocca

CNRS (France) - Balakrishna Prabhu

Universidad ORT Uruguay (Uruguay) - Andrés Ferragut

Duration: 2018 - 2019

Start year: 2018

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: - Mean-field games and their application to load balancing and resource allocations, - Scaling limits for centralized and decentralized load balancing strategies and implementation of practical policies for web servers farms, - Information diffusion and communication protocols in large and distributed wireless networks.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

We have hosted multiple international scientists for short (typically one-week) visits: Jonathan Newton, Paul Duetting, Jason Marden, Bruno Ziliotto

9.4.2. Visits to International Teams

- V. Danjean spent one week at Porto Allegre (Brasil) at UFRGS, hosted by Lucas M. Schnorr to work on the research subject: Tracing of multi-tasked OpenMP Application.
- A. Legrand spent 10 days at Porto Allegre (Brasil) at UFRGS, hosted by Lucas M. Schnorr to teach scientific methodology and Performance Evaluation and to work on the visual performance analysis of dynamic task-based applications.
- G. Huard visited UFRGS (Porto Alegre, Brasil) in the context of the ReDaS Inria associated team from Nov. 27th to Dec 16th along with Alexis Janon. During this visit we worked with Lucas Schnorr on several application trace analysis cases using our own custom analysis framework and leveraging UFRGS expertise on the design and conduct of practical data analysis.
- B. Pradelski was invited for seminars at several places: IHP Game Theory Seminar, Bar-Ilan University Economic Theory seminar, University of Oxford Game Theory seminar. He is also an associate member of the Oxford Man Institute.

9.4.2.1. Research Stays Abroad

P. Mertikopoulos was invited to spend a three-month research visit at the Ecole Polytechnique Fédérale de Lausanne (EPFL). He was hosted by the LIONS lab (headed by V. Cevher).

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- P. Mertikopoulos was a technical program co-chair of the 10th International Conference on NETwork Games, COntrol and OPtimisation (NetGCoop 2020).

- B. Pradelski was a technical program co-chair of the 14th Workshop on the Economics of Networks, Systems and Computation (NetEcon), colocated with ACM SIGMETRICS and EC.

10.1.1.2. Member of the Organizing Committees

- B.Gaujal organized a special day on Potential games at Gamenets (Paris)
- P. Mertikopoulos co-organized the workshop “20PoA: Twenty years of the Price of Anarchy” (Chania, Greece, July 2019).

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

- B. Gaujal: WiOpt, NeuIPS
- J. Anselmi: Valuetools, ASMTA
- P. Mertikopoulos: The 2020 French Days on Optimization and Decision Science (SMAI MODE 2020)
- A. Legrand: ISC HIGH PERFORMANCE
- N. Gast: SIGMETRICS, NeurIPS, ValueTools
- J-M. Vincent: Epew, Valuetools, Simultech
- P. Loiseau: NeurIPS, ICML, SIGMETRICS, PETS, NetEcon
- P. Mertikopoulos: ICML, NeurIPS (area chair).

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- P. Mertikopoulos serves as an associate editor for JDG (Journal of Dynamics and Games) and MCAP (Methodology and Computing in Applied Probability).
- N. Gast serves as an associate editor for Performance Evaluation and Stochastic Models.
- P. Loiseau is an associate editor for ACM Transactions on Internet Technology and IEEE Transactions on Big Data.

10.1.3.2. Reviewer - Reviewing Activities

All members of the team are active reviewers for several international journals and conferences.

10.1.4. Invited Talks

- B. Gaujal:
  - 23/01: ENS Lyon seminars (Le Pleynet) “Evolutionary games and bounded rationality”
  - 7/02: Eva Tardos seminar (Grenoble): “price of anarchy in routing games”
  - 3/05: Workshop for Eitan Altman’s 60th Birthday (Avignon) “Sturmian words at work in optimal routing”
- P. Mertikopoulos:
  - Invited instructor at the CONNECT Summer School on Machine Learning for Communications “Online learning and optimization for wireless systems”, Trinity College, Dublin
  - Invited talk at ICCOPT 2019 (2019 International Conference on Continuous Optimization), Berlin, August 2019
  - Invited talk at NPCG 2019 (Workshop on Network, Population and Congestion Games), Paris, April 2019
  - Invited talk at GDO 2019 (Workshop on Games, Dynamics and Optimization), Cluj-Napoca, April 2019
  - Invited talk at OSL 2019 (Workshop on Optimization and Statistical Learning), Les Houches, March 2019

  - Invited talk at EPFL Machine Learning Seminar, March 2019
  - Invited talk at the Criteo AI Lab, February 2019
- A. Legrand:
  - Simulation of HPC applications and predictions, Scheduling workshop, Bordeaux (27/6/19)
  - Series of talks about reproducible research: TILECS workshop, Grenoble (3/7/19); UFRGS keynote, Porto Alegre (9/10/19); SBAC-PAD conference, Campo Grande (17/10/19); Formidex, UGA (6/11/19); Doctoral school, Neuchatel (7/11/19); Inria Alumni, Paris (12/11/19)
- Nicolas Gast was invited to give a tutorial about “Mean field and refined mean field approximation” at the conference ITC.

10.1.5. Research Administration

- B. Gaujal is a member of the scientific committee of GDR-IM and a member of the council of ‘pole MSTIC’ Grenoble
- P. Mertikopoulos is a member of the steering committee (comité de liaison) of the optimization and decision theory group of the French Society for Industrial and Applied Mathematics (SMAI)
- P. Mertikopoulos is the working group coordinator, core group member and management committee (MC) representative for France in the European Network for Game Theory (GAMENET).
- P. Loiseau is the chair of the steering committee of NetEcon.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

We only list the master level teaching.

- B. Gaujal was involved in multiple courses:
  - M2 course in ENS Lyon with Panayotis Mertikopoulos : Online Optimization
  - M2 course in MPRI (Paris) with Ana Busic: performance evaluation in communication networks
  - M2 course (Ensimag) on network performance models
  - M1 exercice session (Ensimag) applied probability
- P. Mertikopoulos gave an invited PhD level course at EPFL on “Min-max optimization and variational inequalities”.
- V. Danjean was involved in INFO3 and INFO4 at Polytech Grenoble (System Architecture, Internship supervising, ...) and in M1 Info (Operationg systems and Parallel Programming course, Operating System project)
- A. Legrand was involved in multiple courses:
  - Scientific Methodology and Performance Evaluation (M2 MOSIG, UGA)
  - Scientific Methodology and Performance Evaluation (M2 Univ. Federale do Rio Grande do Sul, Porto Alegre)
  - Parallel Systems (M2 MOSIG, UGA)
  - Probability and Simulation (M1, Polytech/UGA)
  - Performance Evaluation (M1, Polytech/UGA)
  - Reproducible Research (Doctoral School MSTII, UGA)
- J. Anselmi taught in the course Probability and Simulation (M1, Polytech/UGA).
- P. Loiseau taught in the courses Probability and Simulation (M1, Polytech/UGA) and “Algorithms for data processing” (M1 INFO, UGA).
- N. Gast is responsible of the master course “Optimization under Uncertainties” (Master 2 ORCO in Grenoble).
- J.-M. Vincent teaches Probability for Informatics and Performance Evaluation at Ensimag, and Mathematics for Computer Science (1st year) and Scientific Methodology and Performance Evaluation (2nd year) at the Master of Computer Science.
- G. Huard taught the course Object Oriented Design class for the M1 INFO, UGA.

10.2.2. Supervision

Supervision of PhD students and postdocs:

- B. Jonglez (Bruno Gaujal and Martin Heusse)
- S. Plassart (Bruno Gaujal and Alain Girault)
- K. Khun (Bruno Gaujal and Nicolas Gast)
- C. Yan (Bruno Gaujal and Nicolas Gast)
- K. Antonakopoulos (P. Mertikopoulos and E. V. Belmega, ETIS/ENSEA)
- B. Roussillon (P. Mertikopoulos and P. Loiseau)
- B. Donassolo (P. Mertikopoulos and A. Legrand):
- P. Rocha Bruel (A. Legrand and Alfredo Goldman)
- T. Cornebize (A. Legrand)
- C. Heinrich (A. Legrand)
- S. Zrigui (A. Legrand and D. Trystram)
- A. Janon (G. Huard and A. Legrand)
- V. Emelianov (N. Gast and P. Loiseau)
- T. Barzolla (N. Gast with Vincent Jost and Van-Dat Cung from G-SCOP laboratory)
- M. Mendil (N. Gast)
- T. Kennouche (N. Gast)
- U. Ozeer (J-M. Vincent)
- Dong Quan Vu (P. Loiseau)
- Vera Sosnovik (O. Goga and P. Loiseau)
- Eleni Gkiouzepi (P. Loiseau)
- Lucas Leandro Nesi (A. Legrand and Lucas Mello Schnorr)
- Dimitrios Moustakas (B. Pradelski and P. Loiseau, with H. Nax from UZH)
- Simon Jantscheg (B. Pradelski and P. Loiseau, with H. Nax from UZH)

Supervision of M2 Students:

- Manal Benaissa (V. Danjean)
- Leo Gayral (Bruno Gaujal and Federica Garin)
- Kimang Khun (Bruno Gaujal and Nicolas Gast)
- Nicolas Rocher (Patrick Loiseau and Panayotis Mertikopoulos)
- Chen Yan (Nicolas Gast)
- Dimitrios Moustakas (B. Pradelski)

10.2.3. Juries

- B. Gaujal was a reviewer of the PhD Thesis of Paulin Jacquot (Ecole Polytechnique).
- V. Danjean was involved in several teaching jurys: INFO3 at Polytech Grenoble, L3 M&I, M1 Info, DU ISN and DIU EIL at UGA.
- A. Legrand was a reviewer of the PhD Thesis of Mohamad El Sayah (Univ. Franche Comté, Besançon)
- N. Gast was member of the PhD Jury of Celine Comte and Eyal Castiel.

10.3. Popularization

Patrick Loiseau wrote with Oana Goga an article “Publicité en Ligne : reprenons la main !”, that was co-published co-published by the blog Binaire (Le Monde) and The Conversation France, June 3, 2019.

10.3.1. Internal or external Inria responsibilities

- B. Gaujal is a member of the CR2 hiring committee in Grenoble.
- J.-M. Vincent is in charge of the relation Rectorat / Inria-Grenoble for the organization of scientific events (Festival of Science, Schools Visits, organization of Conference Cycles on research in CS and Applied Mathematics for teachers in Colleges)
- J.-M. Vincent is
  - Member of the national coordination of the Diplôme Inter-Universitaire “Enseigner l’Informatique au Lycée” (50 universities involved).
  - Local Head of DIU EIL ine Academy of Grenoble
  - Member of the organization of the teaching sessions for all the teachers in CS coming from abroad
  - Member of the national Commission Inter-Irem in Informatics
  - Member of the first national jury for the competitive recruitment of teachers in computer science (Capes NSI 2019-20)

10.3.2. Articles and contents

Arnaud Legrand participated in the writing of a book [39] on Reproducible Research, which aims at helping students and engineers and researchers to find efficient and accessible ways leading them to improve their reproducible research practices.

10.3.3. Education

- V. Danjean is the head of the DU ISN formation (Diplôme Universitaire Informatique et Sciences du Numérique)
- V. Danjean co-organized the new DIU EIL formation (Diplôme Inter-Universitaire Enseigner l’Informatique au Lycée). He is involved both at the national level (for the coordination and the definition of the content of this formation provided in more than 30 universities in France), and at the local level (coordination of the local teams, courses scheduling, conference organization, ...)

10.3.4. Interventions

- V. Danjean participated in “La Fête de la Science”, animating several sessions of “unplugged computer science”
- P. Loiseau co-rganized and animated a workshop “IA, éthique et société”, Forum Ecobiz Grenoble, October 2019.
- P. Loiseau participated in a debate “Ethique et numérique : quels enjeux sociétaux ?”. Festival Transfo, Grenoble, France, January 2019.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] F. HEINRICH.*Modeling, Prediction and Optimization of Energy Consumption of MPI Applications using SimGrid*, Université Grenoble Alpes, May 2019, https://tel.archives-ouvertes.fr/tel-02269894

[2] A. MARCASTEL.*Optimisation en ligne et apprentissage adaptatif pour les réseaux dans les bandes ISM*, Université de Cergy Pontoise, February 2019

[3] P. MERTIKOPOULOS.*Online optimization and learning in games: Theory and applications*, Grenoble 1 UGA - Université Grenoble Alpes, December 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02428077

[4] U. OZEER.*Autonomic Resilience of Distributed IoT Applications in the Fog*, UGA - Université Grenoble Alpes ; MSTII, December 2019

Articles in International Peer-Reviewed Journal

[5] P. ALLIEZ, R. DI COSMO, B. GUEDJ, A. GIRAULT, M.-S. HACID, A. LEGRAND, N. P. ROUGIER.*Attributing and Referencing (Research) Software: Best Practices and Outlook from Inria*, in "Computing in Science & Engineering", 2019, p. 1-14, https://arxiv.org/abs/1905.11123 [*DOI :* 10.1109/MCSE.2019.2949413], https://hal.archives-ouvertes.fr/hal-02135891

[6] J. ANSELMI.*Combining Size-Based Load Balancing with Round-Robin for Scalable Low Latency*, in "IEEE Transactions on Parallel and Distributed Systems", 2019, p. 1-3, forthcoming [*DOI :* 10.1109/TPDS.2019.2950621], https://hal.archives-ouvertes.fr/hal-02276789

[7] J. ANSELMI, J. DONCEL.*Asymptotically Optimal Size-Interval Task Assignments*, in "IEEE Transactions on Parallel and Distributed Systems", 2019, vol. 30, n^o 11, p. 2422-2433 [*DOI :* 10.1109/TPDS.2019.2920121], https://hal.archives-ouvertes.fr/hal-02318576

[8] J. ANSELMI, F. DUFOUR.*Power-of-d-Choices with Memory: Fluid Limit and Optimality*, in "Mathematics of Operations Research", 2019, p. 1-31, forthcoming, https://hal.archives-ouvertes.fr/hal-02394147

[9] I. M. BOMZE, P. MERTIKOPOULOS, W. SCHACHINGER, M. STAUDIGL.*Hessian barrier algorithms for linearly constrained optimization problems*, in "SIAM Journal on Optimization", 2019, vol. 29, p. 2100 - 2127 [*DOI :* 10.1137/18M1215682], https://hal.inria.fr/hal-02403531

[10] J. DONCEL, N. GAST, B. GAUJAL.*Discrete Mean Field Games: Existence of Equilibria and Convergence*, in "Journal of Dynamics and Games", 2019, vol. 6, n^o 3, p. 1-19, https://arxiv.org/abs/1909.01209 [*DOI :* 10.3934/JDG.2019016], https://hal.inria.fr/hal-01277098

[11] A. MARCASTEL, E.-V. BELMEGA, P. MERTIKOPOULOS, I. FIJALKOW.*Online Power Optimization in Feedback-Limited, Dynamic and Unpredictable IoT Networks*, in "IEEE Transactions on Signal Processing", 2019, vol. 67, n^o 11, p. 2987 - 3000, forthcoming [*DOI :* 10.1109/TSP.2019.2910479], https://hal.archives-ouvertes.fr/hal-02189523

[12] P. MERTIKOPOULOS, Z. ZHOU.*Learning in games with continuous action spaces and unknown payoff functions*, in "Mathematical Programming, Series A", 2019, vol. 173, n^o 1-2, p. 465-507, https://arxiv.org/abs/1608.07310 [*DOI :* 10.1007/S10107-018-1254-8], https://hal.archives-ouvertes.fr/hal-01382282

[13] X. WU, P. LOISEAU, E. HYYTIÄ.*Towards Designing Cost-Optimal Policies to Utilize IaaS Clouds with Online Learning*, in "IEEE Transactions on Parallel and Distributed Systems", 2019, vol. 14, forthcoming [*DOI :* 10.1109/TPDS.2019.2935199], https://hal.inria.fr/hal-02303480

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, p. 1-15 [*DOI :* 10.14722/NDSS.2019.23280], https://hal.archives-ouvertes.fr/hal-01959145

[15] K. Antonakopoulos, E.-V. Belmega, P. Mertikopoulos.*An adaptive mirror-prox algorithm for variational inequalities with singular operators*, in "NeurIPS 2019", Vancouver, Canada, 2019, https://hal.inria.fr/hal-02403562

[16] 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*, in "CCGrid 2019 - International Symposium in Cluster, Cloud, and Grid Computing", Larcana, Cyprus, May 2019, p. 1-10 [*DOI :* 10.1109/CCGRID.2019.00026], https://hal.inria.fr/hal-02110868

[17] A. Chakraborty, G. K. Patro, N. Ganguly, K. P. Gummadi, P. Loiseau.*Equality of Voice: Towards Fair Representation in Crowdsourced Top-K Recommendations*, in "FAT* 2019 - ACM Conference on Fairness, Accountability, and Transparency", Atlanta, United States, Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAT*), ACM, January 2019, p. 129-138 [*DOI :* 10.1145/3287560.3287570], https://hal.archives-ouvertes.fr/hal-01959135

[18] T. Cornebize, A. Legrand, F. C. Heinrich.*Fast and Faithful Performance Prediction of MPI Applications: the HPL Case Study*, in "2019 IEEE International Conference on Cluster Computing (CLUSTER)", Albuquerque, United States, 2019 IEEE International Conference on Cluster Computing (CLUSTER), September 2019 [*DOI :* 10.1109/CLUSTER.2019.8891011], https://hal.inria.fr/hal-02096571

[19] B. Donassolo, I. Fajjari, A. Legrand, P. Mertikopoulos.*Fog Based Framework for IoT Service Provisioning*, in "CCNC 2019 - IEEE Consumer Communications & Networking Conference", Las Vegas, United States, IEEE, January 2019, p. 1-6 [*DOI :* 10.1109/CCNC.2019.8651835], https://hal.inria.fr/hal-01859695

[20] J. Doncel, N. Gast, M. Tribastone, M. Tschaikowski, A. Vandin.*UTOPIC: Under-Approximation Through Optimal Control*, in "QEST 2019 - 16th International Conference on Quantitative Evaluation of SysTems", Glasgow, United Kingdom, Springer, September 2019, p. 277-291 [*DOI :* 10.1007/978-3-030-30281-8_16], https://hal.inria.fr/hal-02283189

[21] V. Emelianov, G. Arvanitakis, N. Gast, K. P. Gummadi, P. Loiseau.*The Price of Local Fairness in Multistage Selection*, in "IJCAI-2019 - Twenty-Eighth International Joint Conference on Artificial Intelligence", Macao, France, International Joint Conferences on Artificial Intelligence Organization, May 2019, p. 5836-5842, https://arxiv.org/abs/1906.06613 [*DOI :* 10.24963/IJCAI.2019/809], https://hal.inria.fr/hal-02145071

[22] B. Gaujal, A. Girault, S. Plassart.*A Linear Time Algorithm for Computing Off-line Speed Schedules Minimizing Energy Consumption*, in "MSR 2019 - 12ème Colloque sur la Modélisation des Systèmes Réactifs", Angers, France, November 2019, p. 1-14, https://hal.archives-ouvertes.fr/hal-02372136

[23] Y.-G. Hsieh, F. Iutzeler, J. Malick, P. Mertikopoulos.*On the convergence of single-call stochastic extra-gradient methods*, in "NeurIPS 2019", Vancouver, Canada, 2019, https://arxiv.org/abs/1908.08465 - 27 pages, 3 figures, https://hal.inria.fr/hal-02403555

[24] B. JONGLEZ, S. BIRBALTA, M. HEUSSE.*Persistent DNS connections for improved performance*, in "NETWORKING 2019 - IFIP Networking 2019", Warsaw, Poland, May 2019, p. 1-2, https://hal.inria.fr/hal-02149978

[25] N. LIAKOPOULOS, A. S. DESTOUNIS, G. PASCHOS, T. SPYROPOULOS, P. MERTIKOPOULOS.*Cautious regret minimization: Online optimization with long-term budget constraints*, in "ICML 2019 - 36th International Conference on Machine Learning", Long Beach, United States, June 2019, p. 1-9, https://hal.inria.fr/hal-02405753

[26] 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*, in "ICLR 2019 - 7th International Conference on Learning Representations", New Orleans, United States, May 2019, p. 1-23, https://hal.inria.fr/hal-02111937

[27] 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

[28] 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 - 22nd Conference on Innovation in Clouds, Internet and Networks", Paris, France, IEEE, February 2019, p. 230-232 [*DOI :* 10.1109/ICIN.2019.8685909], https://hal.archives-ouvertes.fr/hal-01979686

[29] D. QUAN VU, P. LOISEAU, A. SILVA, L. TRAN-THANH.*Path Planning Problems with Side Observations—When Colonels Play Hide-and-Seek*, in "AAAI 2020 - Thirty-Fourth AAAI Conference on Artificial Intelligence", New-York, United States, February 2020, p. 1-15, https://hal.inria.fr/hal-02375789

[30] M. STAUDIGL, P. MERTIKOPOULOS.*Convergent Noisy forward-backward-forward algorithms in non-monotone variational inequalities*, in "LSS 2019 - 15th IFAC Symposium on Large Scale Complex Systems: Theory and Applications", Delft, Pays-Bas, May 2019, p. 120-125 [*DOI :* 10.1016/J.IFACOL.2019.06.021], https://hal.inria.fr/hal-02405750

[31] L. VIGNERI, G. PASCHOS, P. MERTIKOPOULOS.*Large-Scale Network Utility Maximization: Countering Exponential Growth with Exponentiated Gradients*, in "INFOCOM 2019 - IEEE International Conference on Computer Communications", Paris, France, IEEE, April 2019, p. 1630-1638 [*DOI :* 10.1109/INFOCOM.2019.8737600], https://hal.inria.fr/hal-02405759

[32] S. YASODHARAN, P. LOISEAU.*Nonzero-sum Adversarial Hypothesis Testing Games*, in "NeurIPS 2019 - Thirty-third Conference on Neural Information Processing Systems", Vancouver, Canada, 2019, p. 1-23, https://hal.inria.fr/hal-02299451

Conferences without Proceedings

[33] E. AGULLO, A. BUTTARI, A. GUERMOUCHE, A. LEGRAND, I. MASLIAH, L. STANISIC.*Simulation of a Sparse Direct Solver on Heterogeneous Systems using Starpu and Simgrid*, in "CSE 2019 - SIAM Conference on Computational Science and Engineering", Spokane, United States, SIAM, February 2019, https://hal.inria.fr/hal-02073725

[34] J. Assunção, J.-M. Vincent, P. Fernandes.*Piecewise Aggregation for HMM fitting. A pre-fitting model for seamless integration with time series data*, in "SEKE 2019 - 31st International Conference on Software Engineering and Knowledge Engineering", Lisbon, Portugal, July 2019, p. 729-734 [*DOI :* 10.18293/SEKE2019-185], https://hal.archives-ouvertes.fr/hal-02409589

[35] T.-E. Kennouche, F. Cadoux, N. Gast, B. Vinot.*ASGriDS: Asynchronous Smart-Grids Distributed Simulator*, in "APPEEC 2019 - 11th IEEE PES Asia-Pacific Power and Energy Engineering Conference", Macao, Macau SAR China, IEEE, December 2019, p. 1-5, https://hal.archives-ouvertes.fr/hal-02384051

[36] A. Legrand, D. Trystram, S. Zrigui.*Adapting Batch Scheduling to Workload Characteristics: What can we expect From Online Learning?*, in "IPDPS 2019 - 33rd IEEE International Parallel & Distributed Processing Symposium", Rio de Janeiro, Brazil, IEEE, May 2019, p. 686-695 [*DOI :* 10.1109/IPDPS.2019.00077], https://hal.archives-ouvertes.fr/hal-02044903

[37] A. Marcastel, E.-V. Belmega, P. Mertikopoulos, I. Fijalkow.*Gradient-free Online Resource Allocation Algorithms for Dynamic Wireless Networks*, in "SPAWC 2019 - 20th IEEE International Workshop on Signal Processing Advances in Wireless Communications", Cannes, France, IEEE, July 2019, p. 1-4 [*DOI :* 10.1109/SPAWC.2019.8815409], https://hal.archives-ouvertes.fr/hal-02189108

[38] D. Quan Vu, P. Loiseau, A. Silva.*Combinatorial Bandits for Sequential Learning in Colonel Blotto Games*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://arxiv.org/abs/1909.04912 , https://hal.archives-ouvertes.fr/hal-02283535

Scientific Books (or Scientific Book chapters)

[39] L. Desquilbet, S. Granger, B. Hejblum, A. Legrand, P. Pernot, N. P. Rougier, E. de Castro Guerra, M. Courbin-Coulaud, L. Duvaux, P. Gravier, G. Le Campion, S. Roux, F. Santos.*Towards reproducible research : Evolve your practices*, Unité régionale de formation à l'information scientifique et technique de Bordeaux, May 2019, p. 1-161, https://hal.archives-ouvertes.fr/hal-02144142

Research Reports

[40] T. Cornebize, A. Legrand.*DGEMM performance is data-dependent*, Université Grenoble Alpes ; Inria ; CNRS, December 2019, n° RR-9310, https://arxiv.org/abs/1912.05381 , https://hal.inria.fr/hal-02401760

[41] B. Gaujal, A. Girault, S. Plassart.*A Discrete Time Markov Decision Process for Energy Minimization Under Deadline Constraints*, Grenoble Alpes ; Inria Grenoble Rhône-Alpes, Université de Grenoble, December 2019, n° RR-9309, 46, https://hal.inria.fr/hal-02391948

[42] B. Gaujal, A. Girault, S. Plassart.*Exploiting Job Variability to Minimize Energy Consumption under Real-Time Constraints*, Inria Grenoble Rhône-Alpes, Université de Grenoble ; Université Grenoble - Alpes, November 2019, n° RR-9300, 23, https://hal.inria.fr/hal-02371742

[43] B. Gaujal, A. Girault, S. Plassart.*Feasibility of on-line speed policies in real-time systems*, Inria Grenoble Rhône-Alpes, Université de Grenoble ; Univ. Grenoble Alpes, November 2019, n° RR-9301, 38, https://hal.inria.fr/hal-02371996

Software

[44] S. ARCHIPOFF, C. AUGONNET, O. AUMAGE, G. BEAUCHAMP, B. BRAMAS, A. BUTTARI, A. CASSAGNE, J. CLET-ORTEGA, T. COJEAN, N. COLLIN, V. DANJEAN, A. DENIS, L. EYRAUD-DUBOIS, N. FURMENTO, S. HENRY, A. HUGO, M. JUHOOR, A. JUVEN, M. KERYELL-EVEN, Y. KHORSI, T. LAMBERT, E. LERIA, B. LIZÉ, M. MAKNI, S. NAKOV, R. NAMYST, L. NESI LUCAS, J. PABLO, D. PASQUALINOTTO, S. PITOISET, N. QUÔC-DINH, C. ROELANDT, C. SAKKA, C. SALINGUE, L. MELLO SCHNORR, M. SERGENT, A. SIMONET, L. STANISIC, S. BÉRANGÈRE, F. TESSIER, S. THIBAULT, V. BRICE, L. VILLEVEYGOUX, P.-A. WACRENIER.*StarPU*, January 2020, Version : 1.3.3
[SWH-ID : swh:1:dir:b6e19d99449a78805e7a55a341fbaba2bc431973], Software, https://hal.inria.fr/hal-02443512

Other Publications

[45] R. I. BOT, P. MERTIKOPOULOS, M. STAUDIGL, P. T. VUONG.*Forward-backward-forward methods with variance reduction for stochastic variational inequalities*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02405776

[46] B. DONASSOLO, I. FAJJARI, A. LEGRAND, P. MERTIKOPOULOS.*Load Aware Provisioning of IoT Services on Fog Computing Platform*, March 2019, working paper or preprint, https://hal.inria.fr/hal-02056743

[47] B. JONGLEZ, S. BIRBALTA, M. HEUSSE.*Persistent DNS connections for improved performance*, May 2019, 1, NETWORKING 2019 - IFIP Networking 2019, Poster, https://hal.inria.fr/hal-02149975

[48] D. QUAN VU, P. LOISEAU, A. SILVA.*Approximate Equilibria in Non-constant-sum Colonel Blotto and Lottery Blotto Games with Large Numbers of Battlefields*, October 2019, https://arxiv.org/abs/1910.06559 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02315698

[49] D. QUAN VU, P. LOISEAU, A. SILVA, L. TRAN-THANH.*Colonel Blotto and Hide-and-Seek Games as Path Planning Problems with Side Observations*, May 2019, https://arxiv.org/abs/1905.11151 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02139519

[50] S. ZRIGUI, R. Y. DE CAMARGO, D. TRYSTRAM, A. LEGRAND.*Improving the Performance of Batch Schedulers Using Online Job Size Classification*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02334116

References in notes

[51] 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

[52] 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

[53] 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

[54] 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

[55] 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

[56] 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

[57] 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

[58] 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

[59] 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

[60] L. BORTOLUSSI, R. LANCIANI.*Fluid Model Checking of Timed Properties*, in "Formal Modeling and Analysis of Timed Systems", Springer International Publishing, 2015

[61] 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

[62] 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

[63] 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

[64] 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

[65] 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

[66] 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

[67] K. Coulomb, M. Faverge, J. Jazeix, O. Lagrasse, J. Marcoueille, P. Noisette, A. Redondy, C. Vuchener. *Visual trace explorer (ViTE)*, October, 2009

[68] J. Doncel, N. Gast, B. Gaujal. *Mean-Field Games with Explicit Interactions*, February 2016, https://hal.inria.fr/hal-01277098

[69] 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

[70] 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

[71] 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

[72] D. Fudenberg, D. K. Levine. *The Theory of Learning in Games*, Economic learning and social evolution, MIT Press, Cambridge, MA, 1998, vol. 2

[73] R. M. Fujimoto. *Parallel Discrete Event Simulation*, in "Commun. ACM", October 1990, vol. 33, n° 10, http://doi.acm.org/10.1145/84537.84545

[74] N. Gast, B. Gaujal. *Markov chains with discontinuous drifts have differential inclusion limits*, in "Performance Evaluation", 2012, vol. 69, n° 12, p. 623-642 [*DOI :* 10.1016/j.peva.2012.07.003], https://hal.inria.fr/hal-00787999

[75] 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° 9, p. 2266–2280 [*DOI :* 10.1109/TAC.2012.2186176], https://hal.inria.fr/hal-00787996

[76] 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

[77] 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

[78] J. Gonzalez, J. Gimenez, 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

[79] M. Heath, J. Etheridge. *Visualizing the performance of parallel programs*, in "IEEE software", 1991, vol. 8, n° 5

[80] 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

[81] L. HU, J.-Y. LE BOUDEC, M. VOJNOVIĆ.*Optimal channel choice for collaborative ad-hoc dissemination*, in "INFOCOM, 2010 Proceedings IEEE", IEEE, 2010

[82] 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

[83] T. G. KURTZ.*Approximation of population processes*, SIAM, 1981, vol. 36

[84] 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

[85] 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

[86] 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

[87] 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

[88] 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

[89] 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

[90] 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

[91] 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

[92] 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

[93] M. L. PUTERMAN.*Markov decision processes: discrete stochastic dynamic programming*, John Wiley & Sons, 2014

[94] 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

[95] W. H. SANDHOLM.*Population Games and Evolutionary Dynamics*, Economic learning and social evolution, MIT Press, Cambridge, MA, 2010

[96] W. H. SANDHOLM, M. STAUDIGL.*A Sample Path Large Deviation Principle for a Class of Population Processes*, in "arXiv preprint arXiv:1511.07897", 2015

[97] 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

[98] 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

[99] 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

[100] 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

[101] 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

[102] P. VELHO, L. MELLO 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

[103] 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

[104] F. WOLF, B. MOHR.*Automatic performance analysis of hybrid MPI/OpenMP applications*, in "Journal of Systems Architecture", 2003, vol. 49, n^o 10-11

[105] 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

[106] 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

[107] 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

[108] 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

[109] 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 CENTERS
Grenoble - Rhône-Alpes
Sophia Antipolis - Méditerranée

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]
Vincent Roca [Team leader, Inria, Researcher, HDR]
Nataliia Bielova [Inria, Researcher, since October 2019]
Cédric Lauradoux [Inria, Researcher]
Daniel Le Métayer [Inria, Senior Researcher, HDR]

Faculty Members

Antoine Boutet [INSA Lyon, Associate Professor]
Mathieu Cunche [INSA Lyon, Associate Professor]

Technical Staff

Adrien Baud [Inria, Engineer, from Oct 2019]

PhD Students

Supriya Sreekant Adhatarao [Inria, PhD Student]
Coline Boniface [Univ Grenoble Alpes, PhD Student]
Guillaume Celosia [INSA Lyon, PhD Student]
Imane Fouad [Inria, Phd Student]
Clement Henin [Ministère de l'Ecologie, de l'Energie, du Développement durable et de la Mer, PhD Student]
Theo Jourdan [INSERM, PhD Student]
Raouf Kerkouche [Univ Grenoble Alpes, PhD Student]
Victor Morel [Inria, PhD Student]
Mathieu Thiery [Inria, PhD Student]
Michael Toth [Inria, Phd Student]

Post-Doctoral Fellows

Raul Pardo Jimenez [Inria, Post-Doctoral Fellow, until Jul 2019]
Celestin Matté [Inria, Post-Doctoral Fellow]

Visiting Scientists

Jeremie Decouchant [Université du Luxembourg, Oct 2019]
Gergely Acs [Budapest University of Technology and Economics, Jul 2019]

Administrative Assistant

Helen Pouchot-Rouge-Blanc [Inria, Administrative Assistant]

2. Overall Objectives

2.1. Context

Since its creation in 2014, the PRIVATICS project-team focusses on privacy protection in the digital world. It includes, on one side, activities that aim at understanding the domain and its evolution, both from theoretical and practical aspects, and, on the other side, activities that aim at designing privacy-enhancing tools and systems. The approach taken in PRIVATICS is fundamentally inter-disciplinary and covers theoretical, legal, economical, sociological and ethical aspects by the means of enriched collaborations with the members of these disciplines.

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

PRIVATICS members have written several position documents for policy makers: a report on facial recognition, algorithmic decision-making, pseudonymisation and a white book on cybersecurity.

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. Wombat

Wi-Fi tracking system for testing and demonstrational purpose

KEYWORDS: Wi-Fi - Privacy - Multimodal tracking of human activity - Wireless network

FUNCTIONAL DESCRIPTION: Wombat is a fully functional Wi-Fi tracking platform supporting three main features: collection, storage/processing, query/output. These three features are implemented through a distributed infrastructure composed of:

Sensor nodes: small devices with wireless monitoring capabilities. They collect information sent on wireless channels and forward it to the server. Central server: the central entity of the system. It receives data sent by sensor nodes and then stores it in an internal data structure. It is also in charge of answering queries related to the stored data.

To ensure communication between the sensor nodes and the server, the Wombat system relies on a wired network (Ethernet). In addition, Wombat can be enriched with a user interface and an opt-out node:

User interface: a device in charge of displaying detailed information about one or several tracked devices (see figure below). The device to display can be specified manually by its MAC address or through proximity detection. Opt-out node: an element in charge of implementing an opt-out mechanism for users refusing to be tracked by the system.

The system is made to work on a dedicated network (the server includes a DHCP server). Nodes can be switched off at any time (they function in read-only mode to be crash-proof).

- Partner: Insa de Lyon
- Contact: Mathieu Cunche
- URL: https://github.com/Perdu/wombat

5.3. Cookie glasses

KEYWORDS: GDPR - Cookie - Consent

SCIENTIFIC DESCRIPTION: In the paper Do Cookie Banners Respect my Choice? Measuring Legal Compliance of Banners from IAB Europe's Transparency and Consent Framework, we show that Consent Management Providers (CMPs) of IAB Europe's Transparency & Consent Framework (TCF) do not always respect user's choice. This extension allows users to verify that their consent is stored appropriately by themselves.

This extension for Firefox and Chrome queries CMPs of IAB Europe's TCF in the same position as a third-party advertiser, making it possible to see consent set by CMPs in real time. In other words, you can see whether consent registered by cookie banners is actually the consent you gave. Will only work with cookie banners of IAB Europe's TCF.

We also added a functionality to manually decode a so-called "consent string" of the framework.

- Participants: Célestin Matte and Nataliia Bielova
- Contact: Alain Prette

5.4. BELL

Browser fingerprinting via Extensions and Login-Leaks

KEYWORDS: Browser Extensions - Security and Privacy in Web Services - Social Networks Security and Privacy

FUNCTIONAL DESCRIPTION: Recent studies show that users can be tracked based on their web browser properties. This software is designed to conduct an experiment on such kinds of user tracking. In this experiment, we demonstrate that a Web user can also be tracked by

- her browser extensions (such as AdBlock, Pinterest, or Ghostery), and

- the websites she has logged in (such as Facebook, Gmail, or Twitter).

In the experiment, we collect user's browser fingerprint, together with the browser extensions installed and a list of websites she has logged in. We only collect anonymous data during the experiment (more details in our Privacy Policy [0]), we will securely store the data on an Inria server, use it only for research purposes and not share it with anyone outside of Inria.

- Contact: Gabor Gulyas
- URL: https://extensions.inrialpes.fr/

5.5. SWIF-codec

An open-source sliding window FEC codec

KEYWORD: Error Correction Code

FUNCTIONAL DESCRIPTION: This development is done in the context of the "Coding for Efficient Network Communications" IRTF Research Group (NWCRG, [https://datatracker.ietf.org/rg/nwcrg]) and IETF hackathon.

[0] https://extensions.inrialpes.fr/privacy.php

This work has strong relationships with the Generic API I-D [https://datatracker.ietf.org/doc/draft-roca-nwcrg-generic-fec-api/] and RFC 8681 on RLC codes [https://www.rfc-editor.org/rfc/rfc8681] as examples of sliding window codes.

- Authors: Vincent Roca, Cédric Adjih, Oumaima Attia and François Michel
- Contact: Vincent Roca
- URL: https://github.com/irtf-nwcrg/swif-codec

6. New Results

6.1. Differential Inference Testing

Participant: Claude Castelluccia.

In order to protect individuals' privacy, data have to be "well-sanitized" before sharing them, i.e. one has to remove any personal information before sharing data. However, it is not always clear when data shall be deemed well-sanitized. In [10], we argue that the evaluation of sanitized data should be based on whether the data allows the inference of sensitive information that is specific to an individual, instead of being centered around the concept of re-identification. We propose a framework to evaluate the effectiveness of different sanitization techniques on a given dataset by measuring how much an individual's record from the sanitized dataset influences the inference of his/her own sensitive attribute. Our intent is not to accurately predict any sensitive attribute but rather to measure the impact of a single record on the inference of sensitive information. We demonstrate our approach by sanitizing two real datasets in different privacy models and evaluate/compare each sanitized dataset in our framework.

6.2. Analyse des impacts de la reconnaissance faciale - Quelques éléments de méthode (in French)

Participants: Claude Castelluccia, Daniel Le Métayer.

Significant technical progress has been made in recent years in the field of image processing, in particular in facial recognition. The deployments and experiments of this type of systems are more and more numerous. However, opinions differ on their use, especially in public space. Noting the lack of consensus on a technology that can have a significant impact on society, many organizations have alerted public opinion and asked for a public debate on the subject. We believe that such a debate is indeed necessary. However, for it to be truly productive, it is necessary to be able to confront the arguments in a rigorous manner while avoiding, as far as possible, the preconceptions, and by distinguishing established facts from assumptions or opinions. The purpose of this document [14] is precisely to help put the terms of the debate on solid foundations. It is therefore not a question here of taking a position on facial recognition in general nor of providing an exhaustive review of its applications but of proposing elements of method, illustrated by a few examples. We first present a quick overview of the applications of facial recognition before detailing the reasons that make it a particularly sensitive subject, emphasizing in particular the risks linked to a possible generalization of its use. We then present an incremental, comparative and rigorous approach to analyze the impacts of a facial recognition system.

6.3. Towards a generic framework for black-box explanation methods

Participants: Daniel Le Métayer, Clément Hénin.

Explainability has generated increased interest during the last decade because the most accurate ML techniques often lead to opaque Algorithmic Decision Systems (ADS) and opacity is a major source of mistrust. Indeed, even if explanations are not a panacea, they can play a key role, not only to enhance trust in the system, but also to allow its users to better understand its outputs and therefore to make a better use of it. In addition, they are necessary to make it possible to challenge the decisions resulting from an ADS. Explanations can take different forms, they can target different types of users and different types of methods can be used to produce them. Our work on this topic [15] focuses on a category of methods, called "black-box", that do not make any assumption about the availability of the code of the ADS or its implementation techniques. Our first contribution is to bring to light a common structure for Black-box Explanation Methods and to define a generic framework allowing us to compare and classify different approaches. This framework consists of three components, called respectively Sampling, Generation and Interaction. Beyond its interest as a systematic presentation of the state of the art, we believe that this framework can also provide new insights for the design of new explanation systems. For example, it may suggest new combinations of Sampling and Generation components or criteria to choose the most appropriate combination to produce a given type of explanation.

6.4. A generic information and consent framework for the IoT

Participants: Daniel Le Métayer, Mathieu Cunche, Victor Morel.

The development of the Internet of Things (IoT) raises specific privacy issues especially with respect to information and consent. People are generally unaware of the devices collecting data about them and do not know the organizations operating them. Solutions such as stickers or wall signs are not effective information means in most situations. As far as consent is concerned, individuals do not have simple means to express and communicate it to the entities collecting data. Furthermore, the devices used to collect data in IoT environments have scarce resources; some of them do not have any user interface, are battery-operated or operate passively. The Working Party 29 (now "European Data Protection Board") advocates the design of new consent mechanisms, such as "privacy proxies", on the devices themselves. Starting from their recommendations, we have defined general requirements that have to be met to ensure that information and consent are managed in a manner that is satisfactory both for data subjects and for data controllers. We have shown in [8] how these requirements can be implemented in different situations, in particular through declaration registers and beacons. Depending on the context and the types of devices involved, not all technical options are always possible. In order to provide guidance to IoT system designers, we have outlined the main choice factors in the design pace are illustrated the framework with several challenging case studies. We have also implemented a Proof of Concept prototype implementation of these techniques.

6.5. Analysis of privacy policies to enhance informed consent

Participant: Daniel Le Métayer.

A privacy policy language must meet a number of requirements to be able to express the valid consent of the data subject for the processing of their personal data. For example, under the GDPR, valid consent must be freely given, specific, informed and unambiguous. Therefore, the language must be endowed with a formal semantics in order to avoid any ambiguity about the meaning of a privacy policy. However, the mere existence of a semantics does not imply that DSs properly understand the meaning of a policy and its potential consequences. One way to enhance the understanding of the data subjects is to provide them information about the potential risks related to a privacy policy. This is in line with Recital 39 of the GDPR which stipulates that data subjects should be "made aware of the risks, rules, safeguards and rights in relation to the processing of personal data and how to exercise their rights in relation to such processing". To address this need, we have defined a language in [11], called PILOT, meeting these requirements and shown its benefits to define precise privacy policies and to highlight the associated privacy risks. In order to automatically answer questions related to privacy risks, we use the verification tool SPIN and the modeling language PROMELA. Risk properties are encoded in Linear Temporal Logic properties that can be automatically checked by SPIN.

6.6. Understanding algorithmic decision-making: Opportunities and challenges, Study for the European Parliament (STOA)

Participants: Claude Castelluccia, Daniel Le Métayer.

Algorithms are far from being a recent invention but they are increasingly involved in systems used to support decision making. Algorithmic Decision Systems (ADS) often rely on the analysis of large amounts of personal data to infer correlations or, more generally, to derive information deemed useful to make decisions. Humans may have a role of varying degree in the decision making and may even be completely out of the loop in entirely automated systems. In many situations, the impact of the decision on people can be significant: access to credit, employment, medical treatment, judicial sentences, etc. Entrusting ADS to make or to influence such decisions raises a variety of issues that differ in nature such as ethical, political, legal, technical, etc. and great care must be taken to analyse and address these issues. If they are neglected, the expected benefits of these systems may be offset by the variety of risks for individuals (discrimination, unfair practices, loss of autonomy, etc.), the economy (unfair practices, limited access to markets, etc.) and society as a whole (manipulation, threat to democracy, etc.).

We have written a report for the European Parliament reviewing the opportunities and risks related to the use of ADS. We present existing options to reduce these risks and explain their limitations. We sketch some recommendations to benefit from the tremendous possibilities of ADS while limiting the risks related to their use. Beyond providing an up-to-date and systematic review of the situation, the report gives a precise definition of a number of key terms and an analysis of their differences. This helps clarify the debate. The main focus of the report is the technical aspects of ADS. However, other legal, ethical and social dimensions are considered to broaden the discussion.

6.7. Saving Private Addresses: An Analysis of Privacy Issues in the Bluetooth-Low-Energy Advertising Mechanism

Participants: Mathieu Cunche, Guillaume Celiosa.

The Bluetooth Low Energy (BLE) protocol is being included in a growing number of connected objects such as fitness trackers and headphones. As part of the service discovery mechanism of BLE, devices announce themselves by broadcasting radio signals called advertisement packets that can be collected with off-the-shelf hardware and software. To avoid the risk of tracking based on those messages, BLE features an address randomization mechanism that substitutes the device address with random temporary pseudonyms,called Private addresses. We analyze the privacy issues associated with the advertising mechanism of BLE, leveraging a large dataset of advertisement packets collected in the wild. First, we identified in [7] that some implementations fail at following the BLE specifications on the maximum lifetime and the uniform distribution of random identifiers. Furthermore, we found that the payload of the advertisement packet can hamper the randomization mechanism by exposing counters and static identifiers. In particular, we discovered that advertising data of Apple and Microsoft proximity protocols can be used to defeat the address randomization scheme. Finally, we discuss how some elements of advertising data can be leveraged to identify the type of device, exposing the owner to inventory attacks

6.8. Fingerprinting Bluetooth-Low-Energy Devices Based on the Generic Attribute Profile

Participants: Mathieu Cunche, Guillaume Celiosa.

Bluetooth Low Energy (BLE) is a short range wireless technology included in many consumer devices such as smartphones, earphones and wristbands. As part of the Attribute (ATT) protocol, discoverable BLE devices expose a data structure called Generic Attribute (GATT) profile that describes supported features using concepts of services and characteristics. This profile can be accessed by any device in range and can expose users to privacy issues. We study how the GATT profile can be used to create a fingerprint that can be exploited to circumvent anti-tracking features of the BLE standard (i.e. MAC address randomization). Leveraging a dataset of more than 13000 profiles, we analyze the potential of this fingerprint and show that it can be used to uniquely identify a number of devices. We also shed light in [6] on several issues where GATT profiles can be mined to infer sensitive information that can impact privacy of users. Finally, we suggest solutions to mitigate those issues.

6.9. Privacy implications of switching ON a light bulb in the IoT world

Participants: Vincent Roca, Mathieu Thiery.

The number of connected devices is increasing every day, creating smart homes and shaping the era of the Internet of Things (IoT), and most of the time, end-users are unaware of their impacts on privacy. We analyze in [23] the ecosystem around a Philips Hue smart white bulb in order to assess the privacy risks associated to the use of different devices (smart speaker or button) and smartphone applications to control it. We show that using different techniques to switch ON or OFF this bulb has significant consequences regarding the actors involved (who mechanically gather information on the user's home) and the volume of data sent to the Internet (we measured differences up to a factor 100, depending on the control technique we used). Even when the user is at home, these data flows often leave the user's country, creating a situation that is neither privacy friendly (and the user is most of the time ignorant of the situation), nor sovereign (the user depends on foreign actors), nor sustainable (the extra energetic consumption is far from negligible). We therefore advocate a complete change of approach, that favors local communications whenever sufficient.

6.10. Security Analysis of Subject Access Request Procedures How to authenticate data subjects safely when they request for their data

Participants: Cédric Lauradoux, Coline Boniface.

With the GDPR in force in the EU since May 2018, companies and administrations need to be vigilant about the personal data they process. The new regulation defines rights for data subjects and obligations for data controllers but it is unclear how subjects and controllers interact concretely. In [4], we try to answer two critical questions: is it safe for a data subject to exercise the right of access of her own data? When does a data controller have enough information to authenticate a data subject? To answer these questions, we have analyzed recommendations of Data Protection Authorities and authentication practices implemented in popular websites and third-party tracking services. We observed that some data controllers use unsafe or doubtful procedures to authenticate data subjects. The most common flaw is the use of authentication based on a copy of the subject's national identity card transmitted over an insecure channel. We define how a data controller should react to a subject's request to determine the appropriate procedures to identify the subject and her data. We provide compliance guidelines on data access response procedures.

6.11. Plausible Deniability for Practical Privacy-Preserving Live Streaming

Participant: Antoine Boutet.

Video consumption is one of the most popular Internet activities worldwide. The emergence of sharing videos directly recorded with smartphones raises important privacy concerns. In this work we propose P3LS, the first practical privacy-preserving peer-to-peer live streaming system. To protect the privacy of its users, P3LS relies on k-anonymity when users subscribe to streams, and on plausible deniability for the dissemination of video streams. Specifically, plausible deniability during the dissemination phase ensures that an adversary is never able to distinguish a user's stream of interest from the fake streams from a statistical analysis (i.e., using an analysis of variance). We exhaustively evaluate P3LS and show that adversaries are not able to identify the real stream of a user with very high confidence. Moreover, P3LS consumes 30% less bandwidth than the standard k-anonymity approach where nodes fully contribute to the dissemination of k streams.

6.12. Protecting motion sensor data against sensitive inferences through an adversarial network approach

Participants: Antoine Boutet, Théo Jourdan.

With the widespread development of the quantified self movement, more and more motion sensor data are captured and transmitted through the intermediary of smartphones. However, granting to applications a direct access to sensor data expose users to many privacy risks, including in particular the possibility of inferring their activities and transportation mode to more sensitive inferences such as their demographic attributes or even mobility deficiency. In this work, we propose a privacy-preserving scheme to protect sensor data for activity recognition while at the same time preventing unwanted sensitive inferences on specific information. To achieve this objective, we leverage on the powerful framework of generative adversarial networks (GANs) to sanitize the sensor data. More precisely in our framework three neural networks are jointly trained, a generator that aim at sanitizing the data given at input as well two discriminators that try to infer respectively the sensitive attributes and the current activity of the user. By letting these neural networks compete against each other, the mechanism improves the protection while providing a good accuracy in terms of activity recognition and limiting sensitive inferences on specified attributes. Preliminary results demonstrate that the approach is promising in terms of achieving a good utility-privacy trade-off.

6.13. Inria white book on Cybersecurity: Current challenges and Inria's research directions

Participant: Vincent Roca.

This book provides an overview of research areas in cybersecurity, illustrated by contributions from Inria teams. The first step in cybersecurity is to identify threats and define a corresponding attacker model. Threats, including malware, physical damage or social engineering, can target the hardware, the network, the operating system, the applications, or the users themselves.

Then, detection and protection mechanisms must be designed to defend against these threats. One of the core mechanisms is cryptography, in order to ensure the confidentiality and integrity of data. These primitives must be the object of continuous cryptanalysis to ensure the highest level of security. However, secure cryptographic primitives alone are not sufficient for secure communications and services: cryptographic protocols, implementing richer interactions on top of the primitives, are needed. These protocols are distributed systems. Ensuring that they achieve their goals in the presence of an adversary requires the use of formal verification techniques, which have been extremely successful in this field.

Additional security services, such as authentication and access control, are needed to enforce a security policy. These security services, usually provided by the operating system or the network devices, can themselves be attacked and sometimes bypassed. Therefore, activities on the information system are monitored in order to detect any violation of the security policy. Finally, as attacks can spread extremely fast, the system must react automatically or at least reconfigure itself to avoid propagating attacks.

Privacy has also become an intrinsic part of cybersecurity. Privacy has its own properties, techniques, and methodology. Moreover, the study of privacy often requires to take legal, economical, and sociological aspects into account.

All these security mechanisms need to be carefully integrated in security-critical applications. These applications include traditional safety-critical applications that are becoming increasingly connected and therefore more vulnerable to security attacks, as well as new infrastructures running in the cloud or connected to a multitude of Things (IoT).

6.14. Inspect what your location history reveals about you - Raising user awareness on privacy threats associated with disclosing his location data

Participant: Antoine Boutet.

Location is one of the most extensively collected personal data on mobile by applications and third-party services. However, how the location of users is actually processed in practice by the actors of targeted advertising ecosystem remains unclear. Nonetheless, these providers have a strong incentive to create very detailed profile of users to better monetize the collected data. End users are usually not aware about the strength and wide range of inference that can be performed from their mobility traces. In this work, users interact with a web-based application to inspect their location history and to discover the inferential power of this kind of data. Moreover to better understand the possible countermeasures, users can apply a sanitization to protect their data and visualize the impact on both the mobility traces and the associated inferred information. The objective of this work is to raise the user awareness on the profiling capabilities and the privacy threats associated with disclosing his location data as well as how sanitization mechanisms can be efficient to mitigate these privacy risks. In addition, by collecting users feedbacks on the personal information revealed and the usage of a geosanitization mechanism, we hope that this work will also be useful to constitute a new and valuable dataset on users perceptions on these questions.

6.15. Pseudonymisation techniques and best practices

Participant: Cédric Lauradoux.

This ENISA report explores further the basic notions of pseudonymisation, as well as technical solutions that can support implementation in practice. Starting from a number of pseudonymisation scenarios, the report defines first the main actors that can be involved in the process of pseudonymisation along with their possible roles. It then analyses the different adversarial models and attacking techniques against pseudonymisation, such as brute force attack, dictionary search and guesswork. Moreover, it presents the main pseudonymisation techniques and policies available today.

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.

7.1.5. DARC

- Title: DARC - the Data Anonymization and Re-identification Competition
- Type: Innovation Pédagogique - IDEX LYON
- Duration: September 2019 - 2020.
- Coordinator: INSA.
- Abstract: In order to increase awareness and empower future digital engineers in a fun way on privacy issues, the DARC project offers learning through play through a challenge carried out jointly by three different training courses of INSA students in Bourges and in Lyon. This challenge consists first of all in anonymizing a dataset from an online sales site, then secondly in trying to re-identify the anonymized data of the other groups.

7.2. National Initiatives

7.2.1. ADT PRESERVE

- Title: PRESERVE: Plate-foRme wEb de SEnsibilisation aux pRoblèmes de Vie privéE
- Duration: 2019 - 2020
- Coordinator: INSA.
- Abstract: The goal of this project is to develop a web platform to increase the user awareness on privacy issues. This platform will gather multiple works investigated in the team and will be used to conduct demonstration and stimulate new collaborations and dissemination actions to end users and media.

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-CNIL collaboration

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. 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."

7.3.1.2. SPARTA

Title: Strategic Programs for Advanced Research and Technology in Europe (SPARTA)

Programm: H2020-SU-ICT-03-2018

Duration: February 2019 - January 2022

Coordinator: CEA

Inria contact: Thomas Jensen (Inria), Vincent Roca (for PRIVATICS)

SPARTA Cybersecurity European Competence Network. The consortium consists of 44 partners from 14 different countries, with the goal to demonstrate the setup and assessment of a European SPARTA Cybersecurity Competence Network.

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

- Jeremy Decouchant (University of Luxembourg) visited Privatics from 14/10/2019 to 25/10/2019 through the Erasmus Staff Mobility For Teaching program. During the visit, Jeremie Decouchant participated in network programming lectures and practical sessions at the INSA Lyon engineering school at the M1 level. In addition, the existing scientific collaborations with the team have been also extended around the usage of Intel Software Guard Extensions (SGX) to implement a privacy-preserving recommendation systems and genome studies.
- Gergely Acs, assistant professor at Budapest University (Hungary), visited our team in June. He worked together with Claude Castelluccia on the security and privacy of Federated machine learning.

- Rosin Claude Ngueveu (UQAM) visited the team in Lyon in July 2019 for two weeks to increase the DATA collaboration. During the visit, Rosin Claude Ngueveu presented joint work at APVP 2019 and advanced existing collaboration to include fairness in our work on protection of motion sensor data.

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, 10/10/2018, Lyon, France.

Claude Castelluccia: *AI & Information Disorder* as part of the Global Forum on AI for Humanity forum, October 2019, Paris, France.

8.1.1.2. Member of the Organizing Committees

Antoine Boutet: Workshop on data transparency, 10/10/2019, Lyon, France.

Antoine Boutet: Winter School on Distributed Systems and Networks 2019, 4-8/02/2019, Sept Laux, France.

Antoine Boutet: SRDS 2019, 01-04/10/19 Lyon, France.

Daniel Le Métayer: Panel *Influence or manipulation? What protections in the digital world?*, CPDP 2019, 30/01/2019, Brussels, Belgium.

Claude Castelluccia, *Building trust in AI, building trust with AI*, Global Science Week, 01/06/2019, Grenoble, France.

8.1.2. Scientific Events Selection

8.1.2.1. Member of the Conference Program Committees

Antoine Boutet: Social Network Analysis and Mining 2019, Nature 2019, MDPI 2019, IEEE Transactions on Services Computing 2019, APVP 2019, Compas 2019, Location Privacy Workshop 2019.

Mathieu Cunche: ACM WiSec 2019, APVP 2019, AlgoTel 2019, IEEE WCNC 2019.

Claude Castelluccia: APF 2019.

Cédric Lauradoux: ACM CCSW 2019, Prix Gilles Kahn.

Daniel Le Métayer: XAI 2019, IWPE 2019, CPDP 2019, APF 2019.

Vincent Roca: SPACOMM 2019.

8.1.3. Invited Talks

Claude Castelluccia, *A Risk Analysis Framework for Facial Recognition Applications*, MIAI chair on AI&Ethics (T. Menissier), 04/12/2019, Grenoble, France.

Claude Castelluccia, *Cognitive security*, closing the regulatory gap for consumer neurotechnology" workshop, Brocher Fondation, 25-27/11/2019, Geneva, Switzerland.

Claude Castelluccia, *Influence or manipulation ? What protections in the digital world? (panel)*, CPDP 2019, 30/01/2019, Brussels, Belgium.

Cédric Lauradoux, *Subject Access Request and Proof of Ownership*, SoSySec Seminar, 25/10/2019, Rennes, France.

Cédric Lauradoux, *Influence or manipulation ? What protections in the digital world? (panel)*, CPDP 2019, 30/01/2019, Brussels, Belgium.

Cédric Lauradoux, *Y-aura-t-il un Cambridge Analytica de nos data santé ? (panel)*, FUTUR.E.S, 13/06/2019, Paris, France.

Cédric Lauradoux, *Pseudonymisation*, École de Cybersécurité de l'Université de Nice, 09/07/2019, Nice, France.

Cédric Lauradoux, *Subject Access Right*, Laboratoire d'Innovation Numérique de la CNIL, 12/07/2019, Paris, France.

Cédric Lauradoux, *Subject Access Request and Proof of Ownership*, CyberAlps Workshop on GDPR, 09/10/2019, Grenoble, France.

Daniel Le Métayer, *Inaugural session: How to promote a responsable design and usage of decision making systems ?* , Center for Internet and Society, 27/09/2019, Paris, France.

Daniel Le Métayer and Clément Hénin *Social Responsability of Algorithms*, SRA 2019, 12/12/2019, Paris, France.

Daniel Le Métayer, *HumanAI*, workshop on transparency and accountability for algorithmic decision systems, 11/09/2019, Montreal, Canada.

Mathieu Cunche, *Mécanismes anti-traçage dans les réseaux sans-fil*, journées nationales du GDR sécurité 2019, 12/06/19, Paris, France.

Vincent Roca, *Vers un habitat intelligent... mais fortement indiscret : la maison connectée sous l'angle de la vie privée*, Séminaire "Vie privée, mobile et sécurité", Festival des Libertés Numériques, 06/02/2019, Rennes, France.

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

Master : Antoine Boutet, *Security and Privacy*, 14h, INSA-Lyon, France.

Master : Antoine Boutet, *Security and Privacy 16h*, Polytech Annecy, France.

Undergraduate course: Antoine Boutet, *System and Network*, 160h, L3, INSA-Lyon, France.

Master: Antoine Boutet, *Network, 24h*, Polytech Annecy, 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 Métayer, *Privacy*, 12h, M2 MASH, Université Paris Dauphine, France.

Master : Daniel Le Métayer, *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*, 3h, L-Pro, University of Grenoble Alpes, France.

Undergraduate course : Vincent Roca, *C-Language Programming*, 24h, L-Pro, University of Grenoble Alpes, France.

Master : Vincent Roca, *On Security and Privacy in smartphones*, 3h, M2, France.

8.2.2. E-learning

MOOC "Protection de la vie privée dans le monde numérique:"

Nataliia Bielova, Cédric Lauradoux, Vincent Roca, Session 3, (open during 2 months), FUN-MOOC, Inria, public targeted: around 30000 for the three sessions since 2018, https://www.fun-mooc.fr/courses/course-v1:inria+41015+session03/about.

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* , April 2017, Vincent Roca with Arnaud Legout (DIANA Inria team).
- 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 Métayer.
- PhD in progress: Théo Jourdan, *Privacy-preserving machine learning in medical domain, October 2018, Antoine Boutet.*
- *PhD in progress : Michale Toth,* Privacy protection of Web users and compliance with GDPR and ePrivacy Regulations , *December 2019, Natallia Bielova and Vincent Roca.*
- *Intern (M2): Michael Toth ,* Analyse et présentation accessible aux utilisateurs de Chartes de Vie Privée dans le contexte du RGPD, Vincent Roca.
- Intern (M2): Piyush Patil, *Privacy Leak Analysis in the Context of Smart Homes, Vincent Roca.*
- *Intern (M1): Jan Aalmoes - Understanding how location data influences personalized content in the mobile context, Antoine Boutet.*
- *Intern (M1): Vincent Prax - Privacy Analysis of Email Providers, Cédric Lauradoux.*
- *Intern (L3): Félix Fonteneau - Activity recognition using federated learning, Antoine Boutet.*
- *Intern (L3): Hilaire Bouaddi - Analysis of mobility traces, Antoine Boutet.*
- *Intern (L3): Amine Bahi - Privacy-preserving and scalable machine learning using homomorphic encryption, Antoine Boutet.*

8.2.4. Juries

HDR: Carole Frindel, *Approche computationnelle de l'imagerie médicale : application en neuro-sciences*, INSA Lyon, France, 13/12/2019, Claude Castelluccia.

PhD: Pieter Robyns, *Explicit and Implicit Information Leakage in Wireless Communication*, 11/12/19, Hasselt University, Belgium, Mathieu Cunche (reviewer).

PhD: Timothy CLAEYS, *Security for the Internet of Things: A bottom-up approach to the secure and standardized Internet of Things*, 19/12/19, Université de Grenoble Alpes, France, Mathieu Cunche (examiner).

PhD: Antoine Vastel, *Tracking versus security: investigating the two facets of browser fingerprinting*, Université de Lille, France, 23/10/2019, Daniel Le Métayer.

PhD: Julien Loudet, *Distributed and privacy-preserving personal queries on personal clouds*, Universitée Paris-Saclay, thèse préparée à Université de Versailles Saint-Quentin-en-Yvelines, 24/10/2019, France, Vincent Roca (reviewer).

8.3. Popularization

8.3.1. Hearings

- Claude Castelluccia: *audition at the Council of Europe on AI & Human Rights*, European Parliament, 05/2019, 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).
- Claude Castelluccia is a member of the Grenoble AI institute (MIAI)
- Daniel Le Métayer is a member of the European Commission Multistakeholder expert group to support the application of General Data Protection Regulation (GDPR).
- Daniel Le Métayer is Chair of the CNIL-Inria privacy award.
- Daniel Le Métayer is a member of the steering committee of APVP (Atelier Protection de la Vie Privée).
- Daniel Le Métayer is a member of the steering committee of the chair « Transformation de l'action publique » of Sciences Po Lyon.
- Antoine Boutet is the communication manager of CITI laboratory.
- Antoine Boutet is the manager the Cybersecurity and privacy option of the 5th year of computer science at INSA-Lyon.
- Cédric Lauradoux is a representative of Inria ethical committee COERLE at Inria Grenoble Rhone-Alpes.
- Cédric Lauradoux is a member of the scientific committee of the research action "Cyber-Physical System" of Labex Persyval (Grenoble).
- Cédric Lauradoux is a member of the ethics committee of ComUE Université Grenoble Alpes.
- Cédric Lauradoux is a member of Inria Grenoble Rhône-Alpes committee for technological development.
- Vincent Roca is the PRIVATICS Team Leader (since Nov. 2020).
- Vincent Roca is the co-chair of the NWCRG ("Coding for Efficident Network Communications" Research Group), Internet Research Task Force (IRTF) / IETF.
- Vincent Roca is the ANR 2017 DAPCODS/IOTics project leader.

- Vincent Roca is the chair of the CUMI ("Commision des Utilisateurs des Moyens Informatiques") of Inria Grenoble Rhone-Alpes.
- Vincent Roca is a member of Inria ethical committee COERLE.
- Vincent Roca is co-editor (with Ludovic Mé and Steve Kremer) of a series of articles dedicated to cybersecurity on the Blog Binaire, Le Monde.

8.3.3. Articles and contents

- Antoine Boutet et Mathieu Cunche: *Souriez, vous êtes géolocalisés !*, "Tout compte fait", 06/05/2019
- Antoine Boutet: *L'algorithme : cette formule arbitraire, miroir de l'intention humaine*, En Vue (INSA Lyon newsletter), 18/12/2019 .
- Claude Castelluccia: *Manipulation informationnelle et psychologique*, Le blog binaire du Monde, 05/2018.
- Mathieu Cunche: *Des chercheurs français découvrent des fuites d'informations sensibles sur les appareils Apple*, Science et Avenir.fr, 16/12/2019.
- Mathieu Cunche, *Vers une reconnaissance faciale généralisée*, Techniques de l'ingénieur, 05/12/2019.
- Cédric Lauradoux, *Bientot un identifiant numérique pour tous*, Les Echos, 07/05/2019.
- Cédric Lauradoux, *Facebook, mon fil d'actu, ma bataille*, Mediapart, 05/12/2019.
- Daniel Le Métayer and Claude Castelluccia: *Algorithmic decision making : risks and opportunities for society*, Inria Interview, 25/06/2019.
- Daniel Le Métayer: *Interview about algorithmic decision systems.*, Atlantico, 21/09/2019.
- Vincent Roca and Cédric Lauradoux: *Prêts à tout pour protéger leurs données*, 01NET, 02/10/2019.
- Vincent Roca: *Suite au Livre Blanc Inria sur la cybersécurité : L'ordinateur quantique est une menace pour une partie des techniques de chiffrement actuellement utilisée*, L'Hebdo, BFM Business, 09/02/2019.
- Vincent Roca: *Les objets connecte´s nous espionnent-ils ?*, "Les Idées Claires", France Culture, 12/03/2019.
- Vincent Roca: *Rencontre Sciences et Politique : Cybersécurité*, Acade´mie des sciences - Acade´mie nationale de me´decine - Office parlementaire d'e´valuation des choix scientifiques et technologiques (OPECST), Sénat, 19/06/2019.
- Vincent Roca: *Objets connectés et vie privée*, Les Echos, 20/03/2019.
- Vincent Roca: *Nouvelles technologies : tous espionnés*, Journal de 20 Heures, France 2, 12/04/2019.
- Vincent Roca: *Vie privée et maison intelligente*, Café des Sciences, 14/05/2019, Grenoble, France.

8.3.4. Education

- Nataliia Bielova: *Animation du forum du MOOC Protection de la vie privée dans le monde numérique*, May-June 2019.
- Antoine Boutet: Contest of data anonymization between three different groups of INSA students (from Lyon and Bourges), 25-26/11/19.
- Cédric Lauradoux: *Action vie privée*, Maison pour la science, 14/02/2019, Grenoble, France.
- Cédric Lauradoux: *Protéger sa vie privée*, DIU EIL (enseignant d'informatique au lycée), 12/04/2019, Grenoble, France.
- Cédric Lauradoux: *Animation du forum du MOOC Protection de la vie privée dans le monde numérique*, May-June 2019.
- Vincent Roca: : *Animation du forum du MOOC Protection de la vie privée dans le monde numérique*, May-June 2019.

8.3.5. Interventions

- Mathieu Cunche: *Cybersécurité: Protéger ses objets physiques*, Pint of science, 22/05/2019, Lyon, France.
- Cédric Lauradoux: *Procès du robot 42*, Festival TRANSFO, 24/01/2019, Grenoble, France.
- Cédric Lauradoux: *Challenges de cryptologie*, Printemps du numérique, 08/04/19, Die, France.
- Cédric Lauradoux: *Challenges de cryptologie*, MathC2+ internship, 25/06/19, Grenoble, France.
- Cédric Lauradoux: *Atelier cryptographie*, Fête de la Science, 10/10/2019, Grenoble, France.
- Cédric Lauradoux: *Procès du robot 42*, Fête de la Science, 12/10/2019, Grenoble, France.
- Cédric Lauradoux: *Les données en question*, Festival Imaginascience, 16/10/2019, Annecy, France.
- Cédric Lauradoux: *Le Numérique nous menace-t-il ?*, Telecom Saint Etienne, 03/12/2019, Saint Etienne, France.
- Cédric Lauradoux: *Vie privée et Liberté dans le monde numérique*, Collège Les Dauphins, 03/12/2019, Saint Jean de Soudain, France.
- Cédric Lauradoux: *Challenges de cryptologie*, Cité scolaire Jean PREVOST, 12/12/2019, Villard de Lans, France.
- Cédric Lauradoux: *Avoir un usage Internet éclairé*, UIAD, 18/12/2019, Grenoble, France.
- Vincent Roca: *Maison intelligente : le point de vue du respect de la vie privée*, Café des Sciences, 14/05/2019, Grenoble, France.
- Vincent Roca: *Habitat intelligent, service de vélo partagé et vie privée*, Fête de la Science, 12/10/2019, Grenoble, France.

8.3.6. Internal action

- Cédric Lauradoux: *Formation Éthique*, Inria Paris, 26/01/2019, Paris, France.
- Cédric Lauradoux: *Formation Éthique*, Inria Saclay, 25/04/2019, Saclay, France.
- Cédric Lauradoux: *Formation Éthique*, Inria Grenoble - Rhône-Alpes, 13/05/2019, Lyon, France.
- Cédric Lauradoux: *Rules of Cyber Engagement*, Inria JSI, 07/06/2019, Lyon, France.
- Cédric Lauradoux: *Privacy: Understanding the GDPR*, Inria Grenoble - Rhône-Alpes, 15/10/2019, Grenoble, France.
- Vincent Roca: *Sécurité numérique : technique, éthique, juridique : qui s'y frotte s'y pique (moderator)*, Inria JSI, 07/06/2019, Lyon, France.

9. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] G. Acs, L. Melis, C. Castelluccia, E. De Cristofaro.*Differentially Private Mixture of Generative Neural Networks*, in "IEEE Transactions on Knowledge and Data Engineering", June 2019, vol. 31, n^{o} 6, p. 1109-1121, A shorter version of this paper appeared at the 17th IEEE International Conference on Data Mining (ICDM 2017). This is the full version, publishedin IEEE Transactions on Knowledge and Data Engineering (TKDE) [*DOI :* 10.1109/TKDE.2018.2855136], https://hal.inria.fr/hal-01921923

[2] G. Celosia, M. Cunche.*Discontinued Privacy: Personal Data Leaks in Apple Bluetooth-Low-Energy Continuity Protocols*, in "Proceedings on Privacy Enhancing Technologies", July 2020, vol. 2020, p. 26 - 46 [*DOI :* 10.2478/popets-2020-0003], https://hal.inria.fr/hal-02394619

[3] T. JOURDAN, A. BOUTET, C. FRINDEL.*Vers la protection de la vie privée dans les objets connectés pour la reconnaissance d'activité en santé*, in "Revue des Sciences et Technologies de l'Information - Série TSI : Technique et Science Informatiques", 2019, p. 1-27, forthcoming [*DOI :* 10.3166/RIA.28.1-27], https://hal.inria.fr/hal-02421854

International Conferences with Proceedings

[4] C. BONIFACE, I. FOUAD, N. BIELOVA, C. LAURADOUX, C. SANTOS.*Security Analysis of Subject Access Request Procedures How to authenticate data subjects safely when they request for their data*, in "APF 2019 - Annual Privacy Forum", Rome, Italy, June 2019, p. 1-20, https://hal.inria.fr/hal-02072302

[5] A. BOUTET, S. GAMBS.*Demo: Inspect what your location history reveals about you Raising user awareness on privacy threats associated with disclosing his location data*, in "CIKM 2019 - 28th ACM International Conference on Information and Knowledge Management", Beijing, China, ACM, November 2019, p. 2861-2864 [*DOI :* 10.1145/3357384.3357837], https://hal.inria.fr/hal-02421828

[6] G. CELOSIA, M. CUNCHE.*Fingerprinting Bluetooth-Low-Energy Devices Based on the Generic Attribute Profile*, in "IoT S&P 2019 - 2nd International ACM Workshop on Security and Privacy for the Internet-of-Things", London, United Kingdom, ACM Press, November 2019, p. 24-31 [*DOI :* 10.1145/3338507.3358617], https://hal.inria.fr/hal-02359914

[7] G. CELOSIA, M. CUNCHE.*Saving Private Addresses: An Analysis of Privacy Issues in the Bluetooth-Low-Energy Advertising Mechanism*, in "MobiQuitous 2019 - 16th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services", Houston, United States, December 2019, p. 1-10 [*DOI :* 10.1145/3360774.3360777], https://hal.inria.fr/hal-02394629

[8] M. CUNCHE, D. LE MÉTAYER, V. MOREL.*A Generic Information and Consent Framework for the IoT*, in "TRUSTCOM 2019 - 18th IEEE International Conference on Trust, Security and Privacy in Computing and Communications", Rotorua, New Zealand, August 2019, p. 1-8, https://hal.inria.fr/hal-02166181

[9] J. DECOUCHANT, A. BOUTET, J. YU, P. ESTEVES-VERISSIMO.*P3LS: Plausible Deniability for Practical Privacy-Preserving Live Streaming*, in "SRDS 2019 - 38th International Symposium on Reliable Distributed Systems", Lyon, France, October 2019, p. 1-10, https://hal.inria.fr/hal-02421820

[10] A. KASSEM, G. ACS, C. CASTELLUCCIA, C. PALAMIDESSI.*Differential Inference Testing: A Practical Approach to Evaluate Sanitizations of Datasets*, in "SPW 2019 - 40th IEEE Symposium on Security and Privacy Workshops", San Francisco, United States, IEEE, May 2019, p. 72-79 [*DOI :* 10.1109/SPW.2019.00024], https://hal.archives-ouvertes.fr/hal-02422992

[11] R. PARDO, D. LE MÉTAYER.*Analysis of Privacy Policies to Enhance Informed Consent*, in "DBSEC 2019 - 33rd Annual IFIP WG 11.3 Conference on Data and Applications Security and Privacy", Charleston, SC, United States, S. N. FOLEY (editor), Data and Applications Security and Privacy XXXIII, Springer International Publishing, July 2019, vol. LNCS-11559, p. 177-198, Part 3: Privacy [*DOI :* 10.1007/978-3-030-22479-0_10], https://hal.inria.fr/hal-02384593

Scientific Books (or Scientific Book chapters)

[12] S. KREMER, L. MÉ, D. RÉMY, V. ROCA.*Cybersecurity : Current challenges and Inria's research directions*, Inria white book, Inria, January 2019, n^{o} 3, 172, https://hal.inria.fr/hal-01993308

[13] S. KREMER, L. MÉ, D. RÉMY, V. ROCA.*Cybersécurité : Défis actuels et axes de recherche à l'Inria*, Inria white book, Inria, May 2019, n^{o} 3, 18, https://hal.inria.fr/hal-02414281

Research Reports

[14] C. CASTELLUCCIA, D. LE MÉTAYER.*Analyse des impacts de la reconnaissance faciale - Quelques éléments de méthode*, Inria Grenoble Rhône-Alpes, November 2019, https://hal.inria.fr/hal-02373093

[15] C. HENIN, D. LE MÉTAYER.*Towards a Generic Framework for Black-box Explanation Methods (Extended Version)*, Inria Grenoble Rhône-Alpes ; Ecole des Ponts ParisTech, May 2019, n^{o} RR-9276, p. 1-28, https://hal.inria.fr/hal-02131174

[16] V. MOREL, R. PARDO.*Three Dimensions of Privacy Policies*, Inria - Research Centre Grenoble – Rhône-Alpes ; CITI - CITI Centre of Innovation in Telecommunications and Integration of services, August 2019, n^{o} RR-9287, https://arxiv.org/abs/1908.06814 , https://hal.inria.fr/hal-02267641

[17] R. PARDO, D. LE MÉTAYER.*Analysis of Privacy Policies to Enhance Informed Consent (Extended Version)*, Inria Rhône-Alpes, March 2019, n^{o} RR-9262, p. 1-22, https://arxiv.org/abs/1903.06068 - Extended Version, https://hal.inria.fr/hal-02067924

Other Publications

[18] G. AVOINE, C. LAURADOUX, T.-R. ROLANDO.*Should Chess Players Learn Computer Security ?*, February 2019, p. 1-11, HACKING IT SECURITY MAGAZINE, https://hal.inria.fr/hal-02082837

[19] G. CELOSIA, M. CUNCHE.*DEMO: Himiko: A human interface for monitoring and inferring knowledge on Bluetooth-Low-Energy objects*, ACM Press, May 2019, p. 292-293, WiSec 2019 - 12th Conference on Security and Privacy in Wireless and Mobile Networks, Poster [*DOI :* 10.1145/3317549.3326297], https://hal.inria.fr/hal-02154148

[20] V. ROCA, A. BEGEN.*Forward Error Correction (FEC) Framework Extension to Sliding Window Codes (RFC 8680)*, RFC Editor (https://www.rfc-editor.org/), January 2020, RFC 8680, Standards Track, TSVWG (Transport Area) working group of IETF (Internet Engineering Task Force), https://www.rfc-editor.org/rfc/rfc8680.html, https://hal.inria.fr/hal-01345125

[21] V. ROCA, B. TEIBI.*Sliding Window Random Linear Code (RLC) Forward Erasure Correction (FEC) Schemes for FECFRAME (RFC 8681)*, RFC Editor (https://www.rfc-editor.org/), January 2020, RFC 8681, Standards Track, TSVWG (Transport Area) working group of IETF (Internet Engineering Task Force), https://www.rfc-editor.org/rfc/rfc8681.html, https://hal.inria.fr/hal-01630089

[22] M. SAITO, M. MATSUMOTO, V. ROCA, E. BACCELLI.*TinyMT32 Pseudorandom Number Generator (PRNG) (RFC 8682)*, RFC Editor (https://www.rfc-editor.org/), January 2020, RFC 8682, Standards Track, TSVWG (Transport Area) working group of IETF (Internet Engineering Task Force), https://www.rfc-editor.org/rfc/rfc8682.html, https://hal.inria.fr/hal-02449210

[23] M. THIERY, V. ROCA, A. LEGOUT.*Privacy implications of switching ON a light bulb in the IoT world*, July 2019, working paper or preprint, https://hal.inria.fr/hal-02196544

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, until Jan 2019, HDR]
Loris Marchal [CNRS, Researcher, HDR]
Bora Uçar [CNRS, Researcher, HDR]

Faculty Members

Anne Benoit [École Normale Supérieure de Lyon, Associate Professor, HDR]
Grégoire Pichon [Univ Claude Bernard, Associate Professor, from Sep 2019]
Yves Robert [École Normale Supérieure de Lyon, Professor, HDR]

Samuel Thibault [Univ de Bordeaux, Associate Professor, from Feb 2019 until Jul 2019, HDR]

PhD Students

Yishu Du [China Scholarship Council, PhD Student, from Dec 2019]
Yiqin Gao [Univ de Lyon, PhD Student]
Changjiang Gou [China Scholarship Council, PhD Student]
Li Han [China Scholarship Council, PhD Student]
Aurélie Kong Win Chang [École Normale Supérieure de Lyon, PhD Student]
Valentin Le Fèvre [École Normale Supérieure de Lyon, PhD Student]
Ioannis Panagiotas [Inria, PhD Student]
Filip Pawlowski [CIFRE Huawei, PhD Student]

Administrative Assistants

Solene Audoux [Inria, Administrative Assistant, until Apr 2019]
Evelyne Blesle [Inria, Administrative Assistant, from Sep 2019]
Virginie Bouyer [Inria, Administrative Assistant, from Apr 2019 until Aug 2019]

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 [45] each includes more than 500,000 cores and deliver a sustained performance of more than 10 Peta FLOPS. The volunteer computing platform BOINC [41] 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 [44], had an average of 2.33 faults per day during the period from August 2008 to February 2010 [68]. 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” [43] 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 [39] 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 [42]).

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 [50], [56], [57], [67], 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 $2 \leq \alpha \leq 3$ [48], [49], [54]; 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 [69] 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 [65]. 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 [69]; (ii) *replication* consists in executing the same task on several processors simultaneously, in order to meet the reliability constraints [47]; and (iii) *checkpointing* consists in "saving" the work done at some certain instants, hence reducing the amount of work lost when a failure occurs [64].

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 [40]. 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" [60]. 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 [58], [59], [63]. 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 [52] 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 [51] 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 [53], [66], [61]. 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 [55]. 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 the 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

- Jean-Yves L'Excellent co-created the MUMPS technologies start-up and left the team to work full time for MUMPS technologies.
- Grégoire Pichon joined the team as an Associate Professor of University Claude Bernard, Lyon 1.
- Anne Benoit was elected chair of the IEEE Technical Committee on Parallel Processing.
- Anne Benoit received on February 2019 the award for Editorial Excellence as Associate Editor of the IEEE Transactions on Parallel and Distributed Systems during 2018.
- Yves Robert received the 2020 IEEE-CS Charles Babbage Award *for contributions to parallel algorithms and scheduling techniques.* This award covers all aspects of parallel computing including computational aspects, novel applications, parallel algorithms, theory of parallel computation, parallel computing technologies, among others. Further information about the award, including a list of past recipients, may be found at https://www.computer.org/web/awards/charles-babbage. The award consists of a $1,000 honorarium, certificate, and the invitation to present a paper and/or presentation at the annual IEEE-CS International Parallel and Distributed Processing Symposium (IPDPS).

5.1.1. Awards

BEST PAPERS AWARDS :

[16]

F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Effective heuristics for matchings in hypergraphs*, in "SEA2 2019 - International Symposium on Experimental Algorithms - Special Event", Kalamata, Greece, Springer, 2019, p. 248-264 [*DOI :* 10.1007/978-3-030-34029-2_17], https://hal.inria.fr/hal-02417475

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 has been developed by INPT(ENSEEIHT)-IRIT, Inria, CERFACS, University of Bordeaux, CNRS and ENS Lyon. Since January 2019, it is developed and licensed by Mumps Technologies SAS.

NEWS OF THE YEAR: In June 2019, a new version of MUMPS, MUMPS 5.2.1, was released by Mumps Technologies.

- 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 Brémond, 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 - Mumps Technologies SAS
- Contact: Jean-Yves L'Excellent
- URL: http://mumps-solver.org/

7. New Results

7.1. Creation of the start-up "Mumps Technologies SAS"

In January 2019, Jean-Yves L'Excellent left the ROMA team to co-found with Patrick Amestoy and Chiara Puglisi the company "Mumps Technologies" around the free software library MUMPS (Cecill-C licence). MUMPS solves large systems of sparse linear equations on high-performance computers in a robust and effective way. Mumps Technologies carries on collaborations and R&D activities to keep the MUMPS software library state-of-the-art and freely available, while offering to its clients a set of services.

7.2. Scheduling independent stochastic tasks under deadline and budget constraints

This work discusses 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.

The findings were published in a journal [8].

7.3. Online scheduling of task graphs on heterogeneous 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 by Amaris et al. [46], 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 $\Theta(\sqrt{m/k})$. We also adapt all our results to the case of multiple types of processors. 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.

The findings were published in a journal [9].

7.4. A generic approach to scheduling and checkpointing workflows

This work deals with scheduling and checkpointing strategies to execute scientific workflows on failure-prone large-scale platforms. To the best of our knowledge, this work is 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 revisit classical mapping heuristics such as HEFT and MINMIN and complement them with several checkpointing strategies. The objective is 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.

The findings were published in a journal [10].

7.5. Limiting the memory footprint when dynamically scheduling DAGs on shared-memory platforms

Scientific workflows are frequently modeled as Directed Acyclic Graphs (DAGs) of tasks, which represent computational modules and their dependences 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 computing platforms. However, for some workflows, such a dynamic schedule may run out of memory by processing too many tasks simultaneously. This paper 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 DAGs 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. Our solution consists in adding new fictitious edges, while trying to minimize the critical path of the graph. After proving that this problem is 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 the quality of the final parallel schedule.

The findings were published in a journal [12].

7.6. Scheduling independent stochastic tasks on heterogeneous cloud platforms

This work introduces scheduling strategies to maximize the expected number of independent tasks that can be executed on a cloud platform within a given budget and under a deadline constraint. The cloud platform is composed of several types of virtual machines (VMs), where each type has a unit execution cost that depends upon its characteristics. The amount of budget spent during the execution of a task on a given VM is the product of its execution length by the unit execution cost of that VM. The execution lengths of tasks follow a variety of standard probability distributions (exponential, uniform, half-normal, etc.), which is known beforehand and whose mean and standard deviation both depend upon the VM type. Finally, there is a global available budget and a deadline constraint, and the goal is to successfully execute as many tasks as possible before the deadline is reached or the budget is exhausted (whichever comes first). On each VM, the scheduler can decide at any instant to interrupt the execution of a (long) running task and to launch a new one, but the budget already spent for the interrupted task is lost. The main questions are which VMs to enroll, and whether and when to interrupt tasks that have been executing for some time. We assess the complexity of the problem by showing its NP-completeness and providing a 2-approximation for the asymptotic case where budget and deadline both tend to infinity. Then we introduce several heuristics and compare their performance by running an extensive set of simulations.

This work has been presented at the Cluster 2019 conference [17].

7.7. Improved energy-aware strategies for periodic real-time tasks under reliability constraints

This work revisited the real-time scheduling problem recently introduced by Haque, Aydin and Zhu [62]. In this challenging problem, task redundancy ensures a given level of reliability while incurring a significant energy cost. By carefully setting processing frequencies, allocating tasks to processors and ordering task executions, we improve on the previous state-of-the-art approach with an average gain in energy of 20%. Furthermore, we establish the first complexity results for specific instances of the problem.

This work has been accepted at the RTSS 2019 conference [18].

7.8. Multilevel algorithms for acyclic partitioning of 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 the 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 interpart 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 multiway 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 significant improvements compared to the current state of the art.

This work is published in a journal [11].

7.9. A multi-dimensional Morton-ordered block storage for mode-oblivious tensor computations

Computation on tensors, treated as multidimensional arrays, revolve around generalized basic linear algebra subroutines (BLAS). We propose a novel data structure in which tensors are blocked and blocks are stored in an order determined by Morton order. This is not only proposed for efficiency reasons, but also to induce efficient performance regardless of which mode a generalized BLAS call is invoked for; we coin the term mode-oblivious to describe data structures and algorithms that induce such behavior. Experiments on one of the most bandwidth-bound generalized BLAS kernel, the tensor–vector multiplication, not only demonstrate superior performance over two state-of-the-art variants by up to 18%, but additionally show that the proposed data structure induces a 71% less sample standard deviation for tensor–vector multiplication across d modes, where d varies from 2 to 10. Finally, we show our data structure naturally expands to other tensor kernels and demonstrate up to 38% higher performance for the higher-order power method.

This work is published in a journal [13]

7.10. 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. We first generalize some graph matching heuristics for this problem. We then 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.

This work is published in the proceedings of SEA[2], where it has received the best paper award [16].

7.11. Karp-Sipser based kernels for bipartite graph matching

We consider Karp–Sipser, a well known matching heuristic in the context of data reduction for the maximum cardinality matching problem. We describe an efficient implementation as well as modifications to reduce its time complexity in worst case instances, both in theory and in practical cases. We compare experimentally against its widely used simpler variant and show cases for which the full algorithm yields better performance .

This work appears in the proceedings of ALENEX2020 [20]

7.12. Efficient and effective sparse tensor reordering

This paper formalizes the problem of reordering a sparse tensor to improve the spatial and temporal locality of operations with it, and proposes two reordering algorithms for this problem, which we call BFS-MCS and Lexi-Order. The BFS-MCS method is a Breadth First Search (BFS)-like heuristic approach based on the maximum cardinality search family; Lexi-Order is an extension of doubly lexical ordering of matrices to tensors. We show the effects of these schemes within the context of a widely used tensor computation, the Candecomp/Parafac decomposition (CPD), when storing the tensor in three previously proposed sparse tensor formats: coordinate (COO), compressed sparse fiber (CSF), and hierarchical coordinate (HiCOO). A new partition-based superblock scheduling is also proposed for HiCOO format to improve load balance. On modern multicore CPUs, we show Lexi-Order obtains up to $4.14\times$ speedup on sequential HiCOO-Mttkrp and $11.88\times$ speedup on its parallel counterpart. The performance of COO-and CSF-based Mttkrps also improves. Our two reordering methods are more effective than state-of-the-art approaches.

This work appears in the proceedings of ICS2019 [21].

7.13. High performance tensor–vector multiplication on shared-memory systems

Tensor–vector multiplication is one of the core components in tensor computations. We have recently investigated high performance, single core implementation of this bandwidth-bound operation. Here, we investigate its efficient, shared-memory implementations. Upon carefully analyzing the design space, we implement a number of alternatives using OpenMP and compare them experimentally. Experimental results on up to 8 socket systems show near peak performance for the proposed algorithms.

This work appears in the proceedings of PPAM2019 and is supported with a technical report [22], [36].

7.14. Matrix symmetrization and sparse direct solvers

We investigate algorithms for finding column permutations of sparse matrices in order to have large diagonal entries and to have many entries symmetrically positioned around the diagonal. The aim is to improve the memory and running time requirements of a certain class of sparse direct solvers. We propose efficient algorithms for this purpose by combining two existing approaches and demonstrate the effect of our findings in practice using a direct solver. We show improvements in a number of components of the running time of a sparse direct solver with respect to the state of the art on a diverse set of matrices.

This work will appear in the proceedings of CSC2020 [23].

7.15. 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, under the realistic duplex single-port communication model. 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 more than three times better makespan in cases with many communications.

This work appears in the proceedings of IPDPS 2019 [25].

7.16. Improving Locality-Aware Scheduling with Acyclic Directed Graph Partitioning

We investigate efficient execution of computations, modeled as Directed Acyclic Graphs (DAGs), on a single processor with a two-level memory hierarchy, where there is a limited fast memory and a larger slower memory. Our goal is to minimize execution time by minimizing redundant data movement between fast and slow memory. We utilize a DAG partitioner that finds localized, acyclic parts of the whole computation that can fit into fast memory, and minimizes the edge cut among the parts. We propose a new scheduler that executes each part one-by-one, obeying the dependency among parts, aiming at reducing redundant data movement needed by cut-edges. Extensive experimental evaluation shows that the proposed DAG-based scheduler significantly reduces redundant data movement.

This work will appear in the proceedings of PPAM 2019 [24].

7.17. Replication Is More Efficient Than You Think

We revisit replication coupled with checkpointing for fail-stop errors. Replication enables the application to survive many fail-stop errors, thereby allowing for longer checkpointing periods. Previously published works use replication with the no-restart strategy, which never restart failed processors until the application crashes. We introduce the restart strategy where failed processors are restarted after each checkpoint. which may introduce additional overhead during checkpoints but prevents the application configuration from degrading throughout successive checkpointing periods. We show how to compute the optimal checkpointing period for this strategy, which is much larger than the one with no-restart, thereby decreasing I/O pressure. We show through simulations that using the restart strategy significantly decreases the overhead induced by replication, in terms of both total execution time and energy consumption.

This work appears in the proceedings of SC 2019 [15], [28].

7.18. Generic matrix multiplication for multi-GPU accelerated distributed-memory platforms over PaRSEC

We introduce a generic and flexible matrix-matrix multiplication algorithm $C = A \times B$ for state-of-the-art computing platforms. Typically, these platforms are distributed-memory machines whose nodes are equipped with several accelerators. To the best of our knowledge, SLATE is the only library that provides a publicly available implementation on such platforms, and it is currently limited to problem instances where the C matrix can entirely fit in the memory of the GPU accelerators. Our algorithm relies on the classical tile-based outer-product algorithm, but enhances it with several control dependencies to increase data re-use and to optimize communication flow from/to the accelerators within each node. The algorithm is written with the PaRSEC runtime system, which allows for a fast and generic implementation, while achieving close-to-peak performance.

This work appears in the proceedings of Scala 2019 [19].

7.19. Reservation strategies for stochastic jobs

We are interested in scheduling stochastic jobs on a reservation-based platform. Specifically, we consider jobs whose execution time follows a known probability distribution. The platform is reservation-based, meaning that the user has to request fixed-length time slots. The cost then depends on both (i) the request duration (pay for what you ask); and (ii) the actual execution time of the job (pay for what you use).

A reservation strategy determines a sequence of increasing-length reservations, which are paid for until one of them allows the job to successfully complete. The goal is to minimize the total expected cost of the strategy. We provide some properties of the optimal solution, which we characterize up to the length of the first reservation. We then design several heuristics based on various approaches, including a brute-force search of the first reservation length while relying on the characterization of the optimal strategy, as well as the discretization of the target continuous probability distribution together with an optimal dynamic programming algorithm for the discrete distribution.

We evaluate these heuristics using two different platform models and cost functions: The first one targets a cloud-oriented platform (e.g., Amazon AWS) using jobs that follow a large number of usual probability distributions (e.g., Uniform, Exponential, LogNormal, Weibull, Beta), and the second one is based on interpolating traces from a real neuroscience application executed on an HPC platform. An extensive set of simulation results show the effectiveness of the proposed reservation-based approaches for scheduling stochastic jobs.

This work appears in the proceedings of IPDPS 2019 [14].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

In the context of a consortium (http://mumps-consortium.org) of users of the MUMPS library (http://mumps-solver.org), we had partnership contracts with EDF, ALTAIR, FFT-MSC Software, Michelin, LSTC, Siemens, ESI Group, Total, Safran, LBNL, Airbus, and SHELL. Following the creation of the start-up Mumps Technologies in January 2019, these contracts (scientific exchanges, support, organization of point-to-point and plenary meetings, releases in advance, ...) have been transferred to Mumps Technologies.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

ANR Project SOLHARIS (2019-2013), 4 years. The ANR Project SOLHAR was launched in November 2019, for a duration of 48 months. It gathers five academic partners (the HiePACS, ROMA, RealOpt, STORM and TADAAM) Inria project-teams, and CNRS-IRIT) and two industrial partners (CEA/CESTA and Airbus CRT). This project aims at producing scalable methods for direct methods for the solution of sparse linear systems on large scale and heterogeneous computing platforms, based on task-based runtime systems.

The proposed research is organized along three distinct research thrusts. The first objective deals with the development of scalable linear algebra solvers on task-based runtimes. The second one focuses on the deployement of runtime systems on large-scale heterogeneous platforms. The last one is concerned with scheduling these particular applications on a heterogeneous and large-scale environment.

9.2. International Initiatives

9.2.1. Inria International Labs

9.2.1.1. 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.

9.2.2. Inria International Partners

9.2.2.1. 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.

9.2.3. 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 four 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 Mingsong Chen have lead 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.

9.3. International Research Visitors

9.3.1. Visits to International Teams

9.3.1.1. 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.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Selection

10.1.1.1. Chair of Conference Program Committees

- Anne Benoit is the chair of HPC Algorithms Track of ISC 2020.
- Bora Uçar was the chair of the Algorithms Track of HiPC 2019; he has organized a mini-symposium at ICIAM 2019 on combinatorial scientific computing (12 talks) with Aydin Buluc and Alex Pothen.
- Frédéric Vivien was Global Chair of Topic 10 (Theory and Algorithms for Parallel Computation and Networking) of Euro-Par 2019, Göttingen, Germany, August 26-30 2019.

10.1.1.2. Member of the Conference Program Committees

- Anne Benoit was a member of the program committees of **IPDPS'20**, **IPDPS'19**, **SC'19**, **ESA'19**, and **Compas'19**.
- Loris Marchal was/is a member of the program committees of **HiPC 2019**, **HPCS 2019**, **ICPP 2019** and **ICPP 2020**.
- Grégoire Pichon was a member of the program committee of **Compas'19**.
- Yves Robert was a member of the program committee of **FTXS'19**, **Scala'19**, **PMBS'19** workshops co-located with SC'19 in Denver CO, and of the new conference with proceedings inside **SiamPP'20**.
- Bora Uçar was a member of the program committee of **PPAM 2019**, **ISC**, **ICPP**, **IPDPS 2020 Workshops**.
- Frédéric Vivien was a member of the program committees of **IPDPS'20**, **PDP 2020**, and **PDP 2019**.

10.1.1.3. Reviewer

- Loris Marchal has reviewed paper for the following conferences: **SPAA 2019**, **SC 2019**, **PODC 2019**,

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

- Yves Robert is Associate Editor of JPDC (Elsevier Journal of Parallel and Distributed Computing), TOPC (ACM Trans. On Parallel Computing), IJHPCA (Int. J. High Performance Computing and Applications) and JOCS (J. Computational Science).
- Bora Uçar is a member of the editorial board of Parallel Computing, SIAM Journal on Matrix Analysis and Applications, SIAM Journal on Scientific Computing, and IEEE Transactions on Parallel and Distributed Computing.
- Frédéric Vivien is a member of the editorial board of the Journal of Parallel and Distributed Computing.

10.1.2.2. Reviewer - Reviewing Activities

- Loris Marchal has reviewed papers for the journal CCPE (Concurrency and Computation: Practice and Experience).
- Yves Robert reviewed papers for journals IEEE TC, IEEE TPDS, JPDC, ACM TOPC, IEEE Trans. Cloud Computing.
- Bora Uçar has reviewed papers for journals IEEE TPDS, ACM TOMS, SIAM SISC, Frontiers of Information Technology & Electronic Engineering.

10.1.3. Invited Talks

- Yves Robert delivered a keynote presentation at HPCS'2019 in Dublin (Int. Conf. on High Performance Computing & Simulation)

10.1.4. Leadership within the Scientific Community

- Anne Benoit is a member of the Steering Committee of IPDPS and HCW (Heterogeneity in Computing Workshop, co-located with IPDPS).
- Yves Robert is a member of the Steering Committee of IPDPS and HCW .
- Bora Uçar has participated in the steering committees of IPDPS, CSC, and HiPC; he has also served as vice-chair of IEEE Technical Committee on Parallel Processing.

10.1.5. Scientific Expertise

- Frédéric Vivien was re-elected to the scientific council of the École normale supérieure de Lyon. He is also a member of the scientific council of the IRMIA labex http://labex-irmia.u-strasbg.fr/.
- Yves Robert is an expert for the Horizon 2020 program of the European Commission and has reviewed two projects in 2019.
- Yves Robert chaired the HCERES evaluation committee for the Skoltech PhD program in Computational and Data-Intensive Science and Engineering, in Moscow, Russia.

10.1.6. Research Administration

- Frédéric Vivien is the vice-head of the LIP laboratory since September 2017.

10.2. Teaching - Supervision - Juries

- Yves Robert was a Committee member for the HDR of Bora Uçar. .

10.2.1. Teaching

Licence: Anne Benoit, Responsible of the L3 students at ENS Lyon, France

Master: Yves Robert, Responsible of Master Informatique Fondamentale, ENS Lyon, France

Licence: Anne Benoit, Algorithmique avancée, 48, L3, ENS Lyon, France

Licence: Yves Robert, Algorithmique, 48, L3, ENS Lyon, France

Licence: Yves Robert, Proababilités, 48, L3, ENS Lyon, France

Licence: Grégoire Pichon, Réseaux, 34, L3, Univ. Lyon 1, France

Licence: Loris Marchal, Compilation (practicals), 14, L3, Univ. Lyon 1, France

Licence: Loris Marchal, Operating systems (practicals), 12, L2, Univ. Lyon 1, 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 : Loris Marchal, Data-Aware Algorithms, 30, M2 Informatique Fondamentale, ENS Lyon 1, France.

10.2.2. Supervision

PhD in progress: Yishu Du, "Resilience for numerical methods", started in December 2019, funding: China Scholarship Council and Inria, advisors: Yves Robert and Loris Marchal.

PhD in progress: Yiqin Gao, "Replication Algorithms for Real-time Tasks with Precedence Constraints", started in October 2018, funding: 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, funding: China Scholarship Council 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 in progress: Ioannis Panagiotas, "High performance algorithms for big data graph and hypergraph problems", started in October 2017, funding: Inria, advisor: 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).

10.2.3. Juries

- Loris Marchal is a responsible of the competitive selection of ENS Lyon students for Computer Science, and is thus a member of the jury of this competitive exam.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] B. UÇAR.*Partitioning, matching, and ordering: Combinatorial scientific computing with matrices and tensors*, ENS de Lyon, September 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02377874

Articles in International Peer-Reviewed Journal

[2] P. R. AMESTOY, A. BUTTARI, J.-Y. L'EXCELLENT, T. MARY.*Bridging the gap between flat and hierarchical low-rank matrix formats: the multilevel BLR format*, in "SIAM Journal on Scientific Computing", May 2019, vol. 41, n^{o} 3, p. A1414-A1442 [*DOI :* 10.1137/18M1182760], https://hal.archives-ouvertes.fr/hal-01774642

[3] 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", February 2019, vol. 45, n^{o} 1, p. 1-23 [*DOI :* 10.1145/3242094], https://hal.archives-ouvertes.fr/hal-01955766

[4] P. R. 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", 2019, vol. 41, n^{o} 1, p. A269-A291 [*DOI :* 10.1137/17M1151882], https://hal.inria.fr/hal-01955659

[5] G. Aupy, A. Benoit, B. Goglin, L. Pottier, Y. Robert.*Co-scheduling HPC workloads on cache-partitioned CMP platforms*, in "International Journal of High Performance Computing Applications", April 2019, vol. 33, n^{o} 6, p. 1221-1239 [*DOI :* 10.1177/1094342019846956], https://hal.inria.fr/hal-02093172

[6] O. Beaumont, T. Lambert, L. Marchal, B. Thomas.*Performance Analysis and Optimality Results for Data-Locality Aware Tasks Scheduling with Replicated Inputs*, in "Future Generation Computer Systems", October 2019, p. 1-28 [*DOI :* 10.1016/j.future.2019.08.024], https://hal.inria.fr/hal-02275473

[7] 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*, in "International Journal of Networking and Computing", 2019, vol. 9, n^{o} 1, p. 2-27 [*DOI :* 10.15803/ijnc.9.1_2], https://hal.inria.fr/hal-02082369

[8] L.-C. Canon, A. K. W. Chang, Y. Robert, F. Vivien.*Scheduling independent stochastic tasks under deadline and budget constraints*, in "International Journal of High Performance Computing Applications", June 2019, p. 1-19 [*DOI :* 10.1177/1094342019852135], https://hal.inria.fr/hal-02291031

[9] L.-C. Canon, L. Marchal, B. Simon, F. Vivien.*Online Scheduling of Task Graphs on Heterogeneous Platforms*, in "IEEE Transactions on Parallel and Distributed Systems", 2019, p. 1-12, forthcoming [*DOI :* 10.1109/TPDS.2019.2942909], https://hal.inria.fr/hal-02291268

[10] L. Han, V. Le Fèvre, L.-C. Canon, Y. Robert, F. Vivien.*A Generic Approach to Scheduling and Checkpointing Workflows*, in "International Journal of High Performance Computing Applications", May 2019, p. 1-19 [*DOI :* 10.1177/1094342019866891], https://hal.inria.fr/hal-02140295

[11] J. Herrmann, Y. M. Özkaya, B. Uçar, K. Kaya, U. V. Catalyurek.*Multilevel Algorithms for Acyclic Partitioning of Directed Acyclic Graphs*, in "SIAM Journal on Scientific Computing", July 2019, vol. 41, n^{o} 4, p. A2117-A2145 [*DOI :* 10.1137/18M1176865], https://hal.inria.fr/hal-02306566

[12] L. Marchal, B. Simon, F. Vivien.*Limiting the memory footprint when dynamically scheduling DAGs on shared-memory platforms*, in "Journal of Parallel and Distributed Computing", February 2019, vol. 128, p. 30-42 [*DOI :* 10.1016/j.jpdc.2019.01.009], https://hal.inria.fr/hal-02025521

[13] F. Pawłowski, B. Uçar, A.-J. Yzelman.*A multi-dimensional Morton-ordered block storage for mode-oblivious tensor computations*, in "Journal of computational science", March 2019, p. 1-35, forthcoming [*DOI :* 10.1016/j.jocs.2019.02.007], https://hal.inria.fr/hal-02082524

International Conferences with Proceedings

[14] 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, IEEE, May 2019, p. 166-175 [*DOI :* 10.1109/IPDPS.2019.00027], https://hal.inria.fr/hal-01968419

[15] A. BENOIT, T. HÉRAULT, V. LE FÈVRE, Y. ROBERT.*Replication Is More Efficient Than You Think*, in "SC 2019 - International Conference for High Performance Computing, Networking, Storage, and Analysis (SC'19)", Denver, United States, November 2019, https://hal.inria.fr/hal-02273142

[16] *Best Paper*
F. DUFOSSÉ, K. KAYA, I. PANAGIOTAS, B. UÇAR.*Effective heuristics for matchings in hypergraphs*, in "SEA2 2019 - International Symposium on Experimental Algorithms - Special Event", Kalamata, Greece, Springer, 2019, p. 248-264 [*DOI :* 10.1007/978-3-030-34029-2_17], https://hal.inria.fr/hal-02417475.

[17] Y. GAO, L.-C. CANON, Y. ROBERT, F. VIVIEN.*Scheduling independent stochastic tasks on heterogeneous cloud platforms*, in "IEEE Cluster 2019 - International Conference on Cluster Computing", Albuquerque, United States, IEEE, September 2019, p. 1-11, https://hal.inria.fr/hal-02271675

[18] L. HAN, L.-C. CANON, J. LIU, Y. ROBERT, F. VIVIEN.*Improved energy-aware strategies for periodic real-time tasks under reliability constraints*, in "RTSS 2019 - 40th IEEE Real-Time Systems Symposium", York, United Kingdom, February 2020, https://hal.inria.fr/hal-02271704

[19] T. HERAULT, Y. ROBERT, G. BOSILCA, J. DONGARRA.*Generic Matrix Multiplication for Multi-GPU Accelerated Distributed-Memory Platforms over PaRSEC*, in "ScalA 2019 - IEEE/ACM 10th Workshop on Latest Advances in Scalable Algorithms for Large-Scale Systems", Denver, United States, IEEE, November 2019, p. 33-41 [*DOI :* 10.1109/SCALA49573.2019.00010], https://hal.inria.fr/hal-02436180

[20] K. KAYA, J. LANGGUTH, I. PANAGIOTAS, B. UÇAR.*Karp-Sipser based kernels for bipartite graph matching*, in "SIAM Symposium on Algorithm Engineering and Experiments (ALENEX20)", Salt Lake City, Utah, United States, January 2020, https://hal.inria.fr/hal-02350734

[21] J. LI, B. UÇAR, U. V. CATALYUREK, J. SUN, K. BARKER, R. VUDUC.*Efficient and effective sparse tensor reordering*, in "ICS 2019 - ACM International Conference on Supercomputing", Phoenix, United States, June 2019, p. 227-237 [*DOI :* 10.1145/3330345.3330366], https://hal.inria.fr/hal-02306569

[22] F. PAWŁOWSKI, B. UÇAR, A.-J. YZELMAN.*High performance tensor-vector multiplication on shared-memory systems*, in "PPAM 2019 - 13th International Conference on Parallel Processing and Applied Mathematics", Bialystok, Poland, September 2019, p. 1-11, https://hal.inria.fr/hal-02332496

[23] R. PORTASE, B. UÇAR.*Matrix symmetrization and sparse direct solvers*, in "SIAM Workshop on Combinatorial Scientific Computing 2020", Seattle, United States, 2020, https://hal.inria.fr/hal-02417778

[24] Y. M. ÖZKAYA, A. BENOIT, U. V. CATALYUREK.*Is Acyclic Directed Graph Partitioning Effective for Locality-Aware Scheduling?*, in "PPAM 2019 - 13th International Conference on Parallel Processing and Applied Mathematics", Bialystok, Poland, September 2019, https://hal.inria.fr/hal-02273122

[25] Y. M. ÖZKAYA, A. BENOIT, B. UÇAR, J. HERRMANN, U. V. CATALYUREK.*A scalable clustering-based task scheduler for homogeneous processors using DAG partitioning*, in "IPDPS 2019 - 33rd IEEE International Parallel & Distributed Processing Symposium", Rio de Janeiro, Brazil, IEEE, May 2019, p. 155-165 [*DOI :* 10.1109/IPDPS.2019.00026], https://hal.inria.fr/hal-02082794

Scientific Books (or Scientific Book chapters)

[26] L. MARCHAL, É. SAULE, O. SINNEN.*Special Issue Proposal for the Parallel Computing Journal: HeteroPar 2016 and HCW 2016 Workshops*, Elsevier, April 2019, https://hal.inria.fr/hal-02423211

Research Reports

[27] A. BENOIT, C. GOU, L. MARCHAL.*Partitioning tree-shaped task graphs for distributed platforms with limited memory*, Inria Grenoble Rhône-Alpes, March 2019, n^{o} RR-9115, p. 1-34, https://hal.inria.fr/hal-01644352

[28] A. BENOIT, T. HÉRAULT, V. LE FÈVRE, Y. ROBERT.*Replication Is More Efficient Than You Think*, Inria - Research Centre Grenoble – Rhône-Alpes, 2019, n^{o} RR-9278, https://hal.inria.fr/hal-02265925

[29] A. BENOIT, V. LE FÈVRE, P. RAGHAVAN, Y. ROBERT, H. SUN.*Design and Comparison of Resilient Scheduling Heuristics for Parallel Jobs*, Inria - Research Centre Grenoble – Rhône-Alpes, October 2019, n^{o} RR-9296, p. 1-29, https://hal.inria.fr/hal-02317464

[30] L.-C. CANON, A. KONG WIN CHANG, Y. ROBERT, F. VIVIEN.*Scheduling independent stochastic tasks under deadline and budget constraints. Extended Version*, Inria Grenoble Rhône-Alpes, February 2019, n^{o} RR-9257, p. 1-38, https://hal.inria.fr/hal-02025785

[31] A. GAINARU, B. GOGLIN, V. HONORÉ, G. PALLEZ, P. RAGHAVAN, Y. ROBERT, H. SUN.*Reservation and Checkpointing 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 2019, n^{o} RR-9294, https://hal.inria.fr/hal-02328013

[32] Y. GAO, L.-C. CANON, Y. ROBERT, F. VIVIEN.*Scheduling independent stochastic tasks on heterogeneous cloud platforms*, Inria - Research Centre Grenoble – Rhône-Alpes, May 2019, n^{o} RR-9275, https://hal.inria.fr/hal-02141253

[33] Y. GAO, L.-C. CANON, F. VIVIEN, Y. ROBERT.*Scheduling stochastic tasks on heterogeneous cloud platforms under budget and deadline constraints*, Inria - Research Centre Grenoble – Rhône-Alpes, February 2019, n^{o} RR-9260, p. 1-34, https://hal.inria.fr/hal-02047434

[34] L. HAN, L.-C. CANON, J. LIU, Y. ROBERT, F. VIVIEN.*Improved energy-aware strategies for periodic real-time tasks under reliability constraints*, Inria - Research Centre Grenoble – Rhône-Alpes, February 2019, n^{o} RR-9259, p. 1-38, https://hal.inria.fr/hal-02056520

[35] T. HÉRAULT, Y. ROBERT, G. BOSILCA, J. DONGARRA.*Generic matrix multiplication for multi-GPU accelerated distributed-memory platforms over PaRSEC*, Inria Grenoble - Rhone-Alpes, September 2019, n^{o} RR-9289, https://hal.inria.fr/hal-02282529

[36] F. PAWŁOWSKI, B. UÇAR, A.-J. YZELMAN.*High performance tensor-vector multiplies on shared memory systems*, Inria - Research Centre Grenoble – Rhône-Alpes, May 2019, n^o RR-9274, p. 1-20, https://hal.inria.fr/hal-02123526

Other Publications

[37] A. AZAD, B. UÇAR, A. POTHEN.*Trends in Combinatorial Analysis: Complex Data, Machine Learning, and High-Performance Computing*, SIAM, September 2019, p. 1-3, https://hal.inria.fr/hal-02304457

[38] C. MOMMESSIN, O. BEAUMONT, L.-C. CANON, L. EYRAUD-DUBOIS, G. LUCARELLI, L. MARCHAL, B. SIMON, D. TRYSTRAM.*Scheduling on Two Types of Resources: a Survey*, January 2020, https://arxiv.org/abs/1909.11365 - working paper or preprint, https://hal.inria.fr/hal-02432381

References in notes

[39] *Blue Waters Newsletter*, dec 2012

[40] *Blue Waters Resources*, 2013, https://bluewaters.ncsa.illinois.edu/data

[41] *The BOINC project*, 2013, http://boinc.berkeley.edu/

[42] *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

[43] *System Resilience at Extreme Scale: white paper*, 2008, DARPA, https://pdfs.semanticscholar.org/9fcb/154d6afce23cd9951fd7c116b86255d91b5c.pdf

[44] *Top500 List - November*, 2011, http://www.top500.org/list/2011/11/

[45] *Top500 List - November*, 2012, http://www.top500.org/list/2012/11/

[46] 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

[47] 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

[48] 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

[49] 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

[50] 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

[51] 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

[52] S. Blackford, J. Dongarra.*Installation Guide for LAPACK*, LAPACK Working Note, June 1999, n^{o} 41, originally released March 1992

[53] 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

[54] 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

[55] 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

[56] 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

[57] 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

[58] 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

[59] 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

[60] 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

[61] 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

[62] M. A. Haque, H. Aydin, D. Zhu.*On reliability management of energy-aware real-time systems through task replication*, in "IEEE Transactions on Parallel and Distributed Systems", 2017, vol. 28, n^{o} 3, p. 813–825

[63] J. W. H. Liu.*The multifrontal method for sparse matrix solution: Theory and Practice*, in "SIAM Review", 1992, vol. 34, p. 82–109

[64] 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

[65] 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

[66] 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

[67] Y. ROBERT, F. VIVIEN.*Algorithmic Issues in Grid Computing*, in "Algorithms and Theory of Computation Handbook", Chapman and Hall/CRC Press, 2009

[68] 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

[69] 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

Faculty Members

Tanguy Risset [Team leader, INSA Lyon, Professor, HDR]
Florent Dupont de Dinechin [INSA Lyon, Professor, HDR]
Florin Hutu [INSA Lyon, Associate Professor]
Kevin Marquet [INSA Lyon, Associate Professor]
Guillaume Salagnac [INSA Lyon, Associate Professor]
Guillaume Villemaud [INSA Lyon, Associate Professor, HDR]

External Collaborator

Régis Rousseau [INSA Lyon, Engineer]

Technical Staff

Gero Muller [Inria, Engineer, from Feb 2019]

PhD Students

Gautier Berthou [Inria]
Andrea Bocco [CEA]
Maxime Christ [Univ Grenoble Alpes, from Oct 2019]
Luc Forget [INSA Lyon]
Tarik Lassouaoui [INSA Lyon]
Yohann Uguen [INSA Lyon]
Yanni Zhou [INSA Lyon]
Tristan Delizy [INSA Lyon until Oct 2019]

Administrative Assistants

Anouchka Ronceray [Inria, from July 2019]
Sophie Karsai [Inria, until July 2019]

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).

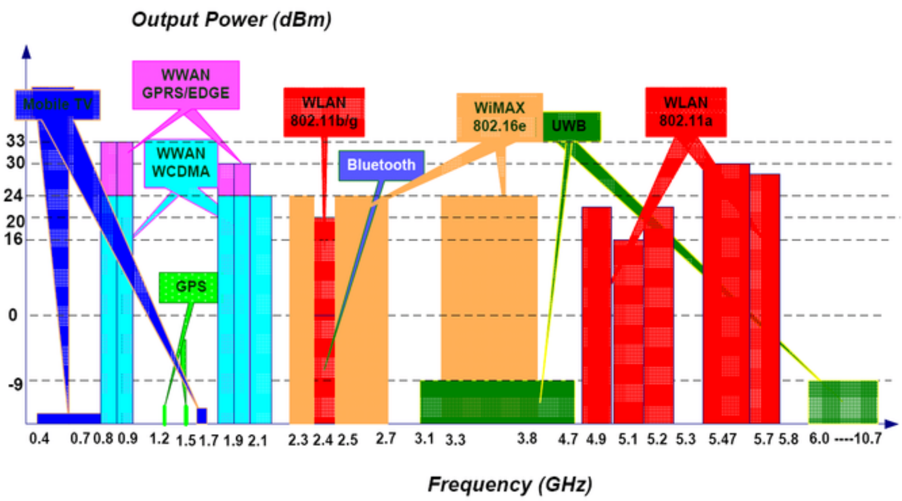


Figure 1. The most recent standards for wireless communications are developed in the UHF and VHF bands. These bands are mostly saturated (source: WPAN/WLAN/WWAN Multi-Radio Coexistence, IEEE 802 Plenary, Atlanta, USA, Nov.2007)

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 technologies, 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.* [37]

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 [36], 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 [26]. The current development of the wireless industry is surely slowed down by the lack of radio resources and the lack of systems flexibility.

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 [33], [34]. 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 [26]. 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 [33], [29].

2.2.1. SDR Technology

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

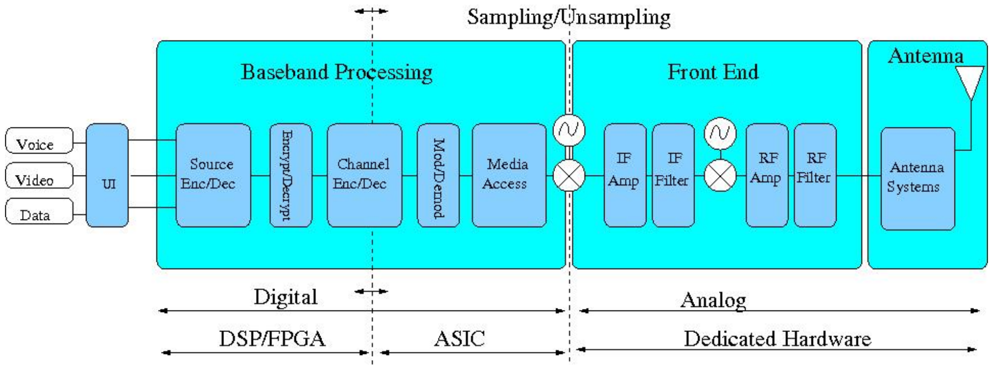


Figure 2. Radio Block Diagram, highlighting separation between digital and analog parts, as well as programmable, configurable and fixed hardware parts.

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" [35]. 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 [30].

3. Research Program

3.1. Flexible Radio Front-End

These are the research axis as they were proposed at the creation of the Socrate Team.

This axis mainly deals with the radio front-end of software radio terminals. 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.

3.2. Multi-User Communications

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.

3.3. Software Radio Programming Model

Finally the third research axis is concerned with software aspect of the software radio terminal. 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.

3.4. Evolution of the Socrate team

In 2018 the Socrate team which was originaly conceived to develop software defined radio 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.

The advent of non-volatile memory technologies (NVRAM) is causing a major evolution in all software layers. On the one hand, the non-volatility of data in the event of a breakdown necessarily leads to fatal inconsistencies if the memory is not managed correctly. On the other hand, these memories have very different performances from the usual DRAM, which tends to the appearance of hybrid and complex memory hierarchies. Many technological and scientific challenges are to be faced in all software layers to deal with these two sets of issues. Above all, the answers to be provided depend on the calculation system considered and for what purpose it is constructed.

Within the framework of very low consumption sensors and devices, the Socrate team proposed, with Sytare [28], a software solution allowing to develop embedded applications on platforms supporting an intermittent power supply (TPS – *Transiently Powered System*) and integrating NVRAM as illustrated in Figure 3. The IPL ZEP (https://project.inria.fr/iplzep/) was also launched by Socrate last year to respond to various scientific challenges related to this issue.

The recent advance in harvesting technologies provides new research direction to Socrate which have skills in radio propagation and low power radio (wake-up radio for instance [31]. Fig 3, illustrates the *future ultra-low sensor* as envisioned by Socrate.

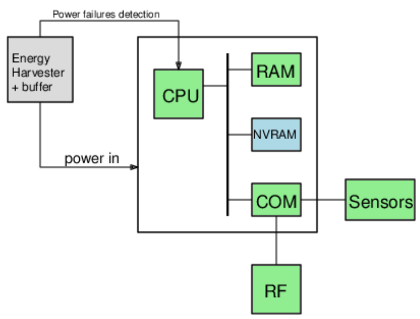


Figure 3. Architecture targeted by Socrate: low energy wireless sensor with peripherals and non volatile memory

4. New Software and Platforms

4.1. 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.

- Participants: Florent de Dinechin and Luc Forget
- Partners: CNRS - ENS Lyon - UCBL Lyon 1 - UPVD
- Contact: Florent de Dinechin
- URL: http://flopoco.gforge.inria.fr/

4.2. 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

4.3. NanoTracer

KEYWORDS: Embedded systems - Power monitoring - Low power consumption

FUNCTIONAL DESCRIPTION: NanoTracer is a high performance ammeter dedicated to power measurements for small devices. The system measures currents between 100nA and 100mA (gain is audo-adjusted dynamically) with a sampling frequency of 2Msps. Data is streamed to a PC over USB which enables long-running experiments, or just real-time visualization of data.

- Contact: Guillaume Salagnac
- URL: https://gitlab.inria.fr/nanotracer/

4.4. marto

Modern Arithmetic Tools

KEYWORDS: High-level synthesis - Arithmetic - FPGA

FUNCTIONAL DESCRIPTION: Marto provides C++ headers to implement custom sized arithmetic operators such as:

Custom sized posits and their environment (including the quire) Custom sized IEEE-754 numbers Custom sized Kulisch accumulators (and sums of products)

- Participants: Yohann Uguen, Florent de Dinechin and Luc Forget
- Contact: Yohann Uguen
- Publication: hal-02130912v4
- URL: https://gitlab.inria.fr/lforget/marto

4.5. hint

High-level synthesis Integer Library

KEYWORD: High-level synthesis

FUNCTIONAL DESCRIPTION: Hint is an header-only arbitrary size integer API with strong semantics for C++. Multiple backends are provided using various HLS libraries, allowing a user to write one operator and synthetize it using the main vendor tools.

- Participants: Yohann Uguen, Florent de Dinechin and Luc Forget
- Contact: Luc Forget
- Publication: hal-02131798v2
- URL: https://github.com/yuguen/hint

5. New Results

5.1. Flexible Radio Front-End

Activities in this axis could globally be divided in three main topics: wake-up radio and wireless power transfer, RFID systems and combination of spatial modulation and full-duplex.

5.1.1. Wake-Up radio and wireless power transfer

The ubiquity of wireless sensor networks (WSN), as well as the rapid development of the Internet of Things (IoT), impel new approaches to reduce the energy consumption of the connected devices. The wake-up radio receivers (WuRx) were born in this context to reduce as much as possible the energy consumption of the radio communication part. We aim at proposing a low-cost, high-efficiency rectifier to improve a quasi-passive WuRx performance in terms of communication range. By optimizing the wideband matching circuit and the proposed rectifier's load impedance, the sensitivity was increased by 5 dB, corresponding to an increase of the communication range (13 meters in free space) [10].

We also studied an original solution to maximize the DC power collected in the case of a wireless power transfer (WPT) scenario. Using state-space model representation, the WPT System is considered as a feedback approach in order to maximize the amount of harvested energy. To do this, a global simulation is performed to show the importance of taking into account the propagation channel and the rectifier circuit aspects in the case of optimizing the waveform to increase the harvested energy. By using an optimized multi-sine signal with zero phase as the excitation, taking into account the characteristics of the channel and the physical contributions of the rectifier, we managed to obtain better output DC values compared to a single tone source or a multi-sine signal without optimization, with the same average power input [14].

We plan now to apply this optimized WPT technique to feed Wireless sensors in the particular case of ventilation ducts (HVAC) [24].

5.1.2. RFID

The ARA (Auvergne Rhone Alpes) RAFTING project mainly deals with the design and analysis of wire antennas for RFID tags in the context of wearable electronics. More specifically, an helical dipole antenna dedicated to the smart textile yarn applications has been designed. Moreover, the performance was analysized with respect to mechanical constraints, together with the extraction of accurate electrical models. This work was done in collaboration with Primo 1D company. In perspective, the integration of the NFC protocol together with RFID UHF and the integration of sensing capabilities is envisaged [6], [19], [7], [12], [21].

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. Theoretical and simulated performance have been studied.

5.1.3. Combination of spatial modulation and full-duplex

Spatial modulation (SM) as a new MIMO technique is based on transmitting part of the information by activating different emitting antennas. SM increases spectral efficiency and uses only one radio frequency chain. Moreover, for full-duplex (FD) communication systems, self-interference (SI) is always a central problem. Therefore, combining FD and SM can dramatically reduce the difficulty of SIC (Self-interference Cancellation) because of the single SI chain. A Full Duplex Spatial Modulation (FDSM) system is proposed and an active analog SIC is designed in this work. Moreover, the impact of SIC accuracy on the system performance is studied. The results demonstrate that the accuracy requirement will increase as the INR (Self-interference-to-noise Ratio) increases. The FDSM system is less sensitive than the FD system, which can get a better BER (Bit Error Rate) performance as errors increase. Furthermore, an SI detector is proposed to resolve the influence of the number of detected symbols.

5.2. Software Radio Programming Model

5.2.1. Transiently powered systems and Non-Volatile Memory

Socrate is studying the new NVRAM (Non-Volatile Radom Access Memory) technology and its use in ultra-low power context. Non-Volatile memory has been extisting 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 [32].

Socrate has started a work on the applicability of NVRAM for *transiently powered systems*, i.e. systems which may undergo power outage at any time. This study resulted in the Sytare software published in IEEE Transaction on Computer [3] and is also studied in an Inria Project Lab ZEP (https://project.inria.fr/iplzep/teams/).

The Sytare software introduces a checkpointing system that takes into account peripherals (ADC, leds, timer, radio communication, etc.) present on all embedded systems. 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.

Another acheivement in this domain is the PhD of Tristan Delizy that concerns memory heterogeneity that results from new NVM technologies. While emerging memory technologies may offer power reduction and high integration density, 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. We aim to provide the embedded application programmer with a transparent software mechanism to leverage this memory heterogeneity. The work of Tristan Delizy studies the interaction between dynamic memory allocation and memory heterogeneity. He provides cycle accurate simulation of embedded platforms with various memory technologies and shows that different dynamic allocation strategies have a major impact on performance. He demonstrates that interesting performance gains can be achieved even for a low fraction of memory using low latency technology, but only with a clever placement strategy between memory banks. This work will soon be proposed to publication.

5.2.2. Sytare integration in Riot

The ADT SytaRiot has been granted to provide transient power management in the Riot operating system [27]. This integration was realized by Gero muller, here is a summary of the technical tasks and correponding pull request on Riot GitHub:

5.2.2.1. Port RIOT to MSP430+FRAM micro-controllers

- Bring-up the chip against the newer msp430-elf compiler and integrate the toolchain into the RIOT CI infrastructure, cf https://github.com/RIOT-OS/riotdocker/pull/67 , https://github.com/RIOT-OS/riotdocker/pull/82 , https://github.com/RIOT-OS/riotdocker/pull/91

- Implement initial support for the MSP430FR59xx in RIOT, including device drivers for key on-chip peripherals (UART, Timers, GPIO, etc). cf https://github.com/RIOT-OS/RIOT/pull/11012
- Implement a board support package for the MSP-EXP430FR5969 Launchpad Development Kit and the Boost-IR daughter-board (Infrared transceiver + keypad), cf https://github.com/geromueller/RIOT/commit/f13d33
- Participate in IETF hackathon 104 (Prague, March 23–29, 2019) to work on SUIT IoT Firmware Update, cf https://trac.ietf.org/trac/ietf/meeting/wiki/104hackathon

5.2.2.2. Explicit checkpointing in RIOT

- Implement the required low-level code (e.g. DMA driver) for saving/restoring the state of the application to FRAM. cf https://github.com/geromueller/RIOT/commits/checkpoint
- Implement save/restore methods in all relevant device drivers (DMA, GPIO, UART, Timers) and design an API to expose checkpointing as a general system service in RIOT. cf https://github.com/geromueller/RIOT/commit/8b301e
- Participate in the RIOT Summit (Helsinki, September 5–6, 2019) to give a talk about checkpointing and power measurement. cf https://summit.riot-os.org/2019/blog/speakers/gero-muller/ Power measurement

5.2.3. A high-performance ammeter for embedded systems

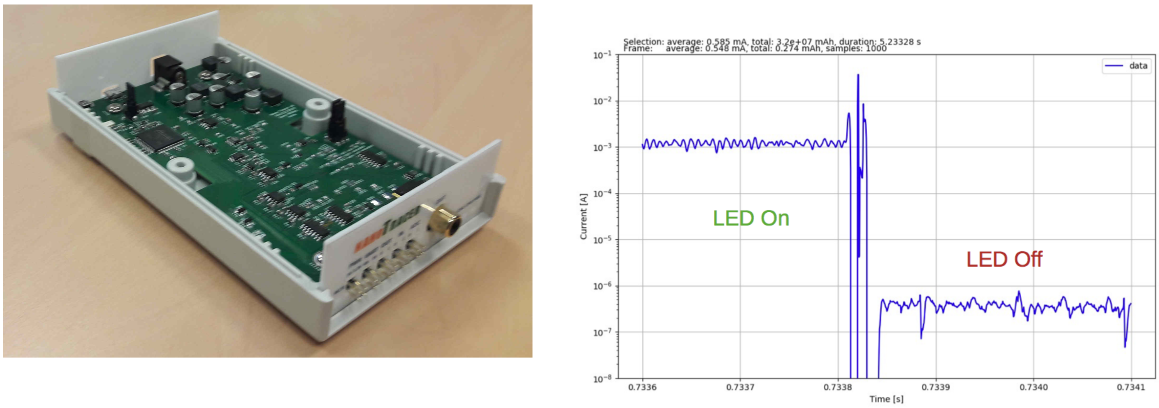


Figure 4. Photo (left) of first packaged nanoTracer prototype and snapshot (right) of a measurement provided by nanoTracer

In embedded low power processing, precise power consumption is a key issue. The Socrate team realized that existing tools could not fullfill the requirements needed for harvesting devices monitoring (measuring from nano-Amperes to milli-Ampere current values at a high sampling rate and continuously).

With the skills of Gero Müller hired on the SytaRiot ADT, the socrate team designed and built a high performance ammeter dedicated to power measurements for small devices. Our prototype measures currents between 100nA and 100mA (gain is audo-adjusted dynamically) with a sampling frequency of 2Msps. Data is streamed to a PC over USB which enables long-running experiments, or just real-time visualization of data (cf screenshot in Fig. 4).

The device, named *nanoTracer*, is referenced in the software section, it is an open project on gitlab (https://gitlab.inria.fr/nanotracer/). A first version is currently tested at Inria (Alexandre Abadie from the IoT SED team) and should soon be available for free for Inria and Academic researcher. We are working on solutions to provide a commercial circuit if requests come from other actors.

5.2.4. Ultra-low latency audio on FPGA

Recently the Socrate team started a collaboration with the researchers of the GRAME group. GRAME is a “Centre National de Création Musicale” (CNCM) organized in three departments: music production, transmission/mediation, and computer music research. Four GRAME researchers have expertise in computer science (compilation), audio DSP, digital lutherie, and human-computer interaction in general. GRAME has been leading the development of the FAUST[0] programming language since its creation in 2004. The GRAME researchers have been associated to CITI as external members in September 2019.

Socrate and GRAME have started a collaboration through the Syfala *(synthèse audio faible latence)* project funded by the Fédération Informatique de Lyon. The goal of Syfala is to design an FPGA-based platform for multichannel ultra-low-latency audio Digital Signal Processing (DSP), programmable at high-level with FAUST and using Socrate’s software FloPoCo (http://flopoco.gforge.inria.fr). This platform is intended to be usable for various applications ranging from sound synthesis and processing to active sound control and artificial sound field/room acoustics.

Two internships have been working on this project. A first result was a presentation by Florent de Dinechin and Tanguy Risset, introducing the use of HLS and FPGA for audio, at the second *Programmable Audio Workshop* (https://faust.grame.fr/paw/) organized by GRAME.

5.2.5. Evaluation of the posit number system

The posit number system is a very elegant way to represent real numbers in a computers. Its proponents promote it as a better replacement for floating-point arithmetic: posits do indeed improve the application-level accuracy of some applications. However, this also comes with accuracy regressions in other cases. Socrate members, along with members of the AriC project-team, first studied some numerical aspects of posits [18]. Socrate then performed a thorough evaluation of the implementation of the main posit operators, improving the state of the art in hardware posit in the process. Posit operators were then compared to IEEE 754-compliant floating-point operators, and were found to be about twice as slow and twice as expensive [20], [15].

5.2.6. Evaluation of the Unum number system

CEA researcher, in collaboration with Socrate members, designed a complete accelerator for the UNUM number system, including hardware [8] and compiler support [11]. A novelty of this work is the use of a variable-length, self-describing, and memory-oriented floating-point number format [23].

5.2.7. General computer arithmetic

The 10th anniversary of the FloPoCo open-source arithmetic core generator project was the occasion to reflect on the evolutions of the field in a special session about arithmetic generator challenges organized at the ARITH conference [16].

A marked evolution over this period has been the deployment of very good High-Level Synthesis tools, thanks to which hardware is described using a software programming language (usually C++). This comes with many new arithmetic optimization opportunities, some of which have been reviewed in collaboration with Steven Derrien, from Inria Cairn [25]

An issue was the lack in this context of a portable, unified, and hardware-oriented library of arbitrary precision integers. In collaboration with David Thomas from Imperial College, London, we worked on such a library, and demonstrated that it enables a safe description of complex small-grain architectures (such as floating-point or posit operators) with a performance matching traditional hardware description languages [9].

[0] FAUST is a domain specific language for real-time audio signal processing primarily developed at GRAME-CNCM and by a worldwide community. FAUST is based on a compiler “translating” DSP specifications written in FAUST into a wide range of lower-level languages (e.g., C, C++, Rust, Java, WASM, LLVM bitcode, etc.). Thanks to its “architecture” system, generated DSP objects can be embedded into template programs (wrappers) used to turn a FAUST program into a specific ready-to-use object (e.g., standalone, plug-in, smartphone app, webpage, etc.).

Meanwhile, we keep studying the most basic operators. There has always existed two main methods of implementing mulitplication by a constant in hardware: Table-Based, and Shift-And-Add. This deserved a qualitative and quantitative comparison [17]. This work (with Martin Kumm, from Fulda Technical University, and Silviu Filip, from Inria Cairn) also includes a refined ILP-based algorithm for the problem of multiplying a fixed-point input number by a real constant.

6. Bilateral Contracts and Grants with Industry

6.1. Bilateral Contracts with Industry

6.1.1. Research Contract with Bosch 2019

In collaboration with Aric, Socrate worked with Bosch on the impementation of the Power function in an embedded context.

7. Partnerships and Cooperations

7.1. National Initiatives

7.1.1. 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.

7.1.2. 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.

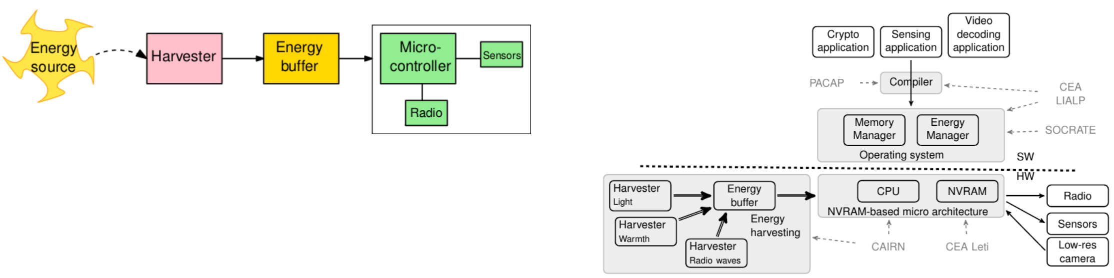


Figure 5. Example of system targeted by the ZEP project on the left, and on the right: the ZEP research program.

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

The project is illustrated by the figure 5, 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.

7.1.3. 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.

7.1.4. ADT SytaRiot

The Riot system (https://www.riot-os.org/) is well known within Inria, it is a joint implementation of Inria and Freie Universität Berlin which is today one of the most widely used open-source OS on small embedded systems. The arrival of non-volatile memories promises a new generation of sensors on which the memory hierarchy will be more heterogeneous than today. The communicating system will be able to undergo a power cut em complete and resume its current activity when power returns.

Sytare (https://gitlab.inria.fr/citi-lab/sytare/), developed for three years by the Socrates team (with the support of an ADT), targets intermittent feeding which will arrive when the technologies of *harvesting* (recovery of ambient energy) are democratized.

The objective of this ADT is to make Riot compatible with NVRAM-based architecture, therefore to integrate Sytare with Riot and thus open Riot to ultra low power platforms containing NVRAM, *eg* Texas microcontrollers Instrument MSP430FR5969.

7.1.5. Digital Hardware AI Architectures

Florent de Dinechin participates to the chair *Digital Hardware AI Architectures* held by Prof. Frédéric Pétrot at the Multidisciplinary Institute in Artificial Intelligence (MIAI) of Grenoble. The other participants are François Duhem (Spintec/CEA) and Fabrice Rastello (LIG/Inria), with industrial partners Google France, Kalray, STMicroelectronics, and Upmem.

This chair funds the PhD of Maxime Christ, which studies how very low-precision arithmetic formats may improve the efficiency of the learning phase of neural networks.

7.2. European Initiatives

7.2.1. 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.

7.3. International Initiatives

7.3.1. Inria International Partners

7.3.1.1. Informal International Partners

Socrate has collaborations with the following international partners.

- **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.
- **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.
- **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
- **Technical University of Fulda, Germany**. This collaboration with the group of Martin Kumm covers many aspects of computer arithmetic, with several joint papers, collaboration on the FloPoCo project, and work in progress on a textbook to appear in 2020. Scientific-in-charge at Inria: Florent de Dinechin
- **Imperial College, London, UK**, departments of Computing and Electrical Engineering. This collaboration with the groups of David Thomas and George Constantinides covers several aspects of reconfigurable computing and reconfigurable arithmetic. Scientific-in-charge at Inria: Florent de Dinechin

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events: Organisation

Tanguy Risset and Leonardo Cardoso have been at the origin ot the First French GNU Radio days in 2018: https://gnuradio-fr-18.sciencesconf.org/. This is a very visible result of the activity on CorteXlab. The second version of the workshop (called now *european GNU Radio days*) has almost double its number of participants in 2019: https://gnuradio-fr-19.sciencesconf.org/. Next year session is planned in June in Poitier.

Kevin Marquet has organized the conference “Vous avez dit sobriété (numérique) ?” within the Ecoinfo CNRS GDS (https://ecoinfo.cnrs.fr/2019/07/11/projets-ecoinfo-en-cours/).

8.1.1.1. Member of the Organizing Committees

- Tanguy Risset is member of the organizing committee of the second European GNU Radio days.

8.1.1.2. Chair of Conference Program Committees

Florin Hutu was session chair during the IEEE International Symposium on Signals Circuits and Systems, 11-12 July 2019, Iasi, Romania.

8.1.1.3. Member of the Conference Program Committees

Tanguy Risset was a member of the following technical program committees:

- IEEE Computer Society Annual Symposium on VLSI (ISVLSI) 2019
- Design Automation and Test in Europe (DATE) 2019
- International Conference on Cognitive Radio Oriented Wireless Networks (CROWNCOM) 2019

Florent de Dinechin was a member of the following technical program committees:

- IEEE Symposium on Computer Arithmetic (ARITH)
- IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM)
- Conférence d'informatique en Parallélisme, Architecture et Système (COMPAS)

Guillaume Villemaud was a member of the following technical program committee:

- EUCAP 2019,
- EuCNC 2019

8.1.2. Journal

8.1.2.1. Member of the Editorial Boards

Guillaume Villemaud is an associate editor of Annals of Telecommunications (Springer).

Florent de Dinechin was Guest Editor for a special issue of IEEE Transactions on Computers dedicated to computer arithmetic [4].

8.1.2.2. Reviewer - Reviewing Activities

Florin Hutu reviewed articles for the following conferences: 2019 International Conference on Advanced Technologies for Communications (ATC), 2019 IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC),

Florin Hutu Florin HUTU reviewed articles for the following journals: IEEE Transactions on Communications, IEEE Communications Letters, IEEE Journal of Radio Frequency Identification, Elsevier Advances in Space Research, MDPI Sensors, MDPI Applied Sciences.

8.1.3. Invited Talks

Florent de Dinechin gave a talk at the PAW workshop (Programmable Audio Workshop) organised by GRAME in Lyon on 14 dec. 2019. the title was: "FPGAs for low latency audio applications?" (https://faust.grame.fr/paw/#fpga)

Florent de Dinechin gave an invited lecture at ICTP (the International Center for Theoretical Physics) for the May 2019 edition of the Advanced Workshop on FPGA-based Systems-On-Chip for Scientific Instrumentation and Reconfigurable Computing.

Tanguy Risset gave a tutorial untitled "Writing a custom GNU Radio processing block"at the second edition of the *european GNU Radio days*, organized in besançon in June 2019 (https://gnuradio-fr-19.sciencesconf.org/resource/page/id/2).

Gero Müller gave an invited talk untitled "Intermittent power and high precision high speed power measurements" at Riot Summit 2019 (https://summit.riot-os.org/2019/blog/speakers/gero-muller/)

8.1.4. Leadership within the Scientific Community

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.

8.1.5. 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).

8.1.6. Research Administration

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.

Kevin Marquet is associate director of the CNRS GDS ECO-info (http://ecoinfo.cnrs.fr/)

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

- Tanguy Risset is professor at the Telecommunications Department of Insa Lyon.
- Florent de Dinechin is a professor at the Computer Science Department of Insa Lyon. He also teaches computer architecture at ENS-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.

8.2.2. Supervision

PhD in progress : **Gautier Berthou** : *Operating system for transiently powered systems*, Inria, (IPL ZEP) since 01/2018.

PhD in progress : **Andrea Bocco** : *Proposition d'une unité de caclul UNUM pour le calcul scientifique*, ANR Metalibm grant, since 12/2016.

PhD in progress : **Luc Forget** : *Algèbre linéaire calculant au plus juste*, ANR Imprenum, since 10/2018.

PhD in progress :**Yanni Zhou** : *Full Duplex and spatial modulation* since 10/2018

PhD in progress : **Tarik Lassouaoui** : *Tag 2 Tag communication* since 10/2018

PhD in progress : **Regis Rousseau** : *Wireless Power Transfer* since 10/2018

PhD starting : **Maxime Christ** : *Learning in Very Low Precision* since 10/2018

PhD defended : **Yohan Uguen** : *High-level synthesis and arithmetic optimizations*, École Doctorale MathInfo,Nov. 2019.

PhD defended : **Tristan Delizy** : *Gestion de la mémoire dynamique pour les systèmes embarqués avec mémoire hétérogène*, École Doctorale MathInfo, Dec. 2019.

8.2.3. Juries

- Tanguy Risset was a member of the jury of the these following theses:
  - Arthur Hugeat (U. Bourgogne Franche-comté)
  - Tristan Delizy (U. Lyon)
- Tanguy Risset was a member of the jury of the HDR of Matthieu Gautier (U. Rennes)
- Florent de Dinechin was a reviewer for
  - the PhD of Geneviève Ndour (U. Rennes 1)
  - the PhD of Yohan Chatelain (U. Paris Saclay)

8.3. Popularization

8.3.1. Education

Florent de Dinechin gave a lecture in Grenoble for high-school teachers, in the framework of the ISN convention between Académie de Grenoble and Inria.

9. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] T. DELIZY.*Heterogeneous dynamic memory management for embedded devices*, Insa Lyon, December 2019, https://hal.archives-ouvertes.fr/tel-02429017

[2] Y. UGUEN.*High-level synthesis and arithmetic optimizations*, INSA Lyon, November 2019, https://hal.archives-ouvertes.fr/tel-02420901

Articles in International Peer-Reviewed Journal

[3] 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", September 2019, vol. 68, n^{o} 9, p. 1390 - 1403 [*DOI :* 10.1109/TC.2018.2889080], https://hal.archives-ouvertes.fr/hal-01954979

[4] J. BRUGUERA, F. DE DINECHIN.*Guest Editors Introduction: Special Section on Computer Arithmetic*, in "IEEE Transactions on Computers", July 2019, vol. 68, n^{o} 7, p. 951-952 [*DOI :* 10.1109/TC.2019.2918447], https://hal.inria.fr/hal-02151757

[5] P. TSAFACK, F. D. HUTU, E. TANYI, B. ALLARD.*RF Communication and IoT Paradigms System Proposal for Effective Consumption and Equity Distribution of Community Water in Developing Countries: A Case Study*, in "Industrial Engineering", July 2019, p. 1-9 [*DOI :* 10.11648/J.IE.20190301.12], https://hal.archives-ouvertes.fr/hal-02383267

Invited Conferences

[6] S. BENOUAKTA, F. D. HUTU, Y. DUROC.*New Approaches For Augmented UHF RFID Textile Yarn*, in "IEEE RFID-TA 2019 - 10th Annual IEEE International Conference on RFID Technology and Applications", Pise, Italy, 2019 IEEE International Conference on RFID Technology and Applications (RFID-TA), IEEE, September 2019, p. 1-5, https://hal.archives-ouvertes.fr/hal-02308051

International Conferences with Proceedings

[7] S. BENOUAKTA, S. KOLEY, F. D. HUTU, Y. DUROC.*Considerations on the equivalent electric models of a UHF RFID Helical Antenna Yarn*, in "ISSCS 2019 - 14-th International Symposium on Signals, Circuits and Systems", Iasi, Romania, International Symposium on Signals, Circuits and Systems, IEEE, July 2019, p. 1-4 [*DOI :* 10.1109/ISSCS.2019.8801787], https://hal.archives-ouvertes.fr/hal-02138670

[8] A. BOCCO, Y. DURAND, F. DE DINECHIN.*SMURF: Scalar Multiple-precision Unum Risc-V Floating-point Accelerator for Scientific Computing*, in "CoNGA 2019 - Conference on Next-Generation Arithmetic", Singapour, Singapore, ACM, March 2019, p. 1-8 [*DOI :* 10.1145/3316279.3316280], https://hal.inria.fr/hal-02087098

[9] L. FORGET, Y. UGUEN, F. DE DINECHIN, D. THOMAS.*A type-safe arbitrary precision arithmetic portability layer for HLS tools*, in "HEART 2019 - International Symposium on Highly Efficient Accelerators and Reconfigurable Technologies", Nagasaki, Japan, June 2019, p. 1-6 [*DOI :* 10.1145/3337801.3337809], https://hal.inria.fr/hal-02131798

[10] A. FUMTCHUM, F. D. HUTU, P. TSAFACK, G. VILLEMAUD, E. TANYI.*High Efficiency Rectifier for a Quasi-Passive Wake-up Radio*, in "ISSCS 2019 - 14-th International Symposium on Signals, Circuits and Systems", Iasi, Romania, 2019 International Symposium on Signals, Circuits and Systems, IEEE, July 2019, p. 1-4 [*DOI :* 10.1109/ISSCS.2019.8801754], https://hal.archives-ouvertes.fr/hal-02138667

[11] T. T. JOST, A. BOCCO, Y. DURAND, C. FABRE, F. DE DINECHIN, A. COHEN.*Variable Precision Floating-Point RISC-V Coprocessor Evaluation using Lightweight Software and Compiler Support*, in "CARRV 2019 - Third Workshop on Computer Architecture Research with RISC-V", Phoenix, AZ, United States, June 2019, p. 1-6, https://hal.inria.fr/hal-02161621

[12] S. KOLEY, S. BENOUAKTA, F. D. HUTU, Y. DUROC.*Passive UHF RFID yarn for relative humidity sensing*, in "EuCAP 2019 - 13th European Conference on Antennas and Propagation", Krakow, Poland, IEEE, March 2019, p. 1-4, https://hal.archives-ouvertes.fr/hal-02137612

[13] A. ROCA, S. RODRIGUEZ, A. SEGURA, K. MARQUET, V. BELTRAN.*A Linux Kernel Scheduler Extension for Multi-Core Systems*, in "HiPC 2019 - 26th IEEE International Conference on High Performance Computing, Data, and Analytics", Hyderabad, India, IEEE, December 2019, p. 1-10, https://hal.inria.fr/hal-02410082

[14] R. ROUSSEAU, G. VILLEMAUD, F. D. HUTU.*Efficiency of Wireless Power transfer with a Multi-sine Source Optimized for the Propagation Channel*, in "WPW 2019 - IEEE MTT's Wireless Power Week", Londre, United Kingdom, IEEE, June 2019, p. 1-4, https://hal.archives-ouvertes.fr/hal-02146410

[15] Y. UGUEN, L. FORGET, F. DE DINECHIN.*Evaluating the hardware cost of the posit number system*, in "FPL 2019 - 29th International Conference on Field-Programmable Logic and Applications (FPL)", Barcelona, Spain, September 2019, p. 106 - 113, https://hal.inria.fr/hal-02130912

[16] F. DE DINECHIN.*Reflections on 10 years of FloPoCo*, in "ARITH 2019 - 26th IEEE Symposium on Computer Arithmetic", kyoto, Japan, IEEE, June 2019, p. 1-3, https://hal.inria.fr/hal-02161527

[17] F. DE DINECHIN, S.-I. FILIP, L. FORGET, M. KUMM.*Table-Based versus Shift-And-Add constant multipliers for FPGAs*, in "ARITH 2019 - 26th IEEE Symposium on Computer Arithmetic", Kyoto, Japan, IEEE, June 2019, p. 1-8, https://hal.inria.fr/hal-02147078

[18] F. DE DINECHIN, L. FORGET, J.-M. MULLER, Y. UGUEN.*Posits: the good, the bad and the ugly*, in "CoNGA 2019 - Conference on Next-Generation Arithmetic", Singapore, Singapore, ACM Press, March 2019, p. 1-10 [*DOI :* 10.1145/3316279.3316285], https://hal.inria.fr/hal-01959581

National Conferences with Proceeding

[19] S. BENOUAKTA, S. KOLEY, F. D. HUTU, Y. DUROC.*Conception d'une antenne hélice pour fil textile RFID UHF extensible*, in "JNM 2019", Caen, France, 2019, p. 865-868, https://hal.archives-ouvertes.fr/hal-02137601

[20] L. FORGET, Y. UGUEN, F. DE DINECHIN.*Hardware cost evaluation of the posit number system*, in "Compas'2019 - Conférence d'informatique en Parallélisme, Architecture et Système", Anglet, France, June 2019, p. 1-7, https://hal.inria.fr/hal-02131982

[21] S. KOLEY, S. BENOUAKTA, F. D. HUTU, Y. DUROC.*Capteur d'humidité passif intégré dans un fil textile RFID UHF*, in "JNM 2019", Caen, France, May 2019, p. 474-478, https://hal.archives-ouvertes.fr/hal-02137608

Scientific Books (or Scientific Book chapters)

[22] K. MARQUET, J. COMBAZ, F. BERTHOUD.*Introduction aux impacts environnementaux du numérique*, in "1024, bulletin de la Société Informatique de France", April 2019, p. 85-97, https://hal.inria.fr/hal-02410129

Patents and standards

[23] A. BOCCO, Y. DURAND, F. DE DINECHIN.*Method and device for floating point representation with variable precision*, March 2019, n° 1902386, https://hal.inria.fr/hal-02169527

[24] G. VILLEMAUD, F. D. HUTU, L. DEMIA, H. BRAVO.*Installation de ventilation et procédé comprenant la transmission d'informations par propagation des ondes dans le réseau de conduites de ventilation*, March 2019, n° F1758247, https://hal.archives-ouvertes.fr/hal-02147025

Other Publications

[25] Y. UGUEN, F. DE DINECHIN, V. LEZAUD, S. DERRIEN.*Application-specific arithmetic in high-level synthesis tools*, October 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02423363

References in notes

[26] 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° 4, p. 40 -48

[27] E. BACCELLI, C. GÜNDOĞAN, O. HAHM, P. KIETZMANN, M. S. LENDERS, H. PETERSEN, K. SCHLEISER, T. C. SCHMIDT, M. WÄHLISCH.*RIOT: an Open Source Operating System for Low-end Embedded Devices in the IoT*, in "IEEE Internet of Things Journal", 2018

[28] G. BERTHOU, T. DELIZY, K. MARQUET, T. RISSET, G. SALAGNAC.*Peripheral State Persistence For Transiently Powered Systems*, Inria, 02 2017, n° RR-9018

[29] 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

[30] EQUIPEX.*Future internet of things*, 2011, http://fit-equipex.fr/

[31] F. D. 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, https://hal.archives-ouvertes.fr/hal-01845177

[32] 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

[33] J. MITOLA III.*The software radio*, in "IEEE National Telesystems Conference", 1992

[34] 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

[35] P. STEENKISTE, D. SICKER, G. MINDEN, D. RAYCHAUDHURI.*Future Directions in Cognitive Radio Network Research*, 2009, NSF Workshop Report

[36] M. WEISER.*The computer for the 21st Century*, in "Pervasive Computing, IEEE", January 2002, vol. 99, n^o 1, p. 19 -25

[37] 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

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]
Alain Girault [Inria, Senior Researcher, HDR]
Pascal Fradet [Inria, 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, Engineer, until Apr 2019]
Roger Pissard-Gibollet [Inria, Engineer, from Mar 2019]

PhD Students

Xiaojie Guo [Univ. Grenoble Alpes, PhD Student]
Maxime Lesourd [Univ. Grenoble Alpes, PhD Student]
Thomas Mari [Institut polytechnique de Grenoble, PhD Student, from Oct 2019]
Stephan Plassart [Univ. Grenoble Alpes, PhD Student]
Christophe Prévot [Inria, PhD Student, until Nov 2019]
Sihem Cherrared [Orange Labs, Univ. Rennes 1, PhD Student]

Arash Shafiei [Orange Labs, PhD Student]
Martin Vassor [Inria, PhD Student]

Post-Doctoral Fellow
Jia Jie Wang [Inria, Post-Doctoral Fellow]

Administrative Assistant
Helen Pouchot-Rouge-Blanc [Inria, Administrative Assistant]

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 distributed embedded systems as dynamic adaptive modular structures?
2. How to program reactive systems with real-time and resource constraints?
3. How to program fault-tolerant and explainable embedded systems?

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 [32], 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 [55]. Witness component-based toolsets such as PTOLEMY [47], BIP [38], or the modular architecture frameworks used, for instance, in the automotive industry (AUTOSAR) [30]. 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 [40]), 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 [36], [45], 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. [36]. 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 [41], [51]. 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 [61], natural sciences, law [62], and statistics [63], 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 [58], to allow the diagnosis of faults in a complex concurrent system [54], or to enforce accountability [57], 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 [53]), or is broken (*e.g.*, by limiting fault propagation [65]).

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. Current Industrial Cooperations

Regarding applications and case studies with industrial end-users of our techniques, we cooperate with Orange Labs on software architecture for cloud services.

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, Arash Shafiei, Jean-Bernard Stefani, Martin Vassor, Souha Ben Rayana.

6.1.1. Hypercells

The location graph framework we have introduced in [66] has evolved into the Hypercell framework presented in [18]. The Hypercell framework allows the definition of different component models for dynamic software architectures featuring both sharing and encapsulation. The basic behavioral theory of hypercells in the form of a contextual bisimulation has been developed and we are currently developing proofs of correctness for encapsulation policies based on this theory.

In collaboration with the Spirals team at Inria Lille – Nord Europe, and Orange, we have used hypercells as a pivot model for developing interpretations, formally defined with the Alloy specification language, of various languages and formalisms for the description of software configurations for cloud computing environments. Configuration languages considered include the TOSCA and OCCI standards, as well as the Open Stack Heat Orchestration Template (HOT), Docker Compose, and the Aeolus component model for cloud deployment. This work, developed as part of a bilateral contract with Orange, allowed the development of a verification tool for the correctness of HOT configurations, and helped uncover several flaws in the ETSI NFV standard.

6.1.2. 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) [49], [39], we have written a survey providing a comprehensive description of the existing parametric dataflow MoCs [42], and we have studied *symbolic* analyses of dataflow graphs [43]. More recently, we have proposed an original method to deal with lossy communication channels in dataflow graphs [48].

We are nowadays studying models allowing *dynamic reconfigurations* of the *topology* of the dataflow graphs. This is required by many modern streaming applications that 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) [13]. 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. Transformations can be seen as graph rewriting rules that match some sub-part of the dataflow graph and replace it by another one. Transformations can be applied an arbitrary number of times during execution and therefore can produce an arbitrary number of new graphs. The major feature and advantage of RDF is that it can be statically analyzed to guarantee that all possible graphs generated at runtime will be connected, consistent, and live. To the best of our knowledge, RDF is the only dataflow MoC allowing an arbitrary number of topological reconfigurations while remaining statically analyzable. It remains to complete the RDF implementation and to evaluate it on realistic case studies. Preliminary results indicate that dynamic reconfigurations can be implemented efficiently.

This is the research topic of Arash Shafiei's PhD, in collaboration with Orange Labs.

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 [46]. 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 [69]. 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 [35].

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 parallel programming language on PREcision Timed (hence deterministic) architectures [71], [72]. The originality resides in the time-triggered model of computation and communication that allows for a very precise control over the thread execution. Synchronization 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 [17]. This is a collaboration with Nicolas Hili and Eric Jenn, the postdoc of Nicolas Hili being funded by the CAPHCA project.

We have also proposed a *multi-rate* extension of FOREC [16]. Indeed, up to now FOREC programs were constrained to operate at a single rate, meaning that all the parallel threads had to share the same execution rate. While this simplified the semantics, it also represented a significant limitation.

Finally, we have extended the compiler of the PRET-C programming language [33], [34] in order to make it energy aware. PRET-C is a parallel programming language in the same sense as Esterel [44], meaning that the parallelism is "compiled away": the PRET-C compiler generates sequential code where the parallel threads from the source program are interleaved according to the synchronous semantics, and produces a classical Control Flow Graph (CFG). This CFG is then turned into a Timed Control Flow Graph (TCFG) by labeling each basic block with the number of clock cycles required to execute it on the chosen processor, based on its micro-architectural characteristics. From the TCFG, we use the method described in Section 6.2.5 to compute a Pareto front of non-dominated (worst-case execution time – WCET, worst-case energy consumption – WCEC) compromises.

6.2.2. 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.*, [67] 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 [64]. 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 [50]. 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.

In previous work we have proposed an approach using mode sequences as symbolic states for our abstractions [59]. 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 [12]. We have shown the effectiveness of the approach on examples inspired by road traffic regulation.

6.2.3. 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 minimize the energy consumption while guaranteeing that each job meets its deadline. The idea is to leverage on the *statistical information* on the jobs' characteristics available at design time: release time, worst-case execution time (WCET), and relative deadline. This is the topic of Stephan Plassart's PhD, funded by the CASERM Persyval project. We have considered several cases depending on the amount of information available at design time:

Offline case: In the offline case, all the information is known and we have proposed the first linear complexity offline scheduling algorithm that minimizes the total energy consumption [15]: our complexity is $\mathcal{O}(n)$ where n is the number of jobs to be scheduled, while the previously best known algorithms were in $\mathcal{O}(n^2)$ and $\mathcal{O}(n \log n)$ [60].

Clairvoyant case: In the clairvoyant case, the characteristics of the jobs are only known statistically, and each job's WCET and relative deadline are only known at release time. We want to compute the *optimal* online scheduling speed policy that minimizes the *expected* energy consumption while guaranteeing that each job meets its deadline. 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 the finite horizon case we use a dynamic programming algorithm, while in the infinite horizon case we use a value iteration algorithm [25].

Non-clairvoyant case: In the non-clairvoyant case, the actual execution time (AET) of a job is only known only when this job completes its execution. This AET is of course assumed to be less than the WCET, which is known at the job's release time. Again, by building an MDP for the system with a well chosen state, we compute the *optimal* online scheduling speed policy that minimizes the *expected* energy consumption [26].

Learning case: In the learning case, the only information known for the jobs are a bound on the jobs' WCETs and a bound on their deadlines. We have proposed two *reinforcement learning* algorithms, one that learns the optimal value of the expected energy (Q-learning), and another one that learns the probability transition matrix of the system, from which we derive the optimal online speed policy.

This work led us to compare several existing speed policies with respect to their feasibility. Indeed, the policies (OA) [70], (AVR) [70], and (BKP) [37] 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 [27].

6.2.4. Formal proofs for schedulability analysis of real-time systems

We contribute to Prosa [31], 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.

We have further developed CertiCAN, a tool produced using Coq for the formal certification of CAN analysis results [14]. 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 [68]. Additional optimizations have been implemented (and proved correct) to make CertiCAN computationally efficient. Experiments demonstrate that CertiCAN is able to certify the results of RTaW-Pegase, an industrial CAN analysis tool, even for large systems.

In addition, we have started investigating how to connect Prosa with implementations and less abstract models. Specifically, we have used Prosa to provide a schedulability analysis proof for RT-CertiKOS, a single-core sequential real-time OS kernel verified in Coq [20]. A connection with a timed-automata based formalization of the CAN specification is also in progress. Our objective with this line of research is to understand and bridge the gap between the abstract models used for real-time systems analysis and actual real-time systems implementation.

Finally, we contributed to a major refactoring of the Prosa library to make it more easily extendable and usable.

6.2.5. Scheduling under multiple constraints and Pareto optimization

We have completed a major work on embedded systems subject to multiple non-functional constraints, by proposing the first of its kind multi-criteria scheduling heuristics for a DAG of tasks onto an homogeneous multi-core chip [9], [23]. 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, namely DAGs of at most 8 tasks. Comparisons showed that the schedules produced by ERPOT are on average only 9% 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.

In a related line of work, we have considered the bi-criteria minimization problem in the (worst-case execution time – WCET, worst-case energy consumption – WCEC) space for real-time programs. To the best of our knowledge, this is the first contribution of this kind in the literature.

A real-time program is abstracted as a Timed Control Flow Graph (TCFG), where each basic block is labeled with the number of clock cycles required to execute it on the chosen processor at the nominal frequency. This timing information can be obtained, for instance, with WCET analysis tools. The target processor is equipped with dynamic voltage and frequency scaling (DVFS) and offers several (frequency f, voltage V) operating points. The goal is to compute a set of non-dominated points in the (WCET, WCEC) plane, non-dominated in the Pareto sense. Each such point is an assignment from the set of basic blocks of the TCFG to the set of available (f, V) pairs.

From the TCFG we extract the longest execution path, therefore deriving the WCET and the WCEC for a chosen fixed (f, V) pair. By construction, all the other execution paths are shorter, so this WCET and this WCEC hold for the whole program. This ensures that each single-frequency assignment is a non-dominated point. Then, we study two frequencies assignments, still for the longest execution path. When the frequency switching costs in time and in energy are assumed to be negligible, we demonstrate that each two frequencies (say with f_i and f_j) assignment is a point in the segment between the single frequency assignment at f_i and the single frequency assignment at f_j. We also propose a linear time heuristic to assign a (f, V) pair to all the other blocks (*i.e.*, those not belonging to the longest path) such that all the other execution paths have a shorter WCET and a lesser WCEC. A key result is that we demonstrate that any two frequencies assignment where the two frequencies are not contiguous is dominated either by a single frequency assignment or by a two frequencies assignment with contiguous frequencies. A corollary is that the Pareto front is a continuous piece-wise affine function. Finally, we generalize these results to the case where the frequency switching costs are not negligible. This is the topic of Jia Jie Wang's postdoc.

We evaluate our method and heuristic on a set of hard real time benchmark programs and we show that they perform extremely well. Our DVFS assignment algorithm can also be used as a back-end for the compiler of the PRET-C programming language [33], [34] in order to make it energy aware, thanks to the ability of this compiler to generate TCFGs (see Section 6.2.1).

6.3. Fault Management and Causal Analysis

Participants: Gregor Goessler, Jean-Bernard Stefani, Sihem Cherrared, Thomas Mari, Martin Vassor.

6.3.1. Fault Ascription in Concurrent Systems

Fault ascription is a precise form of fault diagnosis that relies on counterfactual analysis for pinpointing the causes of system failures. Research on counterfactual causality has been marked, until today, by a succession of definitions of causation that are informally validated against human intuition on mostly simple examples. This approach suffers from its dependence on the tiny number and incompleteness of examples in the literature, and from the lack of objective correctness criteria [52].

We have defined in [28] a set of expected properties for counterfactual analysis, and presented a refined analysis that conforms to our requirements. As an early study of the behavior of our analysis under abstraction we have established its monotony under refinement.

6.3.2. Causal Explanations in Discrete Event Systems

Model-Based Diagnosis of discrete event systems (DES) usually aims at detecting failures and isolating faulty event occurrences based on a behavioural model of the system and an observable execution log. The strength of a diagnostic process is to determine *what* happened that is consistent with the observations. In order to go a step further and explain *why* the observed outcome occurred, we borrow techniques from causal analysis.

In [21] we have presented two constructions of explanations that are able to extract the relevant part of a property violation that can be understood by a human operator. Both support partial observability of events. The first construction is based on minimal sub-sequences of the traces of the log that entail a violation of the property. The second approach is based on a construction of layers similar to [56], in which the explanation is constructed from the choices that definitely move the system closer to the violation of the property. Both approaches are complementary: while subsequence-based explanations are well suited to “condense” the execution trace in sequential portions of the model but are prone to keep non-pertinent parts such as initialisation sequences in the explanation, effective choice explanations highlight the “fateful” choices in an execution, as well as alternative events that would have helped avoid the outcome. Effective choice explanations are therefore able to explain failures stemming from non-deterministic choices, such as concurrency bugs.

6.3.3. Fault Management in Virtualized Networks

From a more applied point of view we have been 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.

In [29] we have proposed a model-based root cause analysis framework called SAKURA. In order to overcome the lack of accurate previous knowledge, SAKURA features a self-modeling algorithm that models the dependencies within and between layers of virtual networks, including auto-recovery and elasticity aspects. Model-based diagnosis is performed using constraint solving on the previous and acquired knowledge. As an illustration we have applied SAKURA to the virtual IpMultimedia Subsystem (vIMS).

Finally, in our survey on fault management in network virtualization environments [11] we have addressed the impact of virtualization on fault management, proposed a new classification of the recent fault management research achievements in network virtualization environments, and compared their major contributions and shortcomings.

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.

7.2. Bilateral Grants with Industry

With Orange:

- Fault Management in Multi-Tenant Programmable Networks. This CIFRE grant funds the PhD of Sihem Cherrared.
- 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.

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).

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 was completed in November 2019.

8.1.2. SEC: Construction of Safe Explainable Cyber-physical systems

Participants: Gregor Goessler, Thomas Mari.

In cyber-physical systems (CPS), software interacts with physical processes so as achieve desired functionalities. CPS are usually subject to safety and reliability requirements. Depending on the application, their failure may have unacceptable consequences, it is therefore crucial to ensure their correctness at design time. In addition, explainability of increasingly autonomous CPS is becoming crucial in order for the CPS to be socially acceptable.

The goal of this project is twofold. First, we will investigate a contract-based design approach for safe CPS in which different aspects – such as functional requirements, real-time constraints, and continuous behaviors – are modeled and verified separately. Second, we will leverage the contracts in order to ensure explainability of the system behavior by construction. By explainability we understand, informally, that for any behavior of the system we can automatically construct, from a log generated by the execution, an excerpt that retains only the events that causally contributed to the outcome, and that is easy to understand by a human expert.

The SEC project is supported by the "Initiatives de Recherche Stratégiques (IRS)" program of the IDEX UGA. It funds the PhD thesis of Thomas Mari, who will be co-advised by Gregor Gössler and Thao Dang (VERIMAG).

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 2022.

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.

The results obtained in 2019 in connection with the RT-proofs project are described in Section 6.2.4.

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 in Toulouse. The general objective of the project is to provide methods and tools to achieve both performance and determinism on modern, high-performance, multi-core and FPGA-enabled SOCs. Our specific contribution lies withing work packages dedicated to the design of novel PRET architectures and programming languages (see Section 6.2.1). This contract has yielded two publications so far [17], [16].

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 and the MPI-SWS in Kaiserslautern (Germany) on formal proofs for the analysis real-time systems. This collaboration is formalized by the ANR-PRCI project called RT-proofs started in 2018, which involves MPI-SWS, TU Braunschweig, Inria, and Onera.

8.4. International Initiatives

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

8.4.1.1. Quasar

Title: Quantitative systems formal verification

International Partner (Institution - Laboratory - Researcher):

CAS (China) - Department of Informatics - Lijun Zhang

Start year: 2019

The general scientific objectives are to extend formal analysis and verification methods such as model checking, process algebra and interactive theorem proving (Coq) to quantitative systems, more specifically probabilistic and quantum computing systems. Application fields include compositional modeling for dynamic real-time probabilistic software architectures and risk analysis. The collaboration will involve active scientists on all these fields not only from Inria and Inst Soft. CAS, but also from CWI, Verimag Grenoble, ECNU Shanghai, and partners of CWI (VU Amsterdam and Twente).

9. Dissemination

9.1. Environmental and societal responsibility

NB: We support the idea of an additional section in future Activity Reports that would be dedicated to actions in connection with our environmental and societal responsibility. We write below content that would fit into it for 2019.

- A discussion on computer ethics was scheduled at our last team seminar. There have been several meetings between permanent researchers to discuss how we can better align our research agenda with the current needs of our society. Our discussions cover aspects such as the environmental and societal impact of ICT (Information and Communication Technologies), and more broadly the role that ICT play in our lives, our role as researchers in computer science, etc. These discussions have led us to consider these issues as a research topic that we should investigate in the near future.
- Sophie Quinton and Jean-Bernard Stefani contributed to the so-called “MakeSEnS” working group appointed by the CEO of Inria to propose a list of concrete actions that the institute could take to tackle the current environmental crisis [24].
- Sophie Quinton has been mandated by Patrick Gros (Director of the Inria Grenoble – Rhône-Alpes research center) to organize discussions and actions regarding the environmental and societal impact of our research at Inria Grenoble Rhône-Alpes.
- Sophie Quinton is a member of the GDS EcoInfo (https://ecoinfo.cnrs.fr/). Her actions as member of EcoInfo include: leading a discussion group on the environmental impact of collaborative platforms as part of a two-day seminar organized at Mines ParisTech [0]; and contributing to a roadmap regarding the environmental and societal impact of ICT prepared by the Conseil National du Numérique (https://cnnumerique.fr/) at the government’s request.
- Sophie Quinton co-chairs a working group of the GDR CIS associated with the Center for Internet and Society (http://cis.cnrs.fr/) focused on environmental issues.
- Sophie Quinton is part of the scientific committee of the upcoming “COP2 étudiante” (https://cop2etudiante.org/).

9.2. Promoting Scientific Activities

9.2.1. Scientific Events: Organisation

9.2.1.1. 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.

9.2.1.2. Member of the Organizing Committees

- Alain Girault was co-organizer of the Workshop on Synchronous Programming (SYNCHRON’19) http://synchron19.org.
- Gregor Gössler was co-organizer of a Shonan seminar on *Causal reasoning in systems*https://project.inria.fr/shonan139.

9.2.2. Scientific Events: Selection

9.2.2.1. Chair of Conference Program Committees

- Sophie Quinton was the program chair of 31st Euromicro Conference on Real-Time Systems (ECRTS’19) https://www.ecrts.org/archives/fileadmin/WebsitesArchiv/ecrts2019/index.html.

9.2.2.2. Member of the Conference Program Committees

[0] http://www.mines-paristech.fr/Actualites/Ingenieurs-et-transitions-environnementales/4116

- Alain Girault served in the program committees of the Forum on specification and Design Languages (FDL'19) and the Conference on Applications of Concurrency to System Design (ACSD'19).
- Gregor Gössler served in the program committees of the ACM/IEEE International Conference on Embedded Software (EMSOFT'19), the 26th SPIN Symposium on Model Checking of Software (SPIN'19), the 17th ACM-IEEE International Conference on Formal Methods and Models for System Design (MEMOCODE'19), and the 4th international Workshop on Formal Reasoning about Causation, Responsibility, and Explanations in Science and Technology (CREST'19).

9.2.2.3. Reviewer

- Alain Girault reviewed an article for the International Conference on Fundamental Approaches to Software Engineering (FASE'19).
- Sophie Quinton reviewed an article for the ACM International Conference on Embedded Software (EMSOFT'19).
- Pascal Fradet reviewed an article for the ACM International Conference on Embedded Software (EMSOFT'19).

9.2.3. Journal

9.2.3.1. Member of the Editorial Boards

- Alain Girault is a member of the editorial board of the Journal on Embedded Systems.

9.2.3.2. Reviewer - Reviewing Activities

- Alain Girault reviewed articles for IEEE Transactions on Service Computing (TSC) and for the International Journal on Software Tools for Technology Transfer (STTT).
- Gregor Gössler reviewed an article for IEEE Transactions on Automatic Control (TAC).

9.2.4. Leadership within the Scientific Community

- Sophie Quinton is a member of the ACM SIGBED Executive Committee and Associate Editor of the SIGBED Review.

9.2.5. 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 Deputy Scientific Director at Inria in charge of the domain "Algorithmics, Programming, Software and Architecture". He is also Scientific Manager of the Cyber-Physical Systems axis of the Persyval-Lab labex https://persyval-lab.org.
- 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.

9.3. Teaching - Supervision - Juries

9.3.1. 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, Analyse de Code pour la Sûreté et la Sécurité, 45 HeqTD, niveau M1, Grenoble INP (Ensimag), France

Licence : Xavier Nicollin, Théorie des Langages 1, 48 HeqTD, niveau L3. Grenoble INP (Ensimag), France

Licence : Xavier Nicollin, Théorie des Langages 2, 37,5 HeqTD, niveau L3, Grenoble INP (Ensimag), France

Licence : Xavier Nicollin, Bases de la Programmation Impérative, 30 HeqTD, niveau L3, Grenoble INP (Ensimag), France

Licence : Sophie Quinton, Théorie des Langages 2, 20 HeqTD, niveau L3, Grenoble INP (Ensimag), France

Master : Sophie Quinton, Performance and Quantitative Properties, 8 HeqTD, MOSIG, Univ. Grenoble Alpes, France

Master: Jean-Bernard Stefani, Formal Aspects of Component Software, 9h, MOSIG, Univ. Grenoble Alpes, France.

9.3.2. 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: Thomas Mari, "Construction of Safe Explainable Cyber-physical systems"; Grenoble INP; since October 2019; co-advised by Gregor Gössler and Thao Dang.
- PhD: Christophe Prévot, "Early Performance assessment for evolving and variable Cyber-Physical Systems", Univ. Grenoble Alpes; defended on November 15th, 2019; co-advised by Alain Girault and Sophie Quinton.
- PhD: Zain A. H. Hammadeh, "Deadline Miss Models for Temporarily Overloaded Systems", TU Braunschweig; defended on May 28th, 2019; co-advised by Rolf Ernst 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, "RDF: A reconfigurable dataflow MoC supporting dynamic topological transformations and static analyzability"; 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.
- Internships: Jonathan Julou and Martin Portalier, "Formal proofs for real-time systems analysis"; ENSIMAG; co-advised by Pascal Fradet, Xiaojie Guo, Maxime Lesourd and Sophie Quinton.

9.3.3. Juries

- Alain Girault was referee for the Habilitation thesis of Julien De Antoni (Université Côte d'Azur) and was member of the PhD thesis of Pierre Donat-Bouillud (Sorbonne Université). He was also vice-president of the Inria Starting Research Position (SRP) Advanced Research Position (ARP) jurys.

- Jean-Bernard Stefani was examiner for the Habilitation (HDR) jury of Thomas Ledoux (Université de Nantes).
- Sophie Quinton was a member of the PhD thesis committee of Hugo Daigmorte (Université de Toulouse).

9.4. Popularization

9.4.1. Articles and contents

Software is a fundamental pillar of modern scientific research, across all fields and disciplines. However, there is a lack of adequate means to cite and reference software due to the complexity of the problem in terms of authorship, roles and credits. This complexity is further increased when it is considered over the lifetime of a software that can span up to several decades. Building upon the internal experience of Inria, we have provided a contribution to the ongoing efforts in order to develop proper guidelines and recommendations for software citation and reference. Namely, we recommend: (1) a richer taxonomy for software contributions with a qualitative scale; (2) to put humans at the heart of the evaluation; and (3) to distinguish citation from reference [10]. This has been a joint work between Alain Girault and colleagues from Software Heritage (Roberto Di Cosmo), from Inria's Evaluation Committee (Pierre Alliez, Benjamin Guedj, Mohand-Saïd Hacid, Nicolas Rougier), and a specialist in reproducible research (Arnaud Legrand).

9.4.2. Interventions

Sophie Quinton played the role of one of the experts in the fictional trial of a robot during the Transfo Festival in Grenoble. She also gave a lecture on the design and verification of real-time systems as part of the ISN curriculum for high school teachers.

Alain Girault played the role of one of the experts in the fictional trial of a robot during the Fête de la Science in Grenoble.

9.4.3. Creation of media or tools for science outreach

All SPADES members have contributed to the scenario of an animated movie, commissioned by the communication service of Inria, in order to popularise the research activities of SPADES.

10. Bibliography

Major publications by the team in recent years

[1] S. ANDALAM, P. S. ROOP, A. GIRAULT, C. TRAULSEN. *A Predictable Framework for Safety-Critical Embedded Systems*, in "IEEE Trans. on Computers", July 2014, vol. 63, n° 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° 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° 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] A. ABDI, A. GIRAULT, H. ZARANDI.*ERPOT: A Quad-Criteria Scheduling Heuristic to Optimize Execution Time, Reliability, Power Consumption and Temperature in Multicores*, in "IEEE Transactions on Parallel and Distributed Systems", October 2019, vol. 30, n^o^ 10, p. 2193-2210 [*DOI :* 10.1109/TPDS.2019.2906172], https://hal.inria.fr/hal-02400019

[10] P. ALLIEZ, R. DI COSMO, B. GUEDJ, A. GIRAULT, M.-S. HACID, A. LEGRAND, N. P. ROUGIER.*Attributing and Referencing (Research) Software: Best Practices and Outlook from Inria*, in "Computing in Science & Engineering", 2019, p. 1-14, https://arxiv.org/abs/1905.11123 [*DOI :* 10.1109/MCSE.2019.2949413], https://hal.archives-ouvertes.fr/hal-02135891

[11] S. CHERRARED, S. IMADALI, E. FABRE, G. GÖSSLER, I. G. B. YAHIA.*A Survey of Fault Management in Network Virtualization Environments: Challenges and Solutions*, in "IEEE Transactions on Network and Service Management", October 2019, p. 1-15 [*DOI :* 10.1109/TNSM.2019.2948420], https://hal.inria.fr/hal-02370378

[12] A. GIRARD, G. GÖSSLER.*Safety Synthesis for Incrementally Stable Switched Systems using Discretization-Free Multi-Resolution Abstractions*, in "Acta Informatica", 2019 [*DOI :* 10.1007/S00236-019-00341-X], https://hal.archives-ouvertes.fr/hal-02286661

International Conferences with Proceedings

[13] P. FRADET, A. GIRAULT, R. KRISHNASWAMY, X. NICOLLIN, A. SHAFIEI.*RDF: Reconfigurable Dataflow*, in "DATE 2019 - Design, Automation & Test in Europe Conference & Exhibition", Florence, Italy, March 2019, p. 1709-1714 [*DOI :* 10.23919/DATE.2019.8714987], https://hal.inria.fr/hal-01960788

[14] P. FRADET, X. GUO, J.-F. MONIN, S. QUINTON.*CertiCAN: A Tool for the Coq Certification of CAN Analysis Results*, in "RTAS 2019 - 25th IEEE Real-Time and Embedded Technology and Applications Symposium", Montreal, Canada, IEEE, April 2019, p. 1-10 [*DOI :* 10.1109/RTAS.2019.00023], https://hal.archives-ouvertes.fr/hal-02119024

[15] B. GAUJAL, A. GIRAULT, S. PLASSART.*A Linear Time Algorithm for Computing Off-line Speed Schedules Minimizing Energy Consumption*, in "MSR 2019 - 12ème Colloque sur la Modélisation des Systèmes Réactifs", Angers, France, November 2019, p. 1-14, https://hal.archives-ouvertes.fr/hal-02372136

[16] A. GIRAULT, N. HILI, É. JENN, E. YIP.*A Multi-Rate Precision Timed Programming Language for Multi-Cores*, in "FDL 2019 - Forum for Specification and Design Languages", Southampton, United Kingdom, IEEE, September 2019, p. 1-8 [*DOI :* 10.1109/FDL.2019.8876950], https://hal.inria.fr/hal-02399998

[17] N. HILI, A. GIRAULT, É. JENN.*Worst-Case Reaction Time Optimization on Deterministic Multi-Core Architectures with Synchronous Languages*, in "RTCSA2019 2019 - 25th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA)", Hangzhou, China, IEEE, August 2019, p. 1-11 [*DOI :* 10.1109/RTCSA.2019.8864570], https://hal.inria.fr/hal-02400009

[18] J.-B. STEFANI, M. VASSOR.*Encapsulation and Sharing in Dynamic Software Architectures: The Hypercell Framework*, in "FORTE 2019 - 39th International Conference on Formal Techniques for Distributed Objects, Components, and Systems (FORTE)", Copenhagen, Denmark, J. A. PÉREZ, N. YOSHIDA (editors), Formal Techniques for Distributed Objects, Components, and Systems, Springer International Publishing, 2019, vol. LNCS-11535, p. 242-260, Part 1: Full Papers [*DOI :* 10.1007/978-3-030-21759-4_14], https://hal.inria.fr/hal-02313751

Conferences without Proceedings

[19] B. GAUJAL, A. GIRAULT, S. PLASSART.*A Linear Time Algorithm for Computing Off-line Speed Schedules Minimizing Energy Consumption*, in "MSR 2019 - 12ème Colloque sur la Modélisation des Systèmes Réactifs", Angers, France, November 2019, p. 1-14, https://hal.archives-ouvertes.fr/hal-02432311

[20] X. GUO, M. LESOURD, M. LIU, L. RIEG, Z. SHAO.*Integrating Formal Schedulability Analysis into a Verified OS Kernel*, in "Computer Aided Verification", New York, United States, July 2019, p. 496-514 [*DOI :* 10.1007/978-3-030-25543-5_28], https://hal.archives-ouvertes.fr/hal-02289494

[21] G. GÖSSLER, T. MARI, Y. PENCOLÉ, L. TRAVÉ-MASSUYÈS.*Towards Causal Explanations of Property Violations in Discrete Event Systems*, in "DX'19 - 30th International Workshop on Principles of Diagnosis", Klagenfurt, Austria, November 2019, p. 1-8, https://hal.inria.fr/hal-02369014

[22] A. N. SYLLA, K. GUILLOUARD, F. KLAMM, M. OUZZIF, P. MERLE, S. BEN RAYANA, J.-B. STEFANI.*Formal Verification of Orchestration Templates for Reliable Deployment with OpenStack Heat*, in "CNSM 2019 - 15th International Conference on Network and Service Management", Halifax, Canada, October 2019, p. 1-5, https://hal.inria.fr/hal-02375386

Research Reports

[23] 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 ; 37, March 2019, n° RR-9196, p. 1-37, https://hal.inria.fr/hal-01848087

[24] F. BERTHOUD, P. GUITTON, L. LEFÈVRE, S. QUINTON, A. ROUSSEAU, J. SAINTE-MARIE, C. SERRANO, J.-B. STEFANI, P. STURM, E. TANNIER.*Sciences, Environnements et Sociétés : Rapport long du groupe de travail MakeSEnS d'Inria*, Inria, October 2019, https://hal.inria.fr/hal-02340948

[25] B. GAUJAL, A. GIRAULT, S. PLASSART.*A Discrete Time Markov Decision Process for Energy Minimization Under Deadline Constraints*, Grenoble Alpes ; Inria Grenoble Rhône-Alpes, Université de Grenoble, December 2019, n^o RR-9309, 46, https://hal.inria.fr/hal-02391948

[26] B. GAUJAL, A. GIRAULT, S. PLASSART.*Exploiting Job Variability to Minimize Energy Consumption under Real-Time Constraints*, Inria Grenoble Rhône-Alpes, Université de Grenoble ; Université Grenoble - Alpes, November 2019, n° RR-9300, 23, https://hal.inria.fr/hal-02371742

[27] B. GAUJAL, A. GIRAULT, S. PLASSART.*Feasibility of on-line speed policies in real-time systems*, Inria Grenoble Rhône-Alpes, Université de Grenoble ; Univ. Grenoble Alpes, November 2019, n° RR-9301, 38, https://hal.inria.fr/hal-02371996

[28] G. GÖSSLER, J.-B. STEFANI.*Causality Analysis and Fault Ascription in Component-Based Systems*, Inria - Research Centre Grenoble – Rhône-Alpes, June 2019, n° RR-9279, p. 1-27, https://hal.inria.fr/hal-02161534

Other Publications

[29] S. CHERRARED, S. IMADALI, E. FABRE, G. GÖSSLER.*SAKURA a Model Based Root Cause Analysis Framework for vIMS*, ACM Press, June 2019, p. 594-595, MobiSys 2019 - 17th ACM International Conference on Mobile Systems, Applications, and Services, Poster, https://hal.inria.fr/hal-02291163

References in notes

[30] *Automotive Open System Architecture*, 2003, http://www.autosar.org

[31] *A Library for formally proven schedulability analysis*, http://prosa.mpi-sws.org/

[32] ARTEMIS JOINT UNDERTAKING.*ARTEMIS Strategic Research Agenda*, 2011

[33] S. ANDALAM, P. S. 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

[34] S. ANDALAM, P. S. ROOP, A. GIRAULT, C. TRAULSEN.*A Predictable Framework for Safety-Critical Embedded Systems*, in "IEEE Transactions on Computers", July 2014, 13, https://hal.inria.fr/hal-01095468

[35] 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

[36] 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° 4

[37] N. BANSAL, T. KIMBREL, K. PRUHS.*Speed Scaling to Manage Energy and Temperature*, in "Journal of the ACM", 2007, vol. 54, n° 1

[38] 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° 3

[39] 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

[40] 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

[41] S. BORKAR.*Designing Reliable Systems from Unreliable Components: The Challenges of Transistor Variability and Degradation*, in "IEEE Micro", 2005, vol. 25, n^{o} 6

[42] 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

[43] 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

[44] F. BOUSSINOT, R. DE SIMONE.*The Esterel Language*, in "Proceedings of the IEEE", September 1991, vol. 79, n^{o} 9, p. 1293–1304

[45] R. DAVIS, A. BURNS.*A Survey of Hard Real-Time Scheduling for Multiprocessor Systems*, in "ACM Computing Surveys", 2011, vol. 43, n^{o} 4

[46] S. A. EDWARDS, E. A. LEE.*The Case for the Precision Timed (PRET) Machine*, in "44th Design Automation Conference (DAC)", IEEE, 2007

[47] 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

[48] 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

[49] P. FRADET, A. GIRAULT, P. POLPAVKO.*SPDF: A schedulable parametric data-flow MoC*, in "Design, Automation and Test in Europe, DATE'12", IEEE, 2012

[50] 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

[51] 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

[52] C. GLYMOUR, D. DANKS, B. GLYMOUR, F. EBERHARDT, J. RAMSEY, R. SCHEINES, P. SPIRTES, C. M. TENG, J. ZHANG.*Actual causation: a stone soup essay*, in "Synthese", 2010, vol. 175, n^{o} 2, p. 169–192

[53] F. C. GÄRTNER.*Fundamentals of Fault-Tolerant Distributed Computing in Asynchronous Environments*, in "ACM Computing Surveys", 1999, vol. 31, n^{o} 1

[54] 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

[55] T. HENZINGER, J. SIFAKIS.*The Embedded Systems Design Challenge*, in "Formal Methods 2006", Lecture Notes in Computer Science, Springer, 2006, vol. 4085

[56] H. JIN, K. RAVI, F. SOMENZI.*Fate and free will in error traces*, in "STTT", 2004, vol. 6, n^o 2, p. 102–116, http://dx.doi.org/10.1007/s10009-004-0146-9

[57] R. KÜSTERS, T. TRUDERUNG, A. VOGT.*Accountability: definition and relationship to verifiability*, in "ACM Conference on Computer and Communications Security", 2010, p. 526-535

[58] 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

[59] E. LE CORRONC, A. GIRARD, G. GÖSSLER.*Mode Sequences as Symbolic States in Abstractions of Incrementally Stable Switched Systems*, in "CDC - 52nd Conference on Decision and Control - 2013", IEEE, December 2013, p. 3225-3230, http://hal.inria.fr/hal-00924815

[60] M. LI, F. YAO, H. YUAN.*An $O(n^2)$ Algorithm for Computing Optimal Continuous Voltage Schedules*, in "Annual Conference on Theory and Applications of Models of Computation, TAMC'17", Bern, Switzerland, LNCS, April 2017, vol. 10185, p. 389–400

[61] P. MENZIES.*Counterfactual Theories of Causation*, in "Stanford Encyclopedia of Philosophy", E. ZALTA (editor), Stanford University, 2009, http://plato.stanford.edu/entries/causation-counterfactual

[62] M. MOORE.*Causation and Responsibility*, Oxford, 1999

[63] J. PEARL.*Causal inference in statistics: An overview*, in "Statistics Surveys", 2009, vol. 3, p. 96-146

[64] 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

[65] J. RUSHBY.*Partitioning for Safety and Security: Requirements, Mechanisms, and Assurance*, NASA Langley Research Center, 1999, n^o CR-1999-209347

[66] 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

[67] P. TABUADA.*Verification and Control of Hybrid Systems - A Symbolic Approach*, Springer, 2009

[68] 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

[69] 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

[70] F. YAO, A. DEMERS, S. SHENKER.*A scheduling model for reduced CPU energy*, in "Proceedings of lEEE Annual Foundations of Computer Science", 1995, p. 374–382

[71] E. YIP, P. S. ROOP, M. BIGLARI-ABHARI, A. GIRAULT.*Programming and Timing Analysis of Parallel Programs on Multicores*, in "International Conference on Application of Concurrency to System Design, ACSD'13", Barcelona, Spain, IEEE, July 2013, p. 167–176, https://hal.inria.fr/hal-00842402

[72] E. YIP, P. S. ROOP, A. GIRAULT, M. BIGLARI-ABHARI.*Synchronous Deterministic Parallel Programming for Multicores with ForeC*, Inria - Research Centre Grenoble – Rhône-Alpes, August 2016, n^{o} RR-8943, https://hal.inria.fr/hal-01351552

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]
Jean-Yves Courtonne [Inria, Starting Research Position, from Sep 2019]
Pierre-Yves Longaretti [CNRS, Researcher]
Peter Sturm [Inria, Senior Researcher, HDR]

Faculty Members
Denis Dupré [Univ Grenoble Alpes, Associate Professor]
Serge Fenet [Univ Claude Bernard, Associate Professor]
Guillaume Mandil [Univ Grenoble Alpes, Associate Professor]
Régis Perrier [Univ Grenoble Alpes, Associate Professor]

Technical Staff
Christophe Boisseau [Inria, Engineer]
Quentin Desvaux [Inria, Engineer, from Oct 2019 until Nov 2019]
Roger Pissard-Gibollet [Inria, Engineer, from Mar 2019]
Steven Quinito Masnada [Inria, Engineer, from Apr 2019 until Dec 2019]

PhD Students
Michela Bevione [Inria, PhD Student]
Mathilde Jochaud Du Plessix [Univ Grenoble Alpes, PhD Student, from Oct 2019]
François-Rémi Mazy [Inria, PhD Student, from Oct 2019]

Post-Doctoral Fellows
Jean-Yves Courtonne [Inria, Post-Doctoral Fellow, until Aug 2019, granted by Agence de l'environnement et de la maîtrise de l'énergie]
Fausto Lo Feudo [Univ Grenoble Alpes, Post-Doctoral Fellow, until Aug 2019]

Administrative Assistant
Marie-Anne Dauphin-Rizzi [Inria, Administrative Assistant]

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.

3.4. Global systemic risks

Modern societies are characterized by a very high level of global interconnections between many sectors of economic, social and political activity, as well as by the environmental impacts produced by these activities and their consequences for human societies themselves. The resulting generalized interdependences induce intrinsic, or "systemic", risks of instability. These risks constitute serious threats for the global socio-ecological system, and the issue of a potential global collapse is part of the threats to be analyzed.

Global systemic risks directly relate to the STEEP team project, i.e., the question of sustainability at various spatial scales. The ability of socio-ecosystems, local communities, nation States and the international community to address these risks is a key factor in a sustainability perspective. However, the topic of global systemic risks was not until recently within the scope of the research activity of the team.

Within academia, the activity of several research institutes is devoted to these risks, with a strong contribution of social sciences. The most representative are probably the *Global Systemic Risk Institute* of Princeton [0], the *Center for Risk Studies* of Cambridge (UK) Cambridge [0], and the *Risk Center* of Zürich [0]. Various teams are also active on these themes, but with a more sectorial focus.

From the point of view of the main processes at work, global systemic risks can be grouped into two categories

1. Risks related to long term mean trends (several decades). For the most part, they are generated by the growing tension between our use of resources and pollution production and our (semi)-natural environment capacity to absorb the associated impacts. They also emerge from the back-reaction of these environmental changes on socio-ecosystems. These risks are underlined and amplified by specific historical, socio-politic and economic dynamics.
2. Risks related to systemic contagion effect, on much shorter terms (weeks or months), but sporadic and random. This type of risk is directly driven by the high level of interconnection of various

[0] https://risk.princeton.edu/
[0] https://www.jbs.cam.ac.uk/faculty-research/centres/risk/
[0] https://riskcenter.ethz.ch/

large scale human activities, to intrinsic instabilities whose very existence is directly tied to these interconnections, and to the potential propagation of these instabilities in all sectors of activity, in a domino-like effect. These risks are amplified by current geopolical dynamics and by the deepening of the various environmental crises.

In such a context, and due to its areas of expertise, the STEEP team has a nearly unique ability and opportunity to make a significant contribution to these questions, most particularly on the modelling front.

The World3 model is without any doubt the most representative of the first category of risks. It was developed by Meadows and coworkers for their famous report on the limits to growth [18], [20]. The reinvestigation by [22], [23] and re-discussion by [13] have renewed the interest for this model, while raising more pinpointed questions on the robustness and validity of the conclusions drawn from it. We have started to address these questions on three different fronts:

1. Through an analysis of the parameterization choices performed in the mode. In practice the team has undertaken a sensitivity analysis that is substantially more comprehensive than previous attempts in this direction.
2. Through an analysis of the modelling choices performed by the Meadows group. This will consist in a partial sectoral and spatial disaggregation of the model.
3. Through some elements of epistemological analysis.

As a matter of fact, two internships have already been devoted to the sensitivity analysis of the model. This is now pursued through a PhD thesis, started last Fall. The student (Mathilde Duplessix) will also address the other two points in the course of her PhD.

The main practical interest of this work is related to the possibility of distinguishing a potential onset of collapse in the first or second half ot the century. These two broad options imply different mitigation/adaptation strategies, that need to be correctly anticipated.

On the front of systemic contagion risks, and although a comprehensive analysis of the whole range of potential risks is impossible in an exploratory phase, the nexus energy/finance/supply chains plays a particular role in our societies and present a specific level of criticity. Some sectorial (and even cross-sectorial) aspects of this nexust have been discussed in the literature (e.g., [16], [17], [14]), but apparently no global, generic has been produced so far. Such a model would constitute by itself a remarkable breakthrough in this topic.

Funding for these research objectives has been obtained this year, in the form of an Inria “Action Exploratoire” project which officially starts in January 2020. This funding covers a PhD (Louis Delannoy, starting in January 2020) and a post-doctoral position (scheduled to start in 2021).

4. Application Domains

4.1. Introduction

In the context described in the previous sections, we can distinguish two connected and complementary strategies for analyzing environmental pressures: a sectorial approach and a spatial one. The first one is more directly connected to ecological accounting, the second one has more direct relations to urban economy and land cover modelling. We start by describing the former.

4.2. Ecological accounting for sectorial pressure assessment

One of the major issues in the assessment of the long-term sustainability of urban areas is related to the concept of “imported sustainability”. Cities bring in from the outside most of their material and energy resources, and reject to the outside the waste produced by their activity. The modern era has seen a dramatic increase in both volume and variety of these material flows and consumption as well as in distance of origin and destination of these flows, usually accompanied by a spectacular increase in the associated environmental impacts. A realistic assessment of the sustainability of urban areas requires to quantify both local and distant environmental impacts; greenhouse gas emissions are only one aspect of this question. Such an assessment brings to light the most relevant direct and indirect lines of action on these issues. In this respect, it is useful to introduce the alternative concepts of consumer versus producer responsibility (or point of view).

The producer point of view is the most useful to pinpoint relevant direct lines of actions on environmental pressures due to production. In other respects, any territory imports and exports goods and services from and to the rest of the world. The consumer point of view provides information on the indirect pressures associated with these exchanges, as production responds to a final demand. Tracking the various supply chains through the analysis of the structure of the local economy and its relations and dependencies to the external world allows us to identify critically important contributions to environmental pressures; this also enables us to define fair environmental indicators in order not to attribute environmental pressures to producers only (whose responsibility is the easier to quantify of the two). In this approach, the producer responsibility follows directly from the measurement of its energy and material uses, while the consumer responsibility is established indirectly through an allocation of the impacts of production to the final consumers, but this second mode of allocation is to some extent virtual and partly subjective. Four methods stand out:

- Material Flow Analysis (MFA)
- Input-Output Analysis (IOA)
- Life-Cycle Analysis (LCA)
- Ecological Footprint (EF)

Each of these is based on a well-defined structuring element: mass conservation for MFA, measure of industrial inter-dependencies for IOA, identification of all the steps from cradle to grave for LCA, measure of biocapacity demand for EF. The different methods have preferred areas of application. For example, EF is more relevant for analyzing primary production such as agricultural staples, wood, etc. IOA is more focused on whole industrial sectors, while LCA is geared towards end-user products, taken as functional units; finally, primary materials (such as metals), waste and emissions are more easily characterized through MFA. Methodological choices are driven by the type of question one needs to address, data availability and collection method and the spatial scales under consideration. Indeed, data can be used in two different ways: bottom-up or top-down. The bottom-up data is more precise, but in general precludes comprehensiveness; on the contrary, the top-down data is by nature more comprehensive, but is not suited for a detailed, fine-scale analysis of the results.

STEEP is pursuing its research program on this theme with three major goals: 1) Creating a comprehensive database enabling pressure analyses; 2) Developing methodologies and models resolving scaling issues, and developing algorithms allowing us to rigorously and automatically obtain adequate assessments; 3) Providing a synthetic analysis of environmental pressures associated to the major material flows, at various geographic levels (employment catchment area, département and région, for France), with the explicit aim of incorporating this type of information in the public decision process on environmental issues, via specifically designed decision-help procedures.

4.3. Urban economy and land use/land cover changes: assessment of spatial distributions of the pressures

The preceding section was focused on territorial metabolism, in particular on the analysis of supply chains. Here territories are examined with a more prominent emphasis on their spatial dimension, with attention to: the spatial distribution of local pressures previously identified (from a land use point of view), and the modeling of future land use and activity location (from an economic point of view). These two questions correspond to very different modeling strategies: the first one is more statistical in nature, extrapolating future land use from past evolution combined with global territory scenarios; the other one has a more fundamental flavor and focuses on an understanding of the processes driving urbanization. For this, we focus more precisely on the question of household and businesses choices of localization, as well as on spatial fluxes within the territory (transportation of goods and persons). The critical point here is to understand and manage urban sprawl and its environmental effects (GHG emission, loss of arable land, ecosystem fragmentation, and so on).

4.3.1. Land Use/Land Cover Change models (LUCC)

LUCC models are mostly used in environmental sciences, e.g. to evaluate the impact of climate change on agriculture, but they can also be used to analyze urban sprawl. There is a variety of models, static or dynamic, grid- or agent- based, local or global, etc., and with varying degrees of sophistication concerning spatio-temporal analysis or decision structures incorporated in the model.

The models of interest here are statistical in nature but spatially explicit. Following decades of development, they are robust, versatile and mature. In principle, agent-models have a larger potential for representing decision processes, but in practice this advantage results in a loss of universality of the models. Among the most well-known and most mature models, one can mention the CLUE family of models, DINAMIC, or LCM (Land Change Modeler). These models are well described in the literature, and will only be briefly presented here.

These models analyze change in land use in a statistical way; they are structured around three different modules:

- The first module determines the probability of change of pixels of the territory (pixels are typically tens to hundreds of meters in size).
- The second module defines the global changes between the various land uses of interest per time step (usually, a few years), based on global scenarios of evolution of the territory under study. These first two modules are independent of one another.
- The last module distributes changes of land use in an explicit manner, pixel per pixel, at each time step, on the basis of the information provided by the first two modules.

Probabilities of change are calibrated on past evolution, from the differences between two past maps of land use in the more favorable cases, or from a single map otherwise (under the assumption that the logic of occupation changes is the same as the logic of land use at this single date). Such changes are then characterized in a statistical way with the help of modeling variables identified by the modeler as having potential explaining or structuring power (typically, a few to a dozen variables are used for one type of land use change). For example, in the case of urban sprawl, typical explaining factors are the distance to existing urbanized zones or distances to roads and other means of transportation, elements of real estate costs, etc. Global scenarios are quantified in terms of global changes in land use over the whole studied area (e.g., how many hectares are transformed from agricultural to urban uses in a given number of years, how does this evolve over time...); this is done either from academic expert knowledge, or from information provided by local planning agencies. Whenever feasible, models are validated by comparing the model predictions with actual evolution at a later date. Therefore, such models need from one to three land use maps at different dates for calibration and validation purposes (the larger the number of maps, the more robust and accurate the model). A large array of statistical tools is available in the literature to perform the calibration and validation of the model.

The horizon of projections of such models is limited in time, typically 20-30 years, due to the inherent uncertainty in such models, although they are occasionally used on longer time-scales. Climate change constraints are included, when needed, through scenarios, as it is not in the scope of such models to incorporate ecological processes that may translate climate change constraints into land cover change dynamics. Note that on such short time-scales, climate change is not dominated by the mean climate evolution but by decade variations which average out on longer time-scales and are not modeled in the global climate models used e.g. for IPCC projections for the end of the century; as a consequence, the various IPCC climate scenarios cannot be distinguished on such a short time horizon.

With regard to LUCC, the STEEP team has been involved for five years in the ESNET project whose funding came to a close in July of 2017, but the scientific production of the project is still underway. This project bears on the characterization of local Ecosystem Services networks; the project has been coordinated by LECA (Laboratoire d'Ecologie Alpine), in collaboration with a number of other research laboratories (most notably, IRSTEA Grenoble, besides our team), and in close interaction with a panel of local stakeholders; the scale of interest is typically a landscape (in the ecologic/geographic sense, i.e., a zone a few kilometers to a few tens of kilometers wide). The project aims at developing a generic modelling framework of ecosystem services, and studying their behavior under various scenarios of coupled urban/environment evolution, at the 2030/2040 horizon, under constraints of climate change. The contribution of the STEEP team is centered on the Land Use/Land Cover Change (LUCC) model that is one of the major building blocks of the whole project modelling effort, with the help of an ESNET funded post-doctoral researcher. In the process, areas of conceptual and methodological improvements of statistical LUCC models have been identified; implementing these improvements will be useful for the LUCC community at large, independently of the ESNET project needs.

4.3.2. Models for Land-Use and Transportation Interactions (LUTI)

Urban transport systems are intricately linked to urban structure and activities, i.e., to land use. Urbanization generally implies an increased travel demand. Cities have traditionally met this additional demand by extending transportation supply, through new highways and transit lines. In turn, an improvement of the accessibility of ever-farther land leads to an expansion of urban development, resulting in a significant feedback loop between transportation infrastructure and land use, one of the main causes of urban sprawl. Transportation models allow us to address questions generally limited to the impacts of new infrastructures, tolls and other legislation on traffic regulation [0], on user behavior [0], or on the environment [0]. LUTI models (Land-Use and Transport Integrated models) can answer a much broader spectrum of issues. For example, they allow us to understand how the localization of households and of economic activities (which generate transportation demand) adapt to changes of transportation supply. They also allow us to assess the impacts of such changes on the increase in real estate value, or more generally on their effects on the economic development of a specific sector or neighborhood. An economic vision interprets all these interactions in terms of equilibrium between demand and supply. Modelling the localization of households and employments (companies) relies on capturing the way stakeholders arbitrate between accessibility, real estate prices, and attractiveness of different areas.

State of the art and operability of LUTI models. The first model that proved able to analyze the interactions between transport and urbanization was developed by Lowry. Since then theories and models have become increasingly complex over time. They can be classified according to different criteria. A first classification retraces the historic path of these theories and models. They can be associated with one or several of the approaches underlying all present theories: economic base theory and gravity models, Input/Output models and theory of urban rent, and micro-simulations. A second possibility consists in classifying the models according to their aims and means. Significant scientific progress has been made over the last thirty years. Nevertheless, modelling tools remain largely restricted to the academic world. Today, only seven models have at least had one recent application outside academia or are commercialized or potentially marketable, in spite of the important needs expressed by the urban planning agencies: Cube Land, DELTA, MARS, OPUS/UrbanSim, PECAS, TRANUS and Pirandello.

To guide their choice of a modelling framework, users can rely on various criteria such as the strength of the theoretical framework, the quality and the diversity of the available documentation, the accessibility of the models (is the model freely available? is the code open source? is the software regularly updated and compatible with the recent operating systems?), the functionality and friendliness of user interfaces (existence of graphic user interface, possibility of interfacing with Geographic Information Systems), existence of technical assistance, volume and availability of the data required to implement the model, etc. For example, among the seven models mentioned above, only two are open source and mature enough to meet professional standards: TRANUS and UrbanSim [0]. These two models are very different but particularly representative of the main current philosophies and trends in this scientific domain. Their comparison is informative.

STEEP implication in LUTI modelling. As yet, very few local planning authorities make use of these strategic models, mostly because they are difficult to calibrate and validate. Systematic improvement on these two critical steps would clearly increase the level of confidence in their results; these limitations hinder their

[0] Congestion, cost and time spent for the transport, etc.
[0] Changes in modality choice.
[0] CO2 emissions, air pollution, noise nuisance, etc.
[0] http://www.urbansim.org

dissemination in local agencies. One of the major goals of STEEP is therefore to meet the need for better calibration and validation strategies and algorithms. This research agenda lies at the core of our project CITiES (ANR Modèles Numériques) that ended in 2017 with the PhD defense of Thomas Capelle . This work is being partly pursued in the QAMECS project.

As for LUTI modeling, we have been using the TRANUS model since the creation of our team. In this framework we work in close collaboration with AURG [0], the local urban planning agency of Grenoble (*Agence d'Urbanisme de la Région Grenobloise*) in order to better understand and to improve the relevance of these tools for such territorial agencies.

5. New Software and Platforms

5.1. 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 [15].

NEWS OF THE YEAR: In 2019 we started a refactoring of this software, together with an overhaul of the associated web service USAT-WEB.

- Participants: Luciano Gervasoni, Serge Fenet, Peter Sturm and Roger Pissard-Gibollet
- Partner: LIRIS
- Contact: Peter Sturm
- URL: https://github.com/lgervasoni/urbansprawl

5.2. 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 [21].

NEWS OF THE YEAR: In 2019 we started a complete overhaul of this software, together with a factorization of the underlying computation core (USAT).

- Participants: Lucas Rezakhanlou, Peter Sturm, Luciano Gervasoni, Serge Fenet and Roger Pissard-Gibollet
- 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

[0] http://www.aurg.org/

5.3. LUM_OSM

Land Use Mix calculation from OpenStreepMap data

KEYWORD: Urban sprawl

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, Serge Fenet and Roger Pissard-Gibollet
- Partners: EPFL - Ecole Polytechnique Fédérale de Lausanne - LIRIS
- Contact: Peter Sturm
- URL: http://github.com/martibosch/landusemix

5.4. 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

5.5. REDEM

REDuction Of EMission

KEYWORD: Climate change

FUNCTIONAL DESCRIPTION: REDEM soft is a tool designed for the benchmarking of national GHG emission reduction trajectories. The actual version of the software is implemented in Visual Basic under Microsoft Excel in order to facilitate handling and diffusion to climate/energy economists.

NEWS OF THE YEAR: In 2019, the database underlying REDEM was updated with recent national emission trajectory information.

- Participants: Constantin Ilasca, Hélène Benveniste, Olivier Boucher, Patrick Criqui and Roger Pissard-Gibollet
- Partners: EDDEN - IPSL
- Contact: Emmanuel Prados
- URL: http://redem.gforge.inria.fr/

5.6. REDEM web

REDEM Web

KEYWORDS: Benchmarking - Climate change - Global warming - Greenhouse gas emissions

FUNCTIONAL DESCRIPTION: Python Library of REDEM model and its web application.

- Participants: Constantin Ilasca, Emmanuel Prados, Hélène Benveniste, Nicolas Assouad, Olivier Boucher, Patrick Criqui and Roger Pissard-Gibollet
- Partners: UPMC - EDDEN
- Contact: Emmanuel Prados
- URL: http://redem.inria.fr/

6. New Results

6.1. Analysis of socio-ecological dimensions of human activities – A case study of Beaufort cheese production in the Maurienne Valley

The PhD thesis of Michela Bevione aims at analysing socio-ecological dimensions of human activities creating wealth by coupling quantitative-biophysical approaches and qualitative and socio-economic methodologies to assess territorial metabolism. By focusing on the interactions between flows and actors, the methodology we propose aims at providing a methodological framework for the understanding of a territory and its capability.

As a case study for this thesis, we chose to focus on the production of the AOC-labelled cheese Beaufort in the Maurienne Valley (Savoie department, Auvergne-Rhône-Alpes region, France). Indeed, agriculture play a structuring role for the economic and social dynamics of the valley, and the landscape construction induced by farming activities contributes to create favourable conditions to the development of the touristic sector. Beaufort represents the flagship product of the agricultural sector in the valley and most of farms are dedicated to milk production for the Beaufort industry.

In [7] we represent the circulation of material flows through flow maps, showing the movement of material and monetary resources and products, their direction, source and destination. We focus on the circulation of flows related to the Beaufort industry within the Maurienne Valley and between the valley and other territories. Through Sankey diagrams (a specific kind of flow maps, where the width of the arrows is proportional to the flow quantity) we present the dominant contributions to the overall material flows circulation. This kind of representation is appropriate to characterise the circulation of material flows, the allocation of environmental pressures throughout the Beaufort industry, as well as the monetary dimension and the added value associated to Beaufort production. Mapping the geographical origin of input resources and the destination of output products and incomes allows to evaluate actors' capacity to create wealth through the activation and mobilisation of local resources and/or their dependence on foreign inputs.

Furthermore, results include schematic representations of the relations between local, extra-territorial actors and the circulation of material, environmental and monetary flows. The influence of immaterial resources (informational flows and traditional savoir-faire) and local infrastructures on the circulation of flows, and vice versa, is illustrated. Finally, positive and negative retroactions induced by output products on input resources for Beaufort production are drawn, as well as the interactions with other sub-systems creating wealth in the valley.

6.2. Sensitivity analysis of World3

World3 is a computer tool created to simulate the interactions between the world population, industrial growth and food production within the limits of the planet. It aims to highlight the problems posed by indefinite material growth in a finite world. The first version of this tool was proposed in 1972 by MIT researchers for the first report to the Club of Rome [19]. This report was both highly successful and polemic. The main detractors of the model criticized it for being too approximate in the choice of the parameters and for beeing too simplistic. We started to work on revisiting some aspects of the scientific validation of this model, in a context where growth is widely debated in the scientific and civil community, through a new and more sophisticated sensitivity analysis of the model, compared to what is available in the literature [11].

6.3. Efficient computation of solution space and conflicts detection for linear systems

Our work on Material Flow Analysis (see e.g. the *AF Filières* project), involves the analysis of systems of linear inequalities, $l \leq Ax \leq u$. There are three different but complementary goals for the analysis: (i) given some known variables x_i, efficiently compute the solution space of unknown variables, (ii) if the set of constraints is infeasible, efficiently identify the conflicts, (iii) efficiently classify variables to determine whether they are redundant, just measured, determinable or non-determinable. A baseline implementation for these tasks was available in the team but proved to be too inefficient for larger problem sizes. Through the internship of Alexandre Borthomieu we worked on various improvements, on the algorithmic and implementation side (e.g. choice of programming language), that eventually led to a reduction of execution by three orders of magnitude, compared to the previous implementation.

6.4. Mapping ecosystem services bundles in a heterogeneous mountain region

2019 was the final year of production of the ESNET project (which officially ended in 2017). This and the following section describe our two most complex pieces of work in that project.

Recent institutional and policy frameworks prescribe the incorporation of ecosystem services (ES) into land use management and planning, favouring co-production of ES assessments by stakeholders, land planners and scientists. Incorporating ES into land management and planning requires models to map and analyze ES. Also, because ES do not vary independently, many operational issues ultimately relate to the mitigation of ES trade-offs, so that multiple ES and their interactions need to be considered. Using a highly accurate LULC (Land Use Land Cover) database for the Grenoble urban region (French Alps), we mapped twelve ES using a range of models of varied complexity [5]. A specific, fine-grained (less than 1 ha) LULC database at regional scale (4450 km^2) added great spatial precision in individual ES models, in spite of limits of the typological resolution for forests and semi-natural areas. We analysed ES bundles within three different socio-ecosystems and associated landscape types (periurban, rural and forest areas). Such type-specific bundles highlighted distinctive ES trade-offs and synergies for each landscape. Advanced approaches combining remote sensing, targeted field data collection and expert knowledge from scientists and stakeholders are expected to provide the significant progress that is now required to support the reduction of trade-offs and enhance synergies between management objectives.

6.5. Co-constructing future land-use scenarios for the Grenoble region, France

Physically and socially heterogeneous mountain landscapes support high biodiversity and multiple ecosystem services. But rapid landscape transformation from fast urbanisation and agricultural intensification around cities to abandonment and depopulation in higher and more remote districts, raises urgent environmental and planning issues. For anticipating their future in a highly uncertain socio-economic context, we engaged stakeholders of a dynamic urban region of the French Alps in an exemplary interactive participatory scenario planning (PSP) for co-creating salient, credible and legitimate scenarios. Stakeholders helped researchers adapt, downscale and spatialize four normative visions from the regional government, co-producing four storylines of trend versus break-away futures. Stakeholder input, combined with planning documents and analyses of recent dynamics, enabled parameterisation of high-resolution models of urban expansion, agriculture and forest dynamics. With similar storylines in spite of stakeholders insisting on different governance arrangements, both trend scenarios met current local and European planning objectives of containing urban expansion and limiting loss and fragmentation of agricultural land. Both break-away scenarios induced considerable conversion from agriculture to forest, but with highly distinctive patterns. Under a commonly investigated, deregulated liberal economic context, encroachment was random and patchy across valleys and mountains. A novel reinforced nature protection scenario affecting primarily mountain and hilly areas fostered deliberate consolidation of forested areas and connectivity. This transdisciplinary approach demonstrated the potential of combining downscaled normative scenarios with local, spatially-precise dynamics informed by stakeholders for local appropriation of topdown visions, and for supporting land planning and subsequent assessment of ecosystem service trade-offs. This work is described in [4].

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

Contract with **ADEME** (French Environment and Energy Management Agency [0]), within a collaboration with FCBA [0], Arvalis [0], Terres Univia [0], and Terres Inovia [0]. Design and development of an interactive spreadsheet application for scenarizing non-food biomass flows in France, from production to consumption (energy and non-energy uses). Visualization in the form of Sankey diagrams.

Contract with **Aura-EE** (Energy and Environment Agency of the Auvergne–Rhône-Alpes Region [0]), within the Interreg Alpine Region program. Estimation of material flows within the wood supply chain in the Alps European Region.

Contract with **Aura-EE** within the European project IMEAS. Estimation of wood flows between the Vercors Regional Natural Parc and the Grenoble metropolitan area.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. 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.

8.2. National Initiatives

8.2.1. AF Filières : Analyse des Flux des Filières biomasse pour des stratégies régionales de bioéconomie

Project funded by ADEME
Duration: 2017-2019

[0] https://www.ademe.fr/en
[0] https://www.fcba.fr
[0] https://www.arvalisinstitutduvegetal.fr/gis-@/view-607-arvstatiques.html
[0] http://www.terresunivia.fr
[0] https://www.terresinovia.fr
[0] https://en.auvergnerhonealpes-ee.fr

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.

8.3. International Initiatives

8.3.1. Inria International Partners

8.3.1.1. Informal International Partners

University of Lausanne (UNIL), Department of Ecology and Evolution (Jérôme Gippet): development of the MoRIS model of propagation of invasive species.

8.3.2. Participation in Other International Programs

Pierre-Yves Longaretti is involved in TARA (Transition adaptation research alliance); he animated the theme *Operationalizing reflexive sustainability* at the TARA Workshop in Bogor, Indonesia, November 2019.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Selection

Jean-Yves Courtonne co-organized a special session entitled *Advances in understanding the physical structures of economies: Materials, energy, and the services they provide* at the 13th International Conference of the European Society for Ecological Economics, Turku, Finland, June 2019. https://esee2019turku.fi/specialsessions

9.1.1.1. Member of Conference Program Committees

- Peter Sturm, ORASIS 2019 (Journées francophones des jeunes chercheurs en vision par ordinateur), Saint-Dié-des-Vosges, France, https://orasis2019.sciencesconf.org
- Peter Sturm: Inria Science Days (Journées Scientifiques Inria), Lyon, France.

9.1.2. Invited Talks

- Jean-Yves Courtonne and Guillaume Mandil: *Durabilité forte : état des lieux et programme de recherche pour l'évaluation d'alternatives socio-techniques*, Colloque Soutenabilité des systèmes socio-techniques, Université Technologique de Troyes, France.
- Pierre-Yves Longaretti, *Global environmental collapse risks: is digital technology a problem or a solution?*, Keynote speech at LIG, Grenoble, France, June 2019.

9.1.3. Scientific Expertise

Jean-Yves Courtonne and Guillaume Mandil participate in the Steering Committe (COPIL) of the CODEC (Contrat d'Objectifs Déchets et Économie Circulaire) of Grenoble Alpes Métropole (Grenoble Metropolitan Area, https://www.lametro.fr/).

Team members intervened in training programs dedicated to IT professionals and teachers organized by EcoInfo [0], a French network dedicated to eco-responsible digital sciences and technologies:

- Guillaume Mandil provided a training session on Life Cycle Analysis for EcoInfo members, Grenoble, May 2019.
- Pierre-Yves Longaretti, Guillaume Mandil, and Peter Sturm gave lectures/training sessions at the week-long training seminar *Impact Environnemental du Numérique – Comprendre et Agir*, Autrans, France, September 2019.

9.1.4. Research Administration

- Emmanuel Prados and Guillaume Mandil are members of the Grenoble section of the *Labos 1.5* initiative: an international, cross-disciplinary collective of academic researchers who share a common goal: to better understand and reduce the environmental impact of research, especially on the Earth's climate, https://labos1point5.org/en/home/
- Emmanuel Prados is member of the working group on sustainable development of the LJK lab, whose goal is to define and environmental charter and an environmental scientific policy for the lab.
- Peter Sturm is Deputy Scientific Director of Inria, in charge of the domain Perception–Cognition–Interaction.
- Peter Sturm co-founded and co-animated (with Céline Serrano) a national working group at Inria whose goal was to provide suggestions for the institute concerning environmental issues, in particular: (i) research to address the increasing environmental challenges (including directions of research and administrative aspects), (ii) how to reduce the ecological impact of the institute, and, more generally, (iii) ways of fostering responsible science ("sciences impliquées"). The working group produced a report [9] and several of its recommendations have already been taken up by Inria.
- Peter Sturm is member (representative of personnel) of the CLHSCT of Inria Grenoble Rhône-Alpes (Local Committee for Hygiene, Security and Working Conditions).
- Peter Sturm is member of the CLDD of Inria Grenoble Rhône-Alpes (Local Sustainable Development Commission).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Jean-Yves Courtonne, Serge Fenet, Fausto Lo Feudo, Pierre-Yves Longaretti, Guillaume Mandil, Régis Perrier, Emmanuel Prados, Peter Sturm, and Françoise Berthoud (CNRS): *Les véritables enjeux environnementaux – compréhension, modélisations et outils quantitatifs*, 18 hours, graduate course, Université Grenoble Alpes, https://team.inria.fr/steep/seminars/cours-ed-2019/

Jean-Yves Courtonne: *Méthodes d'évaluation environnementale*, 1h30, Ensimag (Grenoble INP)

Jean-Yves Courtonne: *Méthodes d'évaluation environnementale*, 2h, Programme IDEX Green University, Université Grenoble Alpes

Pierre-Yves Longaretti: *Global changes, planetary limits and collapse risks*, 3h, Evry ENSIIE engineering school, May 2019.

Pierre-Yves Longaretti: *Global changes, planetary limits and collapse risks*, 3h, Grenoble ENSIMAG engineering school, October 2019.

Serge Fenet, Guillaume Mandil and Régis Perrier have regular teaching duties at the universities employing them.

9.2.2. Supervision

[0] https://ecoinfo.cnrs.fr

PhD in progress: Michela Bevione, *Enjeux socio-écologiques, métabolisme territorial, création de richesse : application à la vallée de la Maurienne*, started in October 2016, supervised by Pierre-Yves Longaretti and Nicolas Buclet (PACTE laboratory)

PhD in progress: François-Rémi Mazy, *Vers une théorie et une implémentation algorithmique cohérentes des modèles statistiques de changement d'usage des sols*, started in October 2019, supervised by Pierre-Yves Longaretti

PhD in progress: Mathilde Jochaud du Plessix, *Robustesse et validité des modèles dynamiques de risques systémiques globaux*, supervised by Serge Fenet and Pierre-Yves Longaretti

9.2.3. Juries

- Pierre-Yves Longaretti was external expert of Grégoire Chambaz masters' thesis (UNIL, Lausanne): *The Evolutionary Dynamics of Societies Critical Synthesis of the Work of J.A. Tainter* et al. *on Sustainability, Collapse, Resilience and Energy Transitions.*
- Peter Sturm was reviewer of the Habilitation thesis of Gilles Simon (Université de Lorraine) and of the PhD thesis of Yilin Zhou (Université Paris-Est) and chaired the PhD thesis committee of Julien Salotti (Université de Lyon).

9.3. Popularization

9.3.1. Education

- Guillaume Mandil gave a half-day training course to a group of high school teachers (Lycée Mounier - CLEPT), July 2019.
- Peter Sturm gave a lecture on the environmental impact of ICT, during a training week of high school teachers of the Hauts-de-France Region, Lille, France, June 2019.

9.3.2. Conference-debate series and YouTube-channel "Understanding and Acting"

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 2019 the YouTube channel has about 5,400 subscribers and reached a total of about 425,000 viewings.

The conference-debates of 2019:

- Olivier Vidal, Université Grenoble Alpes: Resources and energy at the global scale in the context of the energy transition (*Matières premières et énergie à l'échelle mondiale dans le contexte de la transition énergétique*)
- Grégoire Chambaz, Lausanne: How real is the risk of a blackout? (*Le risque de blackout est-il réel ?*)
- Christophe Bonneuil, CNRS and École des hautes études en sciences sociales: When the whites intended to preserve the planet – A history of geo-power, 1865-1914 (*Quand les blancs voulurent conserver la planète. Une histoire du géopouvoir, 1865-1914*)
- Peter Sturm, STEEP: Environmental challenges and what is blocking acting on them (*Enjeux environnementaux et blocages à l'action*)
- Sophie Wahnich, CNRS and Institut interdisciplinaire d'anthropologie du contemporain: Why was King Louis XVI destituted in 1792? Social imaginaries from 1770 to 1792 (*Pourquoi le roi Louis XVI a t-il été destitué en 1792 ? Imaginaires sociaux de 1770 à 1792*)
- Denis Dupré, STEEP: Collapses and finance (part I) – Is Finance a collapse pyromaniac? (*Effondrements et finance (partie I) – La Finance : pyromane des effondrements ?*)

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

9.3.3. Interventions

Guillaume Mandil participated in the preparation of an exposition at Cité des Sciences (Paris).

Team members gave the following general audience talks:

- Serge Fenet: Talk *Design et anthropocène : business as unusual – Une cartographie des risques par le prisme de la systémique* at the *Biennale Internationale du Design*, St-Étienne, France, May 2019.
- Serge Fenet: *Effondrements et anthropocène à la lumière de la dynamique des systèmes* at ESC Clermont-Ferrand, France, May 2019.
- Serge Fenet: *Technologies numériques et catastrophe environnementale : des liens complexes* at École Centrale de Lyon, for a general audience of master students.
- Pierre-Yves Longaretti: Seminar *Voluntary minimal consumption: stakes, obstacles, levers*. EcoInfo thematic school, November 2019.
- Pierre-Yves Longaretti: Seminar *Climate Change and strategic surprise*, Swiss military history and prospective center, December 2019.
- Pierre-Yves Longaretti: Round table discussion, *Will we all be farmers tomorrow? (Demain, tous paysans ?)*, Transition Towns biennial meeting, Grenoble, March 2019.
- Pierre-Yves Longaretti: Round table discussion, *Agriculture: global and local stakes (Agriculture : enjeux locaux et globaux)*, Forum citoyen alimentation & santé : La transition agro-alimentaire dans nos assiettes, Grenoble, November 2019.
- Emmanuel Prados: Institutional Conference *Veille techno à Polytech*, Grenoble, France, December 9-16, 2019. Organized by the MSTII Graduate School of Université Grenoble Alpes (Bernard Tourancheau).
- Emmanuel Prados: Institutional Conference organized by UniLaSalle (Hervé Leyrit), Beauvais, France, October 17, 2019.
- Emmanuel Prados: Conference at the back-to-school day of EPITA LYON (computer intelligence engineering school), Lyon, France, September 10, 2019. Organizer: Epita (Lamia Derrode).

- Emmanuel Prados: Conference on environmental transition for elected politicians of the Grésivaudan community of communes, open to the general public, Le Versoud, France, June 3rd, 2019. Organizer: Le Versoud City Hall (Claire Desmaris).
- Emmanuel Prados: General public conference, Crolles, France, June 22 , 2019. Organizer: Association *Ecocitoyens*.
- Emmanuel Prados: Conference at the *CitizenCampus* entitled *Transition, démocratie, effondrement : la quadrature du cercle ?*, Grenoble, France, March 29, 2019. Organizer: Université Grenoble Alpes. https://twitter.com/YannToma/status/1112094922499330049
- Emmanuel Prados: Conference/debate *Quelles transitions dans un contexte d'effondrement ?*, Palais des Sports de Grenoble, France, March 13, 2019. Organizer: Ville de Grenoble. http://villesentransition.grenoble.fr/events/quelles-transitions-dans-un-contexte-deffondrement/
- Emmanuel Prados: Conference within the ISN cycle, *Le numérique face aux enjeux environnementaux et sociétaux*, Grenoble, France, April 10, 2019. Organizer: Inria. https://www.canal-u.tv/video/inria/le_numerique_face_aux_enjeux_environnementaux_et_societaux.50079
- Emmanuel Prados: Invited conference at the Loria colloquium, *Le numérique face à la catastrophe environnementale et sociétale*, Nancy, France, February 8, 2019. Organizer: Loria (http://www.loria.fr/). https://www.loria.fr/fr/communication/colloquium-loria-exposes-et-videos/
- Peter Sturm made a "Tour de France" by bicycle, to raise awareness on environmental issues. During this trip (May 15 to July 4), he visited all Inria centers and gave a talk on *Environmental challenges and what is blocking acting on them* in ten locations (Sophia Antipolis, Montpellier, Bordeaux, Rennes, Saclay, Rocquencourt, Paris, Nancy, Lyon, Grenoble). His visits were accompanied by various events, including seminars by local researchers, meetings with sustainable development commissions, social gatherings, etc. This initiative is described at https://project.inria.fr/inriavelo/fr/, see also the daily blog https://www.polarsteps.com/PeterSturm3/1692663-inri-a-velo.

10. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] T. Birdal, B. Busam, N. Nassir, S. Ilic, P. Sturm.*Generic Primitive Detection in Point Clouds Using Novel Minimal Quadric Fits*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", 2019, p. 1-14, https://arxiv.org/abs/1901.01255v1 , forthcoming [*DOI :* 10.1109/TPAMI.2019.2900309], https://hal.inria.fr/hal-02368932

[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", September 2019, vol. 77, p. 101146:1-13 [*DOI :* 10.1016/j.compenvurbsys.2017.04.009], https://hal.inria.fr/hal-01519654

[3] R. G. French, C. A. McGhee-French, P. Nicholson, M. M. Hedman, N. J. Rappaport, E. A. Marouf, P.-Y. Longaretti, J. Hahn.*Unusual One-Armed Density Waves in the Cassini Division of Saturn's Rings*, in "Icarus", 2019, p. 1-59, forthcoming, https://hal.archives-ouvertes.fr/hal-02404628

[4] C. Vannier, A. Bierry, P.-Y. Longaretti, B. Nettier, T. Cordonnier, C. Chauvin, N. Bertrand, F. Quetier, R. Lasseur, S. Lavorel.*Co-constructing future land-use scenarios for the Grenoble region, France*, in "Landscape and Urban Planning", 2019, vol. 190, 103614 [*DOI :* 10.1016/j.landurbplan.2019.103614], https://hal.archives-ouvertes.fr/hal-02405231

[5] C. VANNIER, R. LASSEUR, E. CROUZAT, C. BYCZEK, V. LAFOND, T. CORDONNIER, P.-Y. LONGARETTI, S. LAVOREL.*Mapping ecosystem services bundles in a heterogeneous mountain region*, in "Ecosystems and People", February 2019, vol. 15, n^{o} 1, p. 74-88 [*DOI :* 10.1080/26395916.2019.1570971], https://hal.archives-ouvertes.fr/hal-02404663

International Conferences with Proceedings

[6] S. PENG, P. STURM.*Calibration Wizard: A Guidance System for Camera Calibration Based on Modelling Geometric and Corner Uncertainty*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, 2019, p. 1497-1505, https://arxiv.org/abs/1811.03264 , https://hal.inria.fr/hal-02368999

Conferences without Proceedings

[7] M. BEVIONE, N. BUCLET, J.-Y. COURTONNE, P.-Y. LONGARETTI.*Socio-ecological transition, wealth creation and territorial metabolism: the case of the production of the AOC-labelled cheese Beaufort in the Maurienne Valley*, in "ESEE 2019 - 13th International Conference of the European Society for Ecological Economics", Turku, Finland, June 2019, p. 1-4, https://hal.inria.fr/hal-02430879

[8] J.-Y. COURTONNE, J. ALAPETITE, V. WAWRZYIAK, M. BEVIONE.*The AF Filières project: application of PSUT frameworks for regional analyses of agriculture and forestry supply chains and footprints in France*, in "ESEE 2019 - 13th International Conference of the European Society for Ecological Economics", Turku, Finland, June 2019, p. 1-4, https://hal.inria.fr/hal-02430898

Research Reports

[9] F. BERTHOUD, P. GUITTON, L. LEFÈVRE, S. QUINTON, A. ROUSSEAU, J. SAINTE-MARIE, C. SERRANO, J.-B. STEFANI, P. STURM, E. TANNIER.*Sciences, Environnements et Sociétés : Rapport long du groupe de travail MakeSEnS d'Inria*, Inria, October 2019, https://hal.inria.fr/hal-02340948

Other Publications

[10] A. BORTHOMIEU.*Efficient computation of solution space and conflicts detection for linear systems*, Université Grenoble Alpes, June 2019, Mémoire M1 sous la forme d'article, https://hal.inria.fr/hal-02433237

[11] M. JOCHAUD DU PLESSIX.*Analyse du modèle World3 : sensibilité, dynamique, et pistes d'évolution*, Insa Lyon, September 2019, https://hal.inria.fr/hal-02434683

[12] F.-R. MAZY.*Conceptually correct and algorithmically efficient pruning and allocation procedures in statistical LUCC modelling*, Université Grenoble Alpes, September 2019, https://hal.inria.fr/hal-02302133

References in notes

[13] U. BARDI.*The Limits to Growth Revisited*, Springer Verlag, 2011

[14] E. BOVARI, G. GIRAUD, F. MC ISAAC.*Coping With Collapse: A Stock-Flow Consistent Monetary Macrodynamics of Global Warming*, in "Ecological Economics", 2018, vol. 147, p. 383-398

[15] 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

[16] I. GOLDIN, M. MARIATHASAN.*The Butterfly Defect: How Globalization Creates Systemic Risks, and What to Do about It*, Princeton University Press, 2014

[17] D. KOROWICZ.*Trade-off: Financial System Supply-Chain Cross Contagion — a Study in Global Systemic Collapse*, 2012, http://www.feasta.org/wp-content/uploads/2012/10/Trade_Off_Korowicz.pdf

[18] D. MEADOWS, W. BEHRENS III, D. MEADOWS, R. NAILL, J. RANDERS, E. ZAHN.*Dynamics of Growth in a Finite World*, Wright-Allen Press, 1974

[19] D. MEADOWS, D. MEADOWS, J. RANDERS, W. W. BEHRENS.*The Limits to Growth*, Universe Books, 1972

[20] D. MEADOWS, J. RANDERS, D. MEADOWS.*The Limits to growth, The 30-Year Update*, Chelsea Green Publishing, 2004

[21] 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

[22] G. TURNER.*A Comparison of The Limits to Growth with 30 Years of Reality*, in "Global Env. Change", 2008, vol. 18, p. 397-411

[23] G. TURNER.*Is Global Collapse Imminent?, MSSI Research Paper No. 4. Melbourne Sustainable Society Institute, University of Melbourne*, Melbourne Sustainable Society Institute, University of Melbourne, 2014

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]
Cordelia Schmid [Inria, Senior Researcher, HDR]
Karteek Alahari [Inria, Researcher, HDR]
Grégory Rogez [Inria, Starting Research Position, until Jan 2019]
Jakob Verbeek [Inria, Senior Researcher, HDR]

Faculty Member

Jocelyn Chanussot [Institut polytechnique de Grenoble, Professor, from Sep 2019, HDR]

External Collaborator

Ghislain Durif [CNRS, from Feb 2019]

Technical Staff

Ghislain Durif [Inria, Engineer, until Jan 2019]
Ricardo Jose Garcia Pinel [Inria, Engineer, from Jul 2019]
François Gindraud [Inria, Engineer, until Mar 2019]
Igor Kalevatykh [Inria, Engineer, until Nov 2019]
Bruno Lecouat [Inria, Engineer, until Aug 2019]
Xavier Martin [Inria, Engineer]
Alexandre Zouaoui [Inria, Engineer, from Dec 2019]

PhD Students

Minttu Alakuijala [Inria, PhD Student]
Florent Bartoccioni [Inria, PhD Student, from Nov 2019]

Alberto Bietti [Inria, PhD Student]
Mathilde Caron [Facebook, PhD Student, granted by CIFRE]
Dexiong Chen [Univ. Grenoble Alpes, PhD Student]
Mikita Dvornik [Inria, PhD Student]
Maha Elbayad [Univ Grenoble Alpes, PhD Student]
Valentin Gabeur [Inria, PhD Student]
Pierre Louis Guhur [Univ Paris-Saclay, PhD Student, from Sep 2019]
Yana Hasson [Inria, PhD Student]
Ekaterina Iakovleva [Univ Grenoble Alpes, PhD Student]
Roman Klokov [Inria, PhD Student]
Andrei Kulunchakov [Inria, PhD Student]
Bruno Lecouat [Inria, PhD Student, from Sep 2019]
Hubert Leterme [Univ Grenoble Alpes, PhD Student, from Sep 2019]
Pauline Luc [Facebook CIFRE, PhD Student, until May 2019]
Thomas Lucas [Univ Grenoble Alpes, PhD Student]
Lina Mezghani [Inria, PhD Student, from Nov 2019]
Gregoire Mialon [Inria, PhD Student]
Alexander Pashevich [Inria, PhD Student]
Alexandre Sablayrolles [Facebook CIFRE, PhD Student]
Konstantin Shmelkov [Inria, PhD Student, until Mar 2019]
Robin Strudel [École Normale Supérieure de Paris, PhD Student]
Vladyslav Sydorov [Inria, PhD Student]
Gul Varol Simsekli [Inria, PhD Student, until Feb 2019]
Nitika Verma [Univ Grenoble Alpes, PhD Student, until May 2019]
Daan Wynen [Inria, PhD Student, until Sep 2019]
Houssam Zenati [Criteo, PhD Student, from Oct 2019]

Post-Doctoral Fellow

Adria Ruiz Ovejero [Inria, Post-Doctoral Fellow]

Visiting Scientists

Hernan Dario Benitez Restrepo [Pontificia Universidad Javeriana Sede Cali, from Nov 2019]
Pia Bideau [Univ. Massachusetts Amherst, until Jan 2019]
Avijit Dasgupta [IIIT Hyderabad, from Feb 2019 until May 2019]
Ning Huyan [from Dec 2019]
Dou Quan [Inria, from Dec 2019]
Gunnar Atli Sigurdsson [CMU, until Mar 2019]

Administrative Assistant

Nathalie Gillot [Inria, Administrative Assistant]

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.

[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.

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, 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

- Cordelia Schmid received the Royal Society Milner Award, 2019.
- Julien Mairal received the test-of-time award at the International Conference on Machine Learning (ICML), 2019.
- The paper [21] authored by Roman Klokov, Jakob Verbeek, Edmond Boyer [Inria Morpheo] won the “Best Science Paper Award Honourable Mention” at BMVC 2019.
- Jakob Verbeek was awarded as an outstanding reviewer at ICLR 2019.
- Adria Ruiz Ovejero was awarded as an outstanding reviewer at ICCV 2019.

5.1.2. Dissemination

- The team co-organized PAISS 2019, an international AI summer school in Paris. This is the second edition of the school that was first organized in Grenoble in 2018. The 2019 edition brought together over 200 participants. We also provided scholarships to 21 students to encourage diversity among the attendees.

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
- Partner: Naver Labs Europe
- Contact: Nicolas Jourdan
- 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: D. Chen, L. Jacob, and J. Mairal. Biological Sequence Modeling with Convolutional Kernel Networks. Bioinformatics, volume 35, issue 18, pages 3294-3302, 2019.

- 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. 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%. We also provide an extensive ablative analysis to investigate the influence of each component in the proposed framework.

- Participants: Karteek Alahari, Cordelia Schmid and Pavel Tokmakov
- Contact: Pavel Tokmakov
- Publication: hal-01511145v2
- URL: http://lear.inrialpes.fr/research/lvo/

6.4. 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, Xavier Martin, Ivan Laptev and Cordelia Schmid
- Contact: Gül Varol
- Publication: Learning from Synthetic Humans
- URL: http://www.di.ens.fr/willow/research/surreal/

6.5. 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

6.6. Cyanure

Cyanure: An Open-Source Toolbox for Empirical Risk Minimization

KEYWORD: Machine learning

FUNCTIONAL DESCRIPTION: Cyanure is an open-source C++ software package with a Python interface. The goal of Arsenic is to provide state-of-the-art solvers for learning linear models, based on stochastic variance-reduced stochastic optimization with acceleration mechanisms and Quasi-Newton principles. Arsenic can handle a large variety of loss functions (logistic, square, squared hinge, multinomial logistic) and regularization functions (l2, l1, elastic-net, fused Lasso, multi-task group Lasso). It provides a simple Python API, which is very close to that of scikit-learn, which should be extended to other languages such as R or Matlab in a near future.

RELEASE FUNCTIONAL DESCRIPTION: version initiale

- Participant: Julien Mairal
- Contact: Julien Mairal
- URL: http://thoth.inrialpes.fr/people/mairal/arsenic/welcome.html

7. New Results

7.1. Visual Recognition and Robotics

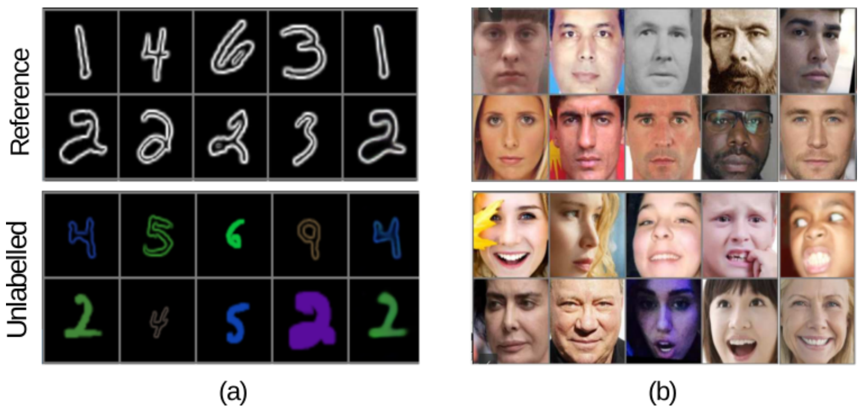


Figure 1. Illustration of different reference-based disentangling problems. (a) Disentangling style from digits. The reference distribution is composed by numbers with a fixed style (b) Disentangling factors of variations related with facial expressions. Reference images correspond to neutral faces. Note that pairing information between unlabelled and reference images is not available during training.

7.1.1. Learning Disentangled Representations with Reference-Based Variational Autoencoders

Participants: Adria Ruiz, Oriol Martinez, Xavier Binefa, 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 [32] 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. 1 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.

7.1.2. Tensor Decomposition and Non-linear Manifold Modeling for 3D Head Pose Estimation

Participants: Dmytro Derkach, Adria Ruiz, Federico M. Sukno.

Head pose estimation is a challenging computer vision problem with important applications in different scenarios such as human-computer interaction or face recognition. In [5], we present a 3D head pose estimation algorithm based on non-linear manifold learning. A key feature of the proposed approach is that it allows modeling the underlying 3D manifold that results from the combination of rotation angles. To do so, we use tensor decomposition to generate separate subspaces for each variation factor and show that each of them has a clear structure that can be modeled with cosine functions from a unique shared parameter per angle (see Fig. 2). Such representation provides a deep understanding of data behavior. We show that the proposed framework can be applied to a wide variety of input features and can be used for different purposes. Firstly, we test our system on a publicly available database, which consists of 2D images and we show that the cosine functions can be used to synthesize rotated versions from an object from which we see only a 2D image at a specific angle. Further, we perform 3D head pose estimation experiments using other two types of features: automatic landmarks and histogram-based 3D descriptors. We evaluate our approach on two publicly available databases, and demonstrate that angle estimations can be performed by optimizing the combination of these cosine functions to achieve state-of-the-art performance.

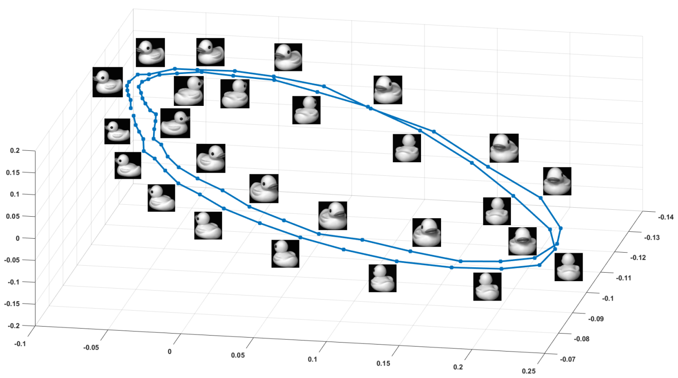


Figure 2. Visualization of the first three coefficients of the pose variation subspace for a dataset of single object rotated about the vertical axis.

7.1.3. Spreading vectors for similarity search

Participants: Alexandre Sablayrolles, Matthijs Douze, Cordelia Schmid, Hervé Jégou.

Discretizing multi-dimensional data distributions is a fundamental step of modern indexing methods. State-of-the-art techniques learn parameters of quantizers on training data for optimal performance, thus adapting quantizers to the data. In this work [29], we propose to reverse this paradigm and adapt the data to the quantizer: we train a neural net which last layer forms a fixed parameter-free quantizer, such as pre-defined points of a hyper-sphere. As a proxy objective, we design and train a neural network that favors uniformity in the spherical latent space, while preserving the neighborhood structure after the mapping. We propose a new regularizer derived from the Kozachenko–Leonenko differential entropy estimator to enforce uniformity and combine it with a locality-aware triplet loss. Experiments show that our end-to-end approach outperforms most learned quantization methods, and is competitive with the state of the art on widely adopted benchmarks. Furthermore, we show that training without the quantization step results in almost no difference in accuracy, but yields a generic catalyzer 3 that can be applied with any subsequent quantizer. The code is available online.

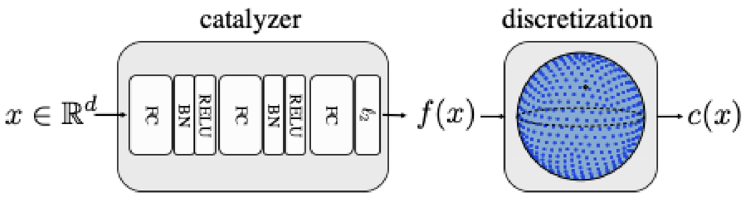


Figure 3. Our method learns a network that encodes the input space $\mathbb{R}^d$ into a code $c(x)$. It is learned end-to-end, yet the part of the network in charge of the discretization operation is fixed in advance, thereby avoiding optimization problems. The learnable function f, namely the "catalyzer", is optimized to increase the quality of the subsequent coding stage.

7.1.4. Diversity with Cooperation: Ensemble Methods for Few-Shot Classification

Participants: Nikita Dvornik, Cordelia Schmid, Julien Mairal.

Few-shot classification consists of learning a predictive model that is able to effectively adapt to a new class, given only a few annotated samples. To solve this challenging problem, meta-learning has become a popular paradigm that advocates the ability to "learn to adapt". Recent works have shown, however, that simple learning strategies without meta-learning could be competitive. In our ICCV'19 paper [17], we go a step further and show that by addressing the fundamental high-variance issue of few-shot learning classifiers, it is possible to significantly outperform current meta-learning techniques. Our approach consists of designing an ensemble of deep networks to leverage the variance of the classifiers, and introducing new strategies to encourage the networks to cooperate, while encouraging prediction diversity, as illustrated in Figure 4. Evaluation is conducted on the mini-ImageNet and CUB datasets, where we show that even a single network obtained by distillation yields state-of-the-art results.

7.1.5. Unsupervised Pre-Training of Image Features on Non-Curated Data

Participants: Mathilde Caron, Piotr Bojanowski [Facebook AI], Julien Mairal, Armand Joulin [Facebook AI].

Pre-training general-purpose visual features with convolutional neural networks without relying on annotations is a challenging and important task. Most recent efforts in unsupervised feature learning have focused on either small or highly curated datasets like ImageNet, whereas using non-curated raw datasets was found to decrease the feature quality when evaluated on a transfer task. Our goal is to bridge the performance gap between unsupervised methods trained on curated data, which are costly to obtain, and massive raw datasets that are easily available. To that effect, we propose a new unsupervised approach, DeeperCluster [13], described in Figure 5 which leverages self-supervision and clustering to capture complementary statistics from large-scale data. We validate our approach on 96 million images from YFCC100M, achieving state-of-the-art results among unsupervised methods on standard benchmarks, which confirms the potential of unsupervised learning

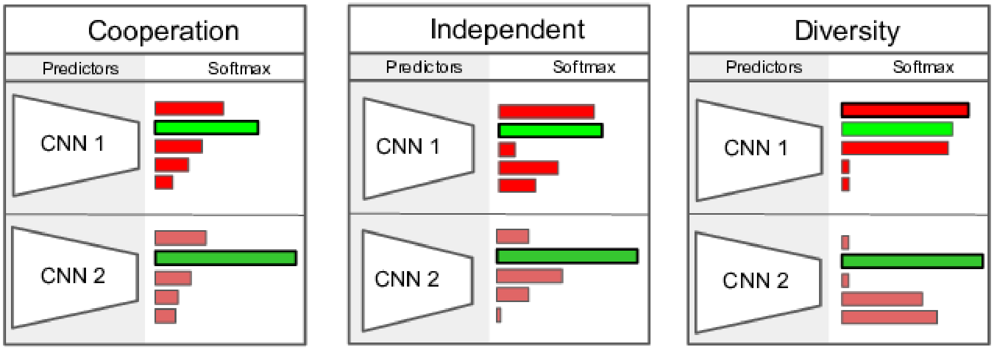


Figure 4. ***Illustration of the cooperation and diversity strategies on two networks.*** *All networks receive the same image as input and compute corresponding class probabilities with softmax. Cooperation encourages the non-ground truth probabilities (in red) to be similar, after normalization, whereas diversity encourages orthogonality.*

when only non-curated raw data are available. We also show that pre-training a supervised VGG-16 with our method achieves 74.9% top-1 classification accuracy on the validation set of ImageNet, which is an improvement of +0.8% over the same network trained from scratch.

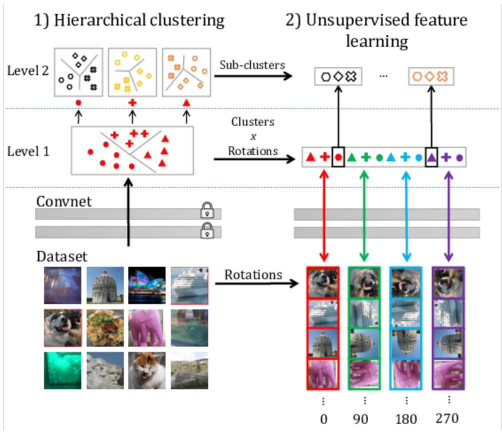


Figure 5. DeeperCluster alternates between a hierachical clustering of the features and learning the parameters of a convnet by predicting both the rotation angle and the cluster assignments in a single hierachical loss.

7.1.6. Learning to Augment Synthetic Images for Sim2Real Policy Transfer

Participants: Alexander Pashevich, Robin Strudel [Inria WILLOW], Igor Kalevatykh [Inria WILLOW], Ivan Laptev [Inria WILLOW], Cordelia Schmid.

Vision and learning have made significant progress that could improve robotics policies for complex tasks and environments. Learning deep neural networks for image understanding, however, requires large amounts of domain-specific visual data. While collecting such data from real robots is possible, such an approach limits the scalability as learning policies typically requires thousands of trials. In this work [25] we attempt to learn manipulation policies in simulated environments. Simulators enable scalability and provide access to the underlying world state during training. Policies learned in simulators, however, do not transfer well to real scenes given the domain gap between real and synthetic data. We follow recent work on domain randomization and augment synthetic images with sequences of random transformations. Our main contribution is to optimize the augmentation strategy for sim2real transfer and to enable domain-independent policy learning, as illustrated in Figure 6. We design an efficient search for depth image augmentations using object localization as a proxy task. Given the resulting sequence of random transformations, we use it to augment synthetic depth images during policy learning. Our augmentation strategy is policy-independent and enables policy learning with no real images. We demonstrate our approach to significantly improve accuracy on three manipulation tasks evaluated on a real robot.

7.1.7. Learning to combine primitive skills: A step towards versatile robotic manipulation

Participants: Robin Strudel [Inria WILLOW], Alexander Pashevich, Igor Kalevatykh [Inria WILLOW], Ivan Laptev [Inria WILLOW], Josef Sivic [Inria WILLOW], Cordelia Schmid.

Manipulation tasks such as preparing a meal or assembling furniture remain highly challenging for robotics and vision. Traditional task and motion planning (TAMP) methods can solve complex tasks but require full state observability and are not adapted to dynamic scene changes. Recent learning methods can operate directly

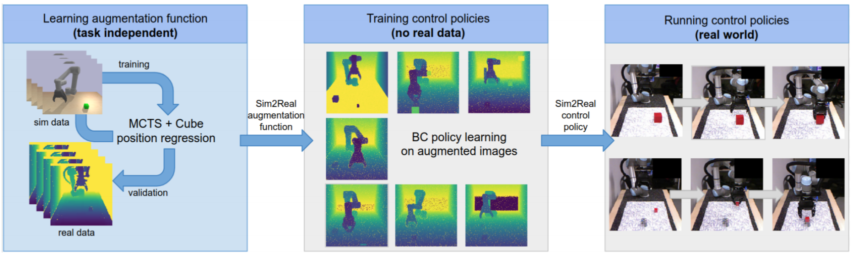


Figure 6. Overview of the method. Our contribution is the policy-independent learning of depth image augmentations (left). The resulting sequence of augmentations is applied to synthetic depth images while learning manipulation policies in a simulator (middle). The learned policies are directly applied to real robot scenes without finetuning on real images.

on visual inputs but typically require many demonstrations and/or task-specific reward engineering. In this work [40] we aim to overcome previous limitations and propose a reinforcement learning (RL) approach to task planning that learns to combine primitive skills illustrated in Figure 7. First, compared to previous learning methods, our approach requires neither intermediate rewards nor complete task demonstrations during training. Second, we demonstrate the versatility of our vision-based task planning in challenging settings with temporary occlusions and dynamic scene changes. Third, we propose an efficient training of basic skills from few synthetic demonstrations by exploring recent CNN architectures and data augmentation. Notably, while all of our policies are learned on visual inputs in simulated environments, we demonstrate the successful transfer and high success rates when applying such policies to manipulation tasks on a real UR5 robotic arm.

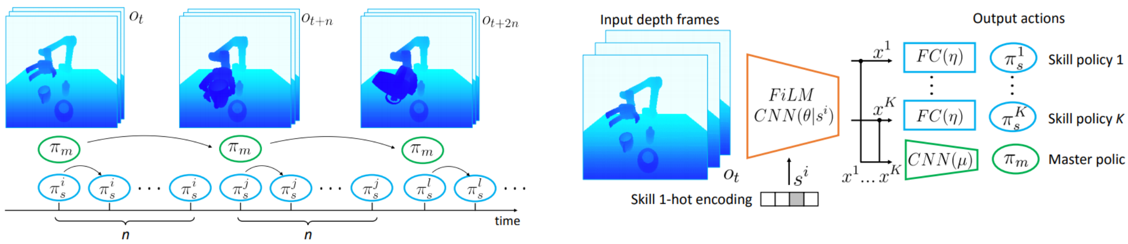


Figure 7. Illustration of our approach. (Left): Temporal hierarchy of master and skill policies. The master policy π_m is executed at a coarse interval of n time-steps to select among K skill policies $\pi_s^1 ... \pi_s^K$. Each skill policy generates control for a primitive action such as grasping or pouring. (Right): CNN architecture used for the skill and master policies.

7.1.8. Probabilistic Reconstruction Networks for 3D Shape Inference from a Single Image

Participants: Roman Klokov, Jakob Verbeek, Edmond Boyer [Inria Morpheo].

In our BMVC'19 paper [21], we study end-to-end learning strategies for 3D shape inference from images, in particular from a single image. Several approaches in this direction have been investigated that explore different shape representations and suitable learning architectures. We focus instead on the underlying probabilistic mechanisms involved and contribute a more principled probabilistic inference-based reconstruction framework, which we coin Probabilistic Reconstruction Networks. This framework expresses image conditioned

3D shape inference through a family of latent variable models, and naturally decouples the choice of shape representations from the inference itself. Moreover, it suggests different options for the image conditioning and allows training in two regimes, using either Monte Carlo or variational approximation of the marginal likelihood. Using our Probabilistic Reconstruction Networks we obtain single image 3D reconstruction results that set a new state of the art on the ShapeNet dataset in terms of the intersection over union and earth mover's distance evaluation metrics. Interestingly, we obtain these results using a basic voxel grid representation, improving over recent work based on finer point cloud or mesh based representations. In Figure 8 we show a schematic overview of our model.

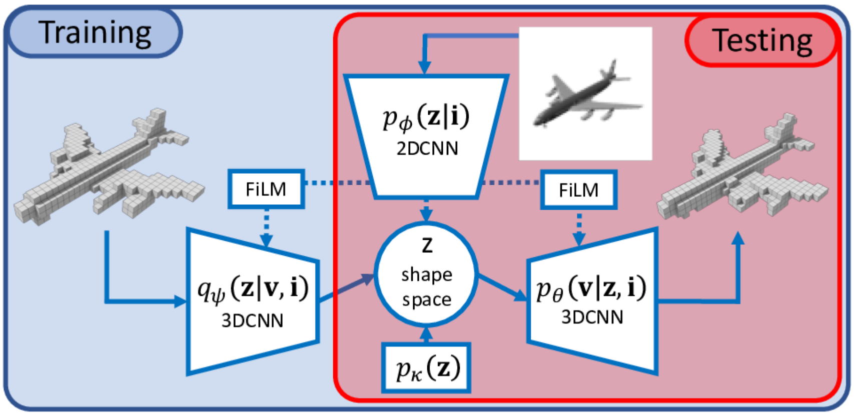


Figure 8. Probabilistic Reconstruction Networks for 3D shape inference from a single image. Arrows show the computational flow through the model, dotted arrows show optional image conditioning. Conditioning between 2D and 3D tensors is achieved by means of FiLM layers. The inference network q_ψ is only used during training for variational inference.

7.1.9. Hierarchical Scene Coordinate Classification and Regression for Visual Localization

Participants: Xiaotian Li [Aalto Univ., Finland], Shuzhe Wang [Aalto Univ., Finland], Li Zhao [Aalto Univ., Finland], Jakob Verbeek, Juho Kannala [Aalto Univ., Finland].

Visual localization is critical to many applications in computer vision and robotics. To address single-image RGB localization, state-of-the-art feature-based methods match local descriptors between a query image and a pre-built 3D model. Recently, deep neural networks have been exploited to regress the mapping between raw pixels and 3D coordinates in the scene, and thus the matching is implicitly performed by the forward pass through the network. However, in a large and ambiguous environment, learning such a regression task directly can be difficult for a single network. In our paper [37], we present a new hierarchical scene coordinate network to predict pixel scene coordinates in a coarse-to-fine manner from a single RGB image. The network consists of a series of output layers with each of them conditioned on the previous ones. The final output layer predicts the 3D coordinates and the others produce progressively finer discrete location labels. The proposed method outperforms the baseline regression-only network and allows us to train single compact models which scale robustly to large environments. It sets a new state-of-the-art for single-image RGB localization performance on the 7-Scenes, 12-Scenes, Cambridge Landmarks datasets, and three combined scenes. Moreover, for large-scale outdoor localization on the Aachen Day-Night dataset, our approach is much more accurate than existing scene coordinate regression approaches, and reduces significantly the performance gap w.r.t. explicit feature matching approaches. In Figure 9 we illustrate the scene coordinate predictions for the Aachen dataset experiments.

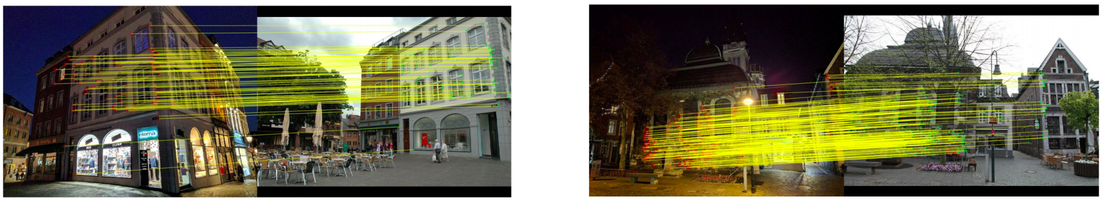


Figure 9. The scene coordinate predictions are visualized as 2D-2D matches between the query (left) and database (right) images. For each pair, the retrieved database image with the largest number of inliers is selected, and only the inlier matches are visualized. We show that our method is able to produce accurate correspondences for challenging queries.

7.1.10. Moulding Humans: Non-parametric 3D Human Shape Estimation from Single Images

Participants: Valentin Gabeur, Jean-Sébastien Franco [Inria Morpheo], Xavier Martin, Cordelia Schmid, Gregory Rogez [NAVER LABS Europe].

While the recent progress in convolutional neural networks has allowed impressive results for 3D human pose estimation, estimating the full 3D shape of a person is still an open issue. Model-based approaches can output precise meshes of naked under-cloth human bodies but fail to estimate details and un-modelled elements such as hair or clothing. On the other hand, non-parametric volumetric approaches can potentially estimate complete shapes but, in practice, they are limited by the resolution of the output grid and cannot produce detailed estimates. In this paper [19], we propose a non-parametric approach that employs a double depth map 10 to represent the 3D shape of a person: a visible depth map and a "hidden" depth map are estimated and combined, to reconstruct the human 3D shape as done with a "mould". This representation through 2D depth maps allows a higher resolution output with a much lower dimension than voxel-based volumetric representations.

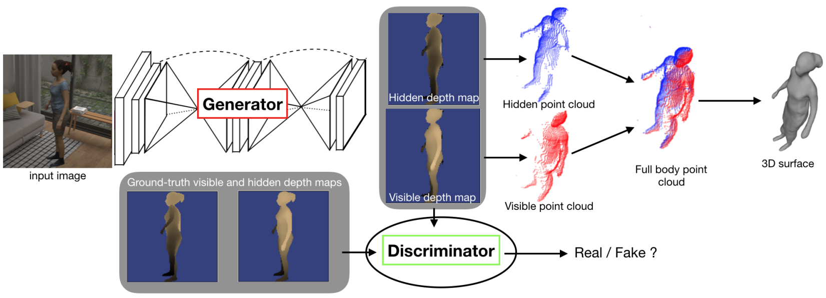


Figure 10. Given a single image, we estimate the "visible" and the "hidden" depth maps. The 3D point clouds of these 2 depth maps are combined to form a full-body 3D point cloud, as if lining up the 2 halves of a "mould". The 3D shape is then reconstructed using Poisson reconstruction. An adversarial training with a discriminator is employed to increase the humanness of the estimation.

7.1.11. Focused Attention for Action Recognition

Participants: Vladyslav Sydorov, Karteek Alahari.

In this paper [30], we introduce an attention model for video action recognition that allows processing video in higher resolution, by focusing on the relevant regions first. The network-specific saliency is utilized to guide the cropping, we illustrate the procedure in Figure 11. We show performance improvement on the Charades dataset with this strategy.

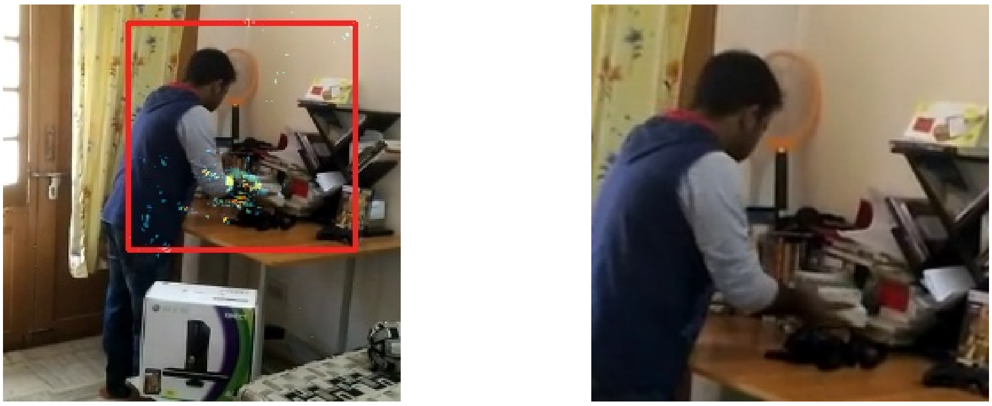


Figure 11. Example of attention on Charades action recognition dataset. (Left) Saliency scores (displayed as a heatmap) are localized around the object, a box maximizing the saliency measure within is selected. (Right) The network is provided with the relevant crop of the video, and can process it at a higher resolution.

7.2. Statistical Machine Learning

7.2.1. 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 , 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.

7.2.2. A Generic Acceleration Framework for Stochastic Composite Optimization

Participants: Andrei Kulunchakov, Julien Mairal.

In [35], we introduce various mechanisms to obtain accelerated first-order stochasticoptimization algorithms when the objective function is convex or stronglyconvex. Specifically, we extend the Catalyst approach originally designed fordeterministic objectives to the stochastic setting. Given an optimizationmethod with mild convergence guarantees for strongly convex problems,the challenge is to accelerate convergence to a noise-dominated region, andthen achieve convergence with an optimal worst-case complexity depending on thenoise variance of the gradients.A side contribution of our work is also a generic analysis that canhandle inexact proximal operators, providing new insights about the robustness of stochastic algorithms when the proximal operator cannot be exactly computed. An illustration from this work is explained in Figure 12.

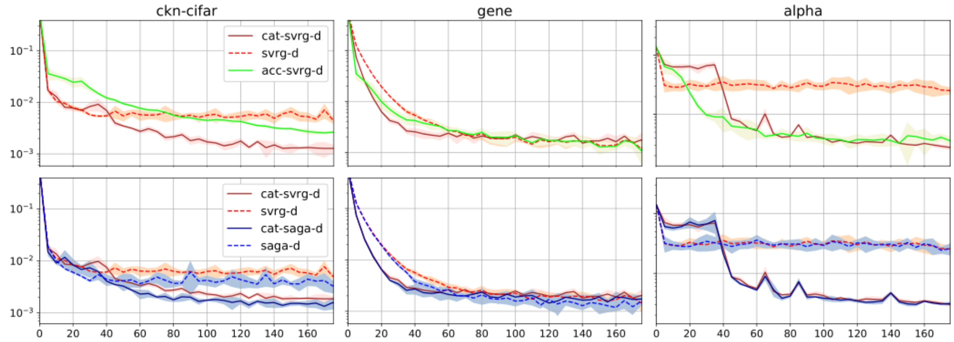


Figure 12. Accelerating SVRG-like (top) and SAGA (bottom) methods for ℓ_2-logistic regression with $\mu = 1/(100n)$ (bottom) for mild dropout, which imitates stochasticity in the gradients. All plots are on a logarithmic scale for the objective function value, and the x-axis denotes the number of epochs. The colored tubes around each curve denote a standard deviations across 5 runs. The curves show that acceleration may be useful even in the stochastic optimization regime.

7.2.3. Estimate Sequences for Variance-Reduced Stochastic Composite Optimization

Participants: Andrei Kulunchakov, Julien Mairal.

In [23], we propose a unified view of gradient-based algorithms for stochastic convex composite optimization. By extending the concept of estimate sequence introduced by Nesterov, we interpret a large class of stochastic optimization methods as procedures that iteratively minimize a surrogate of the objective. This point of view covers stochastic gradient descent (SGD), the variance-reduction approaches SAGA, SVRG, MISO, their proximal variants, and has several advantages: (i) we provide a simple generic proof of convergence for all of the aforementioned methods; (ii) we naturally obtain new algorithms with the same guarantees; (iii) we derive generic strategies to make these algorithms robust to stochastic noise, which is useful when data is corrupted by small random perturbations. Finally, we show that this viewpoint is useful to obtain accelerated algorithms. A comparison with different approaches is shown in Figure 13.

7.2.4. White-box vs Black-box: Bayes Optimal Strategies for Membership Inference

Participants: Alexandre Sablayrolles, Matthijs Douze, Yann Ollivier, Cordelia Schmid, Hervé Jégou.

Membership inference determines, given a sample and trained parameters of a machine learning model, whether the sample was part of the training set. In this paper [28], we derive the optimal strategy for membership inference with a few assumptions on the distribution of the parameters. We show that optimal attacks only depend on the loss function, and thus black-box attacks are as good as white-box attacks. As the optimal strategy is not tractable, we provide approximations of it leading to several inference methods 14, and show that existing membership inference methods are coarser approximations of this optimal strategy. Our membership attacks outperform the state of the art in various settings, ranging from a simple logistic regression to more complex architectures and datasets, such as ResNet-101 and Imagenet.

7.3. Theory and Methods for Deep Neural Networks

7.3.1. Group Invariance, Stability to Deformations, and Complexity of Deep Convolutional Representations

Participants: Alberto Bietti, Julien Mairal.

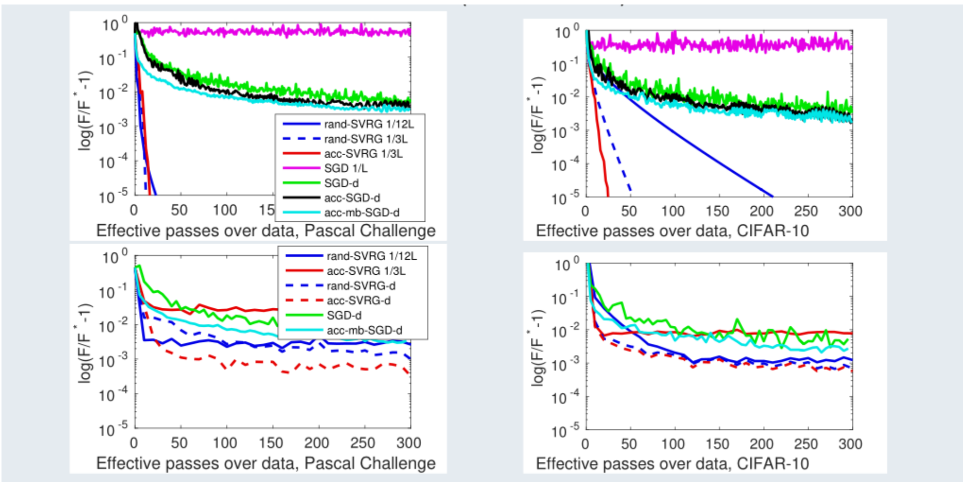


Figure 13. Comparison of different standard approaches with our developed method on two datasets for ℓ_2-logistic regression with mild dropout (bottom) and deterministic case (above). The case of exact gradient computations clearly shows benefits from acceleration, which consist in fast linear convergence. In the stochastic case, we demonstrate either superiority or high competitiveness of the developed method along with its unbiased convergence to the optimum. In both cases, we show that acceleration is able to generically comprise strengths of standard methods and even outperform them.

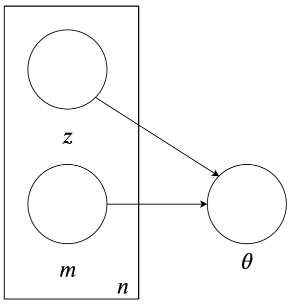


Figure 14. Plate notation of the membership inference problem: for each data point z_i, a binary membership variable m_i is sampled, and z_i belongs to the training set iff $m_i = 1$. Given the trained parameters θ and a sample z_i, we want to infer the value of m_i.

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 the paper [3], 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.

7.3.2. A Kernel Perspective for Regularizing Deep Neural Networks

Participants: Alberto Bietti, Grégoire Mialon, Dexiong Chen, Julien Mairal.

We propose a new point of view for regularizing deep neural networks by using the norm of a reproducing kernel Hilbert space (RKHS) [12]. Even though this norm cannot be computed, it admits upper and lower approximations leading to various practical strategies. Specifically, this perspective (i) provides a common umbrella for many existing regularization principles, including spectral norm and gradient penalties, or adversarial training, (ii) leads to new effective regularization penalties, and (iii) suggests hybrid strategies combining lower and upper bounds to get better approximations of the RKHS norm. We experimentally show this approach to be effective when learning on small datasets, or to obtain adversarially robust models.

7.3.3. On the Inductive Bias of Neural Tangent Kernels

Participants: Alberto Bietti, Julien Mairal.

State-of-the-art neural networks are heavily over-parameterized, making the optimization algorithm a crucial ingredient for learning predictive models with good generalization properties. A recent line of work has shown that in a certain over-parameterized regime, the learning dynamics of gradient descent are governed by a certain kernel obtained at initialization, called the neural tangent kernel. In [12], we study the inductive bias of learning in such a regime by analyzing this kernel and the corresponding function space (RKHS). In particular, we study smoothness, approximation, and stability properties of functions with finite norm, including stability to image deformations in the case of convolutional networks, and compare to other known kernels for similar architectures.

7.3.4. Large Memory Layers with Product Keys

Participants: Guillaume Lample, Alexandre Sablayrolles, Marc'Aurelio Ranzato, Ludovic Denoyer, Hervé Jégou.

This paper introduces a structured memory which can be easily integrated into a neural network. The memory is very large by design and significantly increases the capacity of the architecture, by up to a billion parameters with a negligible computational overhead. Its design and access pattern is based on product keys, which enable fast and exact nearest neighbor search. The ability to increase the number of parameters while keeping the same computational budget lets the overall system strike a better trade-off between prediction accuracy and computation efficiency both at training and test time. This memory layer, shown in Figure 15, allows us to tackle very large scale language modeling tasks. In our experiments we consider a dataset with up to 30 billion words, and we plug our memory layer in a state-of-the-art transformer-based architecture. In particular, we found that a memory augmented model with only 12 layers outperforms a baseline transformer model with 24 layers, while being twice faster at inference time. We release our code for reproducibility purposes.

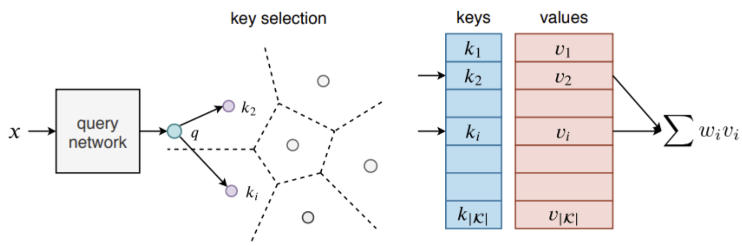


Figure 15. Overview of a key-value memory layer: The input x is processed through a query network that produces a query vector q, which is compared to all the keys. The output is the sparse weighted sum over the memories associated with the selected keys. For a large number of keys $|\mathcal{K}|$, the key selection procedure becomes too expensive in practice. Our product key method is exact and makes this search process very fast.

7.3.5. Understanding Priors in Bayesian Neural Networks at the Unit Level

Participants: Mariia Vladimirova, Jakob Verbeek, Pablo Mesejo [Univ. Granada, Spain], Julyan Arbel [Inria MISTIS].

In our ICML'19 paper [31], we investigate deep Bayesian neural networks with Gaussian weight priors and a class of ReLUlike nonlinearities. Bayesian neural networks with Gaussian priors are well known to induce an L2, "weight decay", regularization. Our results characterize a more intricate regularization effect at the level of the unit activations. Our main result establishes that the induced prior distribution on the units before and after activation becomes increasingly heavy-tailed with the depth of the layer. We show that first layer units are Gaussian, second layer units are sub-exponential, and units in deeper layers are characterized by sub-Weibull distributions. Our results provide new theoretical insight on deep Bayesian neural networks, which we corroborate with experimental simulation results.

7.3.6. Adaptative Inference Cost With Convolutional Neural Mixture Models

Participants: Adria Ruiz, Jakob Verbeek.

Despite the outstanding performance of convolutional neural networks (CNNs) for many vision tasks, the required computational cost during inference is problematic when resources are limited. In this paper [27], we propose Convolutional Neural Mixture Models (CNMMs), a probabilistic model embedding a large number of CNNs that can be jointly trained and evaluated in an efficient manner. Within the proposed framework, we present different mechanisms to prune subsets of CNNs from the mixture, allowing to easily adapt the computational cost required for inference (see Fig. 16). Image classification and semantic segmentation experiments show that our method achieve excellent accuracy-compute trade-offs. Moreover, unlike most of previous approaches, a single CNMM provides a large range of operating points along this trade-off, without any re-training.

7.4. Pluri-disciplinary Research

7.4.1. Biological Sequence Modeling with Convolutional Kernel Networks

Participants: Dexiong Chen, Laurent Jacob, 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

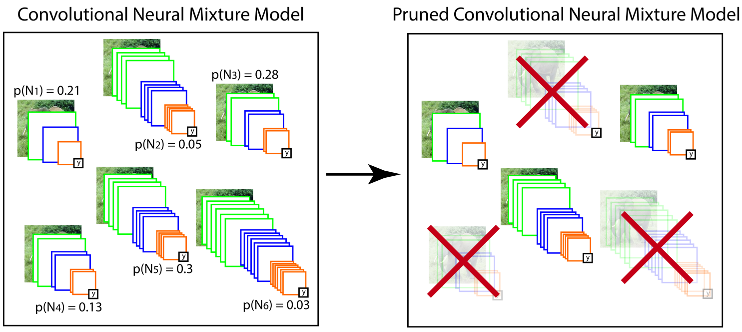


Figure 16. A Convolutional Neural Mixture Model embeds a large number of CNNs. Weight sharing enables efficient joint training of all networks and computation of the mixture output. The learned mixing weights can be used to remove networks from the mixture, and thus reduce the computational cost of inference.

unannotated sequences with good accuracy. Unfortunately, their performance with medium- or small-scale datasets is mitigated, which requires inventing new data-efficient approaches. In this paper [4], [14], we introduce a hybrid approach between convolutional neural networks and kernel methods to model biological sequences. Our method, shown in Figure 17, 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.

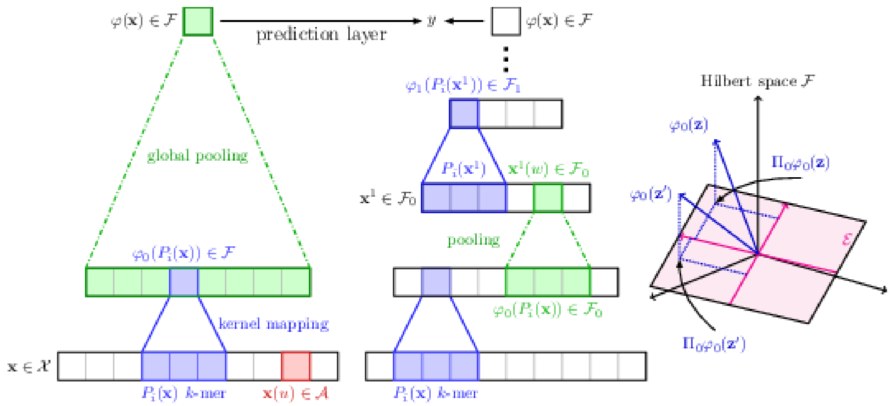


Figure 17. Construction of single-layer (left) and multilayer (middle) CKN-seq and the approximation of one layer (right). For a single-layer model, each k-mer $P_i(\mathbf{x})$ is mapped to $\varphi_0(P_i(\mathbf{x}))$ in $\mathcal{F}$ and projected to $\Pi\varphi_0(P_i(\mathbf{x}))$ parametrized by $\psi_0(P_i(\mathbf{x}))$. 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.

7.4.2. Recurrent Kernel Networks

Participants: Dexiong Chen, Laurent Jacob [CNRS, LBBE Laboratory], Julien Mairal.

Substring kernels are classical tools for representing biological sequences or text. However, when large amounts of annotated data are available, models that allow end-to-end training such as neural networks are often preferred. Links between recurrent neural networks (RNNs) and substring kernels have recently been drawn, by formally showing that RNNs with specific activation functions were points in a reproducing kernel Hilbert space (RKHS). In this paper [15], we revisit this link by generalizing convolutional kernel networks—–originally related to a relaxation of the mismatch kernel—–to model gaps in sequences. It results in a new type of recurrent neural network (Figure 18), which can be trained end-to-end with backpropagation, or without supervision by using kernel approximation techniques. We experimentally show that our approach is well suited to biological sequences, where it outperforms existing methods for protein classification tasks.

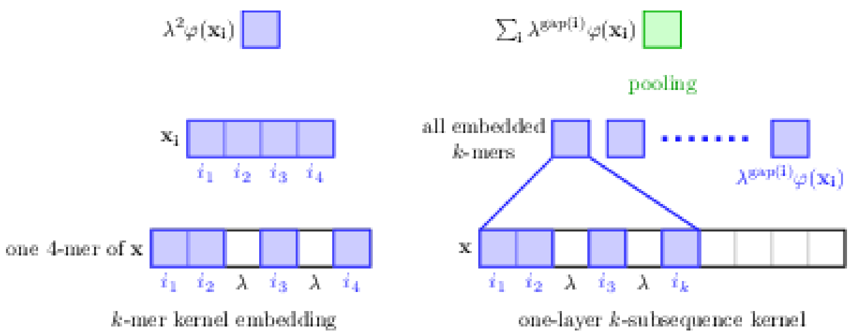


Figure 18. Representation of a sequence in a RKHS based on our kernel.

7.4.3. Depth-adaptive Transformer

Participants: Maha Elbayad, Jiatao Gu [Facebook AI], Edouard Grave [Facebook AI], Michael Auli [Facebook AI].

State of the art sequence-to-sequence models for large scale tasks perform a fixed number of computations for each input sequence regardless of whether it is easy or hard to process. In our ICLR'2020 paper [18], we train Transformer models which can make output predictions at different stages of the network and we investigate different ways to predict how much computation is required for a particular sequence. Unlike dynamic computation in Universal Transformers, which applies the same set of layers iteratively, we apply different layers at every step to adjust both the amount of computation as well as the model capacity. On IWSLT German-English translation our approach matches the accuracy of a well tuned baseline Transformer while using less than a quarter of the decoder layers. Figure 19 illustrates the different halting mechanisms investigated in this work. Namely, a sequence-level approach where we assume all the sequence's tokens are equally difficult and a token-level approach where tokens exit at varying depths.

8. Bilateral Contracts and Grants with Industry

8.1. 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.

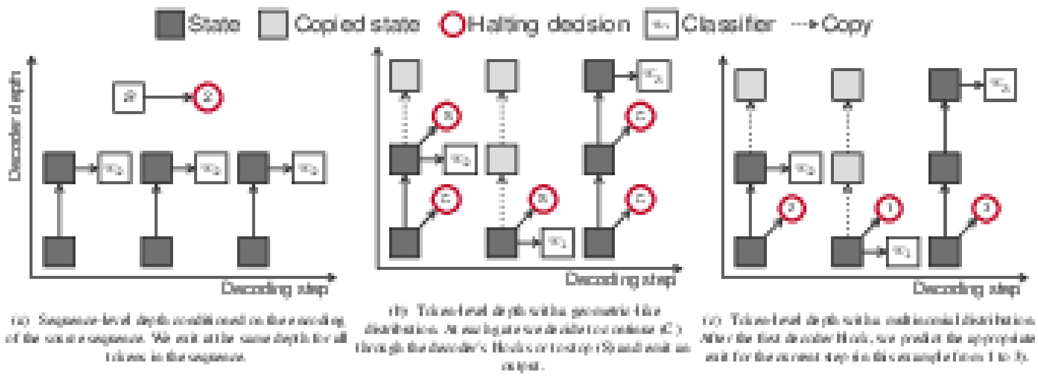


Figure 19. Illustration of the variant adaptive depth predictors: (a) the sequence-level and (b, c) at the token-level.

8.2. Facebook

Participants: Cordelia Schmid, Jakob Verbeek, Julien Mairal, Karteek Alahari, Pauline Luc, Alexandre Sablayrolles, Mathilde Caron, Lina Mezghani.

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. Lina Mezghani is the new PhD student in this collaboration since 2019.

8.3. NAVER LABS Europe

Participant: Karteek Alahari.

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. A new CIFRE PhD contract was submitted to ANRT for approval in October 2019.

8.4. Valeo AI

Participants: Karteek Alahari, Florent Bartoccioni.

This collaboration started in 2019 with the arrival of PhD student Florent Bartoccioni. Despite the progress seen in computer vision, artificial systems lack the capability to address the large disparity between human and machine-based scene understanding. For example, at any road intersection most people have the ability to accurately forecast or anticipate events in this scenario, such as changes in colour of the traffic lights, when and how pedestrians are likely to cross the street. This apparently natural human behaviour is not replicable by state-of-theart computer vision methods, which are ill-equipped to make such forecasts. The goal of this collaborative PhD is to address this forecasting problem.

8.5. Criteo

Participant: Julien Mairal.

This collaboration started in April 2019, with the arrival of a master student, Houssam Zenati, who will pursue a CIFRE PhD starting in 2020. The goal of this collaboration is to develop machine learning techniques for counterfactual loss optimization, which is a fundamental problem in machine learning related to causal inference. The goal is to learn stochastic policies, based on offline logged data. The problem is important for web advertising, which is the main activity of the Criteo company, but the potential scope of application is much larger, with possible applications in medicine and experimental sciences.

8.6. Google

Participants: Karteek Alahari, Minttu Alakuijala, Valentin Gabeur, Julien Mairal.

This collaboration started in February 2019, with the arrival of two CIFRE PhD students, Minttu Alakuijala and Valentin Gabeur, who are respectively working on visual models for robotics, and 3D human pose estimation.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. MIAI chair - Towards more data efficiency in machine learning

Participants: Julien Mairal, Karteek Alahari, Jakob Verbeek.

Julien Mairal holds a chair of the 3IA MIAI institute. The goal is to improve the data efficiency of machine learning algorithms.

9.1.2. MIAI chair - Towards self-supervised visual learning

Participant: Cordelia Schmid.

Cordelia Schmid holds a chair of the 3IA MIAI institute. The goal is to develop new self-supervised learning methods for computer vision.

9.1.3. MIAI chair - Multiscale, multimodal and multitemporal remote sensing

Participant: Jocelyn Chanussot.

Jocelyn Chanussot holds a chair of the 3IA MIAI institute.

9.1.4. 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).

9.1.5. 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. This collaboration ended in 2019.

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, 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 and ended in March 2019. 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 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, Konstantin Shmelkov, Vladyslav Sydorov, Daan Wynen, Nikita 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) - Machine Learning Department - Katerina Fragkiadaki

Start year: 2019

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

We propose to renew the associate team GAYA, with the primary goal of interpreting videos in terms of recognizing actions, understanding the human-human and human-object interactions. In the first three years, the team has started addressing the problem of learning an efficient and robust video representation to attack this challenge. GAYA will now focus on building semantic models, wherein we learn incremental, joint audio-visual models, 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 (Thoth and WILLOW), a

US university (Carnegie Mellon University [CMU]) as the main partner team, and another US university (UC Berkeley) as a secondary partner. It will allow the partners to effectively combine their respective strengths in areas such as inference and machine learning approaches for vision tasks, joint audio-visual models, large-scale learning, geometric reasoning. The main expected outcomes of this collaboration are: new machine learning algorithms for handling minimally annotated multi-modal data, large-scale public datasets for benchmarking, theoretical analysis of objects shapes and contours. This associate team originally started in 2016, and was extended in 2019 for another 3 years.

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. In 2019, the project resulted in the development of an approach for object interaction [20].

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 until Jan 2019.
- Avijit Dasgupta (PhD Student, IIIT Hyderabad, India) was an intern in the team from Feb to May 2019.
- Gunnar Sigurdsson (PhD student, CMU) was an intern in the team from Jan to Mar 2019.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- C. Schmid is a general chair for ECCV 2020, ICCV 2023.

10.1.1.2. Member of the Organizing Committees

- K. Alahari and J. Mairal co-organized the international summer school PAISS 2019.
- J. Mairal is a member of the organizing committee for the international conference SIAM Imaging Science 2020.
- J. Mairal is a co-organizer of the workshop OSL’19 at Les Houches.
- J. Mairal co-organized a discussion session at the Ellis/Dali workshop, San Sebastian, 2019.

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

- K. Alahari: area chair for CVPR 2020, ICCV 2019.
- K. Alahari: senior program committee member for AAAI 2020, IJCAI 2019, IJCAI 2020.
- K. Alahari: doctoral consortium chair for ICCV 2023.
- J. Mairal: area chair for NeurIPS 2019, AISTATS 2020 and ECCV 2020.
- J. Mairal: tutorial chair for CVPR 2022.
- C. Schmid: area chair for ICCV 2019.
- C. Schmid: senior area chair for NeurIPS 2019.
- J. Verbeek: area chair for ICCV 2019.

10.1.2.2. Reviewer

The permanent members, postdocs and senior PhD students of the team reviewed numerous papers for international conferences in artificial intelligence, computer vision and machine learning, including AAAI, AISTATS, CVPR, ICCV, ICML, ICLR, NeurIPS in 2019.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- K. Alahari: Associate editor of the International Journal of Computer Vision, since 2019.
- K. Alahari: Associate editor for Computer Vision and Image Understanding journal, since 2018.
- J. Mairal: Associate editor of the Journal of Machine Learning Research (JMLR), since 2019.
- 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.
- J. Verbeek: Associate editor International Journal on Computer Vision, 2014-2019.
- J. Verbeek: Associate editor IEEE Transactions Pattern Analysis and Machine Intelligence, since 2018.

10.1.3.2. Reviewer - Reviewing Activities

The permanent members, postdocs and senior PhD students of the team reviewed numerous papers for international journals in computer vision (IJCV, PAMI, CVIU), machine learning (JMLR, Machine Learning). Some of them also review for other reputed journals such as PLOS ONE, SIAM Journal on Optimization, SIAM Imaging Science.

10.1.4. Invited Talks

- K. Alahari: Speaker on the Panel on AI and Mathematics, Knowledge Summit, Lyon, France, 2019.
- K. Alahari: Invited talk, LIAMA workshop, Paris, France, 2019.
- A. Bietti: Invited talk, GIPSA Lab, Grenoble, 2019.
- A. Bietti: Invited talk, UC Berkeley, 2019.
- A. Bietti: Seminar, Microsoft Research AI, Redmond, 2019.
- A. Bietti: Seminar, TTI-Chicago, 2019.
- D. Chen: Machine Learning in Computational Biology (MLCB) workshop on recurrent kernel networks, Vancouver, 2019.
- J. Mairal: Invited talk at the YES workshop, Eindhoven, 2019.
- J. Mairal: Talk in mini-symposium, ICCOPT, Berlin, 2019.
- J. Mairal: Invited talk at the Imaging and Machine Learning conference, IHP, Paris, 2019.

- J. Mairal: Seminar. Centrale Lille, 2019.
- R. Klokov: Invited talk at Christmas Colloquium on Computer Vision, Yandex, Moscow, 2019.
- C. Schmid: Invited speaker at BMVA symposium in Video Understanding, London, September 2019.
- C. Schmid: Keynote speaker at BMVC, Cardiff, UK, September 2019.
- C. Schmid: Keynote speaker at SIGIR, Paris, July 2019.
- C. Schmid: Invited speaker at Computer Vision after 5 Years, in conjunction with CVPR, June 2019.
- C. Schmid: Invited speaker at Tutorial on Unifying Human Activity Understanding, in conjunction with CVPR, June 2019.
- C. Schmid: Invited speaker at Facebook AI Video Summit, June 2019.
- C. Schmid: Keynote speaker at AI Experts Workshop in conjunction with the AI for Good Global Summit, Geneva, May 2019.
- C. Schmid: Invited speaker at Women in Data Science Conference, Zürich, April 2019.
- C. Schmid: Invited speaker at Collège de France seminar (chair of Stephane Mallat), February 2019.
- C. Schmid: Talk at Google EMEA research days, Zurich, December 2019.
- C. Schmid: Talk at Workshop on AI for Robotics, Naver, Grenoble, November 2019.
- C. Schmid: Talk at Workshop, Robotics: A Challenge for the Artificial Intelligence, Toulouse, October 2019.
- C. Schmid: Presentation at PRAIRIE inauguration, Paris, October 2019.
- C. Schmid: Seminar at DeepMind, London, September 2019.
- C. Schmid: Seminar at Intel Network on Intelligent Systems, Munich, September 2019.
- C. Schmid: Seminar at Ellis workshop, September 2019.
- C. Schmid: Seminar at MPI Tübingen, July 2019.
- C. Schmid: Seminar at WILLOW/SIERRA retreat, Marseille, June 2019.
- C. Schmid: Dinner speaker at the workshop "Women in Computer Vision", in conjunction with CVPR'19.
- C. Schmid: Seminar at Google MTV, April 2019.
- C. Schmid: Seminar at ETH Zürich, March 2019.
- J. Verbeek: Invited talk at Breaking the Surface Workshop on maritime robotics and its applications, Biograd na Moru, Croatia, Oct 2019.
- J. Verbeek: Invited talk at Dagstuhl Workshop on Joint Processing of Language and Visual Data for Better Automated Understanding, Germany, Jan 2019.
- D. Wynen: SMILE Reading Group Paris, 2019.

10.1.5. Leadership within the Scientific Community

- J. Mairal, J. Verbeek and C. Schmid became Ellis fellows.
- C. Schmid: Participation in a round table on AI, a technology for innovation, forum 5i, Grenoble, May 2019.
- C. Schmid: Animating several mentorship sessions at Women in Data Science Conference, Zürich, April 2019.
- C. Schmid: Mentor at the Doctoral Consortium, in conjunction with ICCV'19, CVPR'19.
- C. Schmid: Mentor for female PhD students at the workshop "Women in Computer Vision", CVPR'19.

10.1.6. Scientific Expertise

- J. Mairal: Judge for the IBM Watson AI Xprize.
- J. Mairal: Expert for ANR.

10.1.7. Research Administration

- K. Alahari: One of the two referents for Human Resources - Excellence in Research (HRS4R) at Inria Grenoble.
- J. Mairal: Jury member for the Inria starting and advanced research positions.
- C. Schmid: Member of Scientific Advisory Committee of the Helmholtz AI Cooperation Unit, 2020—
- C. Schmid: Member of scientific advisory board for the German Competence Centers for AI Research, 2019—
- J. Verbeek: Member steering committee MinaLogic, innovation cluster for digital technologies based in France's Auvergne-Rhône-Alpes region, 2018-2019.
- J. Verbeek: Scientific correspondent national project calls, Inria Grenoble, 2017-2019.
- J. Verbeek: Member Scientific council Advanced Data Mining axis of Persyval Laboratory of Excellence, Grenoble, 2015-2019.
- J. Verbeek: Member Inria Grenoble working group on HPC - Big Data - Machine learning, 2018-2019.
- J. Verbeek: Member of Inria Commission Administrative Paritaire (advises on matters about individual careers: such as promotions, temporary outsourcing, etc.), 2016-2019.

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: J. Mairal, Large-Scale Optimization for Machine Learning, 9h eqTD, Lecture at OBA summer school, Veroli, 2019.
- Doctorat: J. Mairal, Large-Scale Optimization for Machine Learning, 4.5h eqTD. Invited Tutorial at IEEE Data Science Workshop on Minneapolis, 2019.
- Doctorat: J. Mairal, Large-Scale Optimization for Machine Learning, 4.5h eqTD. Invited Tutorial at the Quantitative BioImaging Conference, Rennes, 2019.
- Doctorat: C. Schmid, Course on action recognition at Data Science Summer School, Paris, June 2019.
- Doctorat: C. Schmid, Course on action recognition at Prairie artificial intelligence summer school (PAISS), 2.25h eqTD, Paris, October 2019.
- Master: K. Alahari, C. Schmid, Object recognition, Master-2 Computer Science, Grenoble University, 15.75h eqTD together, 2019.
- 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: J. Mairal, Advanced Learning Models, 13.5h eqTD, M2, UGA, Grenoble.
- Master: C. Schmid, Object recognition and computer vision, Master-2 MVA, ENS, 9h eqTD, 2019.

- Master: A. Sablayrolles, Fundamentals of Machine Learning, African Masters of Machine Intelligence, Kigali, Rwanda.

10.2.2. Supervision

HDR: Karteek Alahari, Human, Motion and Other Priors for Partially-Supervised Recognition, Univ. Grenoble Alpes, 28/1/2019.

PhD: Alberto Bietti, Foundations of deep convolutional models through kernel methods, Univ. Grenoble Alpes, 27/11/2019, director: Julien Mairal.

PhD: Nikita Dvornik, Learning with Limited Annotated Data for Visual Understanding, Univ. Grenoble Alpes, 26/11/2019, thesis directors: Cordelia Schmid and Julien Mairal.

PhD: Konstantin Shmelkov, Approaches for incremental learning and image generation, Univ. Grenoble Alpes, 29/3/2019, thesis directors: Karteek Alahari and Cordelia Schmid.

10.2.3. Juries

- K. Alahari: External examiner for the PhD thesis of Alessandro di Martino, University of Bath, UK.
- K. Alahari: Examiner for the PhD thesis of Thomas Robert, Sorbonne Université, Paris, France.
- K. Alahari: Examiner for the PhD thesis of D. Khuê Lê-Huu, Université Paris-Saclay, France.
- K. Alahari: Member of comité de suivi for the PhD thesis of Miguel Angel Solinas, Univ. Grenoble-Alpes, France.
- J. Mairal: Reviewer for the PhD thesis of Zhenyu Liao, Université Paris-Saclay.
- J. Mairal: Reviewer for the PhD thesis of Belhal Karimi, Université Paris-Saclay
- J. Mairal: Reviewer for the PhD thesis of Martin Bompaire, Université Paris-Saclay
- J. Mairal: Reviewer for the PhD thesis of Yassine Yaakoubi, Polytechnique Montréal.
- J. Mairal: Examinateur for the PhD thesis of Mathurin Massias, Université Paris-Saclay.
- J. Mairal: Member of comité de suivi for the PhD thesis of Olga Permiakova, Univ. Grenoble Alpes.
- J. Verbeek: Member supervisory commitee for PhD of Riccardo Del Chiaro, 2018-2020, Univ. Florence, Italy.
- J. Verbeek: Member supervisory commitee for PhD of Fabien Baradel, 2017-2019, INSA Lyon, France.
- J. Verbeek: External reviewer for Shell Xu Hu, 2019, Ecole des Ponts, Paris Tech, Univ. Paris Est, Paris, France.
- J. Verbeek: Rapporteur for Hedi Ben Younes, 2019, Sorbonne University, Paris, France.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] K. ALAHARI.*Human, Motion and Other Priors for Partially-Supervised Recognition*, Communauté Université Grenoble Alpes, January 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02269024

[2] K. SHMELKOV.*Approaches for incremental learning and image generation*, Université Grenoble Alpes, March 2019, https://tel.archives-ouvertes.fr/tel-02183259

Articles in International Peer-Reviewed Journal

[3] A. BIETTI, J. MAIRAL. *Group Invariance, Stability to Deformations, and Complexity of Deep Convolutional Representations*, in "Journal of Machine Learning Research", 2019, vol. 20, n^o^ 1, p. 1-49, https://arxiv.org/abs/1706.03078 , https://hal.inria.fr/hal-01536004

[4] D. CHEN, L. JACOB, J. MAIRAL. *Biological Sequence Modeling with Convolutional Kernel Networks*, in "Bioinformatics", September 2019, vol. 35, n^o^ 18, p. 3294–3302 [*DOI :* 10.1093/BIOINFORMATICS/BTZ094], https://hal.inria.fr/hal-01632912

[5] D. DERKACH, A. RUIZ, F. M. SUKNO. *Tensor Decomposition and Non-linear Manifold Modeling for 3D Head Pose Estimation*, in "International Journal of Computer Vision", October 2019, vol. 127, n^o^ 10, p. 1565-1585 [*DOI :* 10.1007/S11263-019-01208-X], https://hal.archives-ouvertes.fr/hal-02267568

[6] G. DURIF, L. MODOLO, J. E. MOLD, S. LAMBERT-LACROIX, F. PICARD. *Probabilistic Count Matrix Factorization for Single Cell Expression Data Analysis*, in "Bioinformatics", October 2019, vol. 20, p. 4011–4019, https://arxiv.org/abs/1710.11028 [*DOI :* 10.1093/BIOINFORMATICS/BTZ177], https://hal.archives-ouvertes.fr/hal-01649275

[7] N. DVORNIK, J. MAIRAL, C. SCHMID. *On the Importance of Visual Context for Data Augmentation in Scene Understanding*, in "IEEE Transactions on Pattern Analysis and Machine Intelligence", December 2019, p. 1-15, forthcoming [*DOI :* 10.1109/TPAMI.2019.2961896], https://hal.archives-ouvertes.fr/hal-01869784

[8] H. LIN, J. MAIRAL, Z. HARCHAOUI. *An Inexact Variable Metric Proximal Point Algorithm for Generic Quasi-Newton Acceleration*, in "SIAM Journal on Optimization", May 2019, vol. 29, n^o^ 2, p. 1408-1443, https://arxiv.org/abs/1610.00960 [*DOI :* 10.1137/17M1125157], https://hal.inria.fr/hal-01376079

[9] 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, p. 1-15, forthcoming [*DOI :* 10.1109/TPAMI.2019.2892985], https://hal.archives-ouvertes.fr/hal-01961189

[10] P. TOKMAKOV, C. SCHMID, K. ALAHARI. *Learning to Segment Moving Objects*, in "International Journal of Computer Vision", March 2019, vol. 127, n^o^ 3, p. 282–301, https://arxiv.org/abs/1712.01127 [*DOI :* 10.1007/S11263-018-1122-2], https://hal.archives-ouvertes.fr/hal-01653720

International Conferences with Proceedings

[11] A. BIETTI, J. MAIRAL. *On the Inductive Bias of Neural Tangent Kernels*, in "NeurIPS 2019 - Thirty-third Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, p. 1-24, https://arxiv.org/abs/1905.12173 , https://hal.inria.fr/hal-02144221

[12] A. BIETTI, G. MIALON, D. CHEN, J. MAIRAL. *A Kernel Perspective for Regularizing Deep Neural Networks*, in "ICML 2019 - 36th International Conference on Machine Learning", Long Beach, United States, Proceedings of Machine Learning Research, June 2019, vol. 97, p. 664-674, https://arxiv.org/abs/1810.00363 , https://hal.inria.fr/hal-01884632

[13] M. CARON, P. BOJANOWSKI, J. MAIRAL, A. JOULIN. *Unsupervised Pre-Training of Image Features on Non-Curated Data*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea,

Proceedings of the International Conference on Computer Vision (ICCV), October 2019, p. 1-10, https://hal.archives-ouvertes.fr/hal-02119564

[14] D. CHEN, L. JACOB, J. MAIRAL.*Biological Sequence Modeling with Convolutional Kernel Networks*, in "RECOMB 2019 - 23rd Annual International Conference Research in Computational Molecular Biology", Washington DC, United States, Springer, May 2019, p. 1-2 [*DOI :* 10.1007/978-3-030-17083-7], https://hal.archives-ouvertes.fr/hal-02388776

[15] D. CHEN, L. JACOB, J. MAIRAL.*Recurrent Kernel Networks*, in "NeurIPS 2019 - Thirty-third Conference Neural Information Processing Systems", Vancouver, Canada, December 2019, p. 1-19, https://arxiv.org/abs/1906.03200 , https://hal.inria.fr/hal-02151135

[16] N. CRASTO, P. WEINZAEPFEL, K. ALAHARI, C. SCHMID.*MARS: Motion-Augmented RGB Stream for Action Recognition*, in "CVPR 2019 - IEEE Conference on Computer Vision & Pattern Recognition", Long Beach, CA, United States, IEEE, June 2019, p. 1-10, https://hal.inria.fr/hal-02140558

[17] N. DVORNIK, C. SCHMID, J. MAIRAL.*Diversity with Cooperation: Ensemble Methods for Few-Shot Classification*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, p. 1-12, https://arxiv.org/abs/1903.11341 - Added experiments with different network architectures and input image resolutions, https://hal.archives-ouvertes.fr/hal-02080004

[18] M. ELBAYAD, J. GU, E. GRAVE, M. AULI.*Depth-adaptive Transformer*, in "ICLR 2020 - Eighth International Conference on Learning Representations", Addis Ababa, Ethiopia, December 2019, p. 1-14, https://hal.inria.fr/hal-02422914

[19] V. GABEUR, J.-S. FRANCO, X. MARTIN, C. SCHMID, G. ROGEZ.*Moulding Humans: Non-parametric 3D Human Shape Estimation from Single Images*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, p. 1-10, https://hal.inria.fr/hal-02242795

[20] Y. HASSON, G. VAROL, D. TZIONAS, I. KALEVATYKH, M. J. BLACK, I. LAPTEV, C. SCHMID.*Learning joint reconstruction of hands and manipulated objects*, in "CVPR 2019 - IEEE Conference on Computer Vision and Pattern Recognition", Long Beach, United States, IEEE, June 2019, p. 1-14, https://hal.archives-ouvertes.fr/hal-02429093

[21] R. KLOKOV, J. VERBEEK, E. BOYER.*Probabilistic Reconstruction Networks for 3D Shape Inference from a Single Image*, in "BMVC 2019 - British Machine Vision Conference", Cardiff, United Kingdom, September 2019, p. 1-15, https://arxiv.org/abs/1908.07475 - Awarded with Best Science Paper Honourable Mention Award at BMVC'19., https://hal.inria.fr/hal-02268466

[22] A. KULUNCHAKOV, J. MAIRAL.*A Generic Acceleration Framework for Stochastic Composite Optimization*, in "NeurIPS 2019 - Thirty-third Conference Neural Information Processing Systems", Vancouver, Canada, December 2019, p. 1-24, https://arxiv.org/abs/1906.01164 , https://hal.inria.fr/hal-02139489

[23] A. KULUNCHAKOV, J. MAIRAL.*Estimate Sequences for Variance-Reduced Stochastic Composite Optimization*, in "ICML 2019 - 36th International Conference on Machine Learning", Long Beach, United States, June 2019, p. 1-24, https://arxiv.org/abs/1905.02374 - short version of preprint arXiv:1901.08788, https://hal.inria.fr/hal-02121913

[24] T. LUCAS, K. SHMELKOV, K. ALAHARI, C. SCHMID, J. VERBEEK.*Adaptive Density Estimation for Generative Models*, in "NeurIPS 2019 - Thirty-third Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, p. 1-24, https://hal.archives-ouvertes.fr/hal-01886285

[25] A. PASHEVICH, R. STRUDEL, I. KALEVATYKH, I. LAPTEV, C. SCHMID.*Learning to Augment Synthetic Images for Sim2Real Policy Transfer*, in "IROS 2019 - IEEE/RSJ International Conference on Intelligent Robots and Systems", Macao, China, November 2019, p. 1-6, https://arxiv.org/abs/1903.07740 - 7 pages, https://hal.archives-ouvertes.fr/hal-02273326

[26] J. PEYRE, I. LAPTEV, C. SCHMID, J. SIVIC.*Detecting unseen visual relations using analogies*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, https://arxiv.org/abs/1812.05736v3 , https://hal.archives-ouvertes.fr/hal-01975760

[27] A. RUIZ, J. VERBEEK.*Adaptative Inference Cost With Convolutional Neural Mixture Models*, in "ICCV 2019 - International Conference on Computer Vision", Seoul, South Korea, October 2019, p. 1-12, https://hal.archives-ouvertes.fr/hal-02267564

[28] A. SABLAYROLLES, M. DOUZE, Y. OLLIVIER, C. SCHMID, H. JÉGOU.*White-box vs Black-box: Bayes Optimal Strategies for Membership Inference*, in "ICML 2019 - 36th International Conference on Machine Learning", Long Beach, United States, June 2019, https://arxiv.org/abs/1908.11229 , https://hal.inria.fr/hal-02278902

[29] A. SABLAYROLLES, M. DOUZE, C. SCHMID, H. JÉGOU.*Spreading vectors for similarity search*, in "ICLR 2019 - 7th International Conference on Learning Representations", New Orleans, United States, May 2019, p. 1-13, https://arxiv.org/abs/1806.03198 - Published at ICLR 2019, https://hal.inria.fr/hal-02278905

[30] V. SYDOROV, K. ALAHARI, C. SCHMID.*Focused Attention for Action Recognition*, in "BMVC 2019 - British Machine Vision Conference", Cardiff, United Kingdom, September 2019, p. 1-12, https://hal.archives-ouvertes.fr/hal-02292339

[31] M. VLADIMIROVA, J. VERBEEK, P. MESEJO, J. ARBEL.*Understanding Priors in Bayesian Neural Networks at the Unit Level*, in "ICML 2019 - 36th International Conference on Machine Learning", Long Beach, United States, Proceedings of the 36th International Conference on Machine Learning, June 2019, vol. 97, p. 6458-6467, https://arxiv.org/abs/1810.05193 - 10 pages, 5 figures, ICML'19 conference [*DOI :* 10.05193], https://hal.archives-ouvertes.fr/hal-02177151

Conferences without Proceedings

[32] A. RUIZ, O. MARTINEZ, X. BINEFA, J. VERBEEK.*Learning Disentangled Representations with Reference-Based Variational Autoencoders*, in "ICLR workshop on Learning from Limited Labeled Data", New Orleans, United States, May 2019, p. 1-17, https://hal.inria.fr/hal-01896007

Other Publications

[33] 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

[34] A. ISCEN, G. TOLIAS, Y. AVRITHIS, O. CHUM, C. SCHMID.*Graph Convolutional Networks for Learning with Few Clean and many Noisy Labels*, November 2019, https://arxiv.org/abs/1910.00324 - working paper or preprint [*DOI :* 10.00324], https://hal.inria.fr/hal-02370212

[35] A. KULUNCHAKOV, J. MAIRAL.*Estimate Sequences for Stochastic Composite Optimization: Variance Reduction, Acceleration, and Robustness to Noise*, January 2019, https://arxiv.org/abs/1901.08788 - working paper or preprint, https://hal.inria.fr/hal-01993531

[36] B. LECOUAT, J. PONCE, J. MAIRAL.*Revisiting Non Local Sparse Models for Image Restoration*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02414291

[37] X. LI, S. WANG, Y. ZHAO, J. VERBEEK, J. KANNALA.*Hierarchical Scene Coordinate Classification and Regression for Visual Localization*, November 2019, https://arxiv.org/abs/1909.06216 - working paper or preprint, https://hal.inria.fr/hal-02384675

[38] J. MAIRAL.*Cyanure: An Open-Source Toolbox for Empirical Risk Minimization for Python, C++, and soon more*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02417766

[39] G. MIALON, A. D'ASPREMONT, J. MAIRAL.*Screening Data Points in Empirical Risk Minimization via Ellipsoidal Regions and Safe Loss Functions*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02395624

[40] R. STRUDEL, A. PASHEVICH, I. KALEVATYKH, I. LAPTEV, J. SIVIC, C. SCHMID.*Learning to combine primitive skills: A step towards versatile robotic manipulation*, August 2019, https://arxiv.org/abs/1908.00722 - 11 pages, https://hal.archives-ouvertes.fr/hal-02274969

[41] G. VAROL, I. LAPTEV, C. SCHMID, A. ZISSERMAN.*Synthetic Humans for Action Recognition from Unseen Viewpoints*, January 2020, https://arxiv.org/abs/1912.04070 - working paper or preprint, https://hal.inria.fr/hal-02435731

Project-Team TRIPOP

Modeling, Simulation and Control of Nonsmooth Dynamical Systems

IN COLLABORATION WITH: Laboratoire Jean Kuntzmann (LJK)

IN PARTNERSHIP WITH:
CNRS

Institut polytechnique de Grenoble

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team TRIPOP

Creation of the Project-Team: 2019 June 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 [Team leader, Inria, Senior Researcher, HDR]
Franck Bourrier [IRSTEA, Researcher, HDR]
Bernard Brogliato [Inria, Senior Researcher, HDR]
Arnaud Tonnelier [Inria, Researcher, until Aug 2019, HDR]

Faculty Member

Guillaume James [Institut polytechnique de Grenoble, Professor, HDR]

External Collaborator

Christophe Prieur [CNRS, Senior Researcher, HDR]

Technical Staff

Maurice Brémond [Inria, Engineer]
Franck Pérignon [CNRS, Engineer]

PhD Students

Charlélie Bertrand [ENTPE, PhD Student]
Rami Sayoud [Schneider Electric, PhD Student]

Benoit Viano [Inria, PhD Student, from Nov 2019]

Post-Doctoral Fellows

Achref Elmansour [Inria, Post-Doctoral Fellow, until Aug 2019]
Mohammad Rasool Mojallizadeh [Inria, Post-Doctoral Fellow, from Jul 2019]
Alexandre Rocca [Inria, Post-Doctoral Fellow]

Visiting Scientists

Paul Armand [Univ de Limoges, until Mar 2019]
Mathias Legrand [McGill University, until Jun 2019]

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] [60], [77], [93], [62], [37].

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

$$\begin{cases} \dot{x}(t) = A\,x(t) + a + B\,\lambda(t) \\ y(t) = C\,x(t) + c + D\,\lambda(t) \\ 0 \leq y(t) \perp \lambda(t) \geq 0. \end{cases} \tag{10}$$

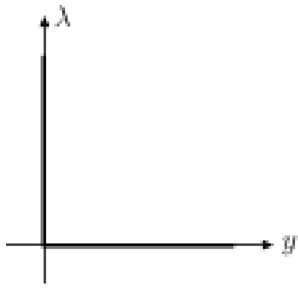


Figure 1. Complementarity condition $0 \leq y \perp \lambda \geq 0$.

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 [66]. 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 [30].

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 piece-wise 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,

$$\begin{cases} \dot{x}(t) = Ax(t) + a + B\lambda(t) \\ y(t) = Cx(t) + a \\ -\lambda(t) \in \mathrm{Sgn}(y(t)), \end{cases} \tag{24}$$

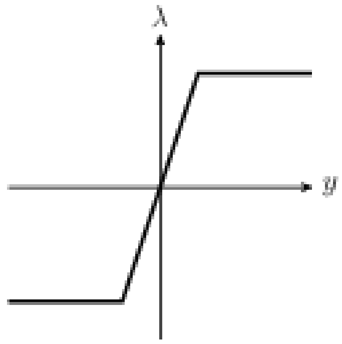
can reformulated in terms of the following set-valued system

$$\begin{cases} \dot{x}(t) = Ax(t) + a + B\lambda(t) \\ y(t) = Cx(t) + c \\ -y(t) \in N_{[-1,1]^m}(\lambda(t)). \end{cases} \tag{25}$$

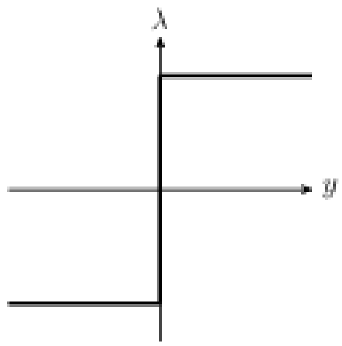
The system (21) appears in a lot of applications; among them, we can cite the sliding mode control, electrical circuits with relay and Zener diodes [26], or mechanical systems with friction [28].

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:

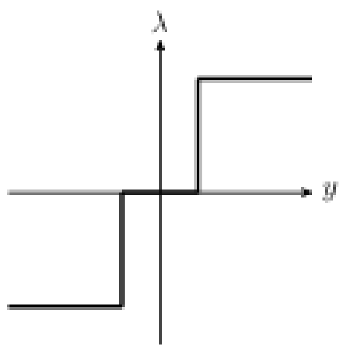
$$\dot{x}(t) = \begin{cases} Ax + a + B & \text{if} \quad Cx + c > 0 \\ Ax + a - B & \text{if} \quad Cx + c < 0. \end{cases} \tag{26}$$



Saturation



Relay



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 [60] 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 [35], 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 [86], [67], [66], complementarity theory [55], and variational inequalities theory [59].
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 [39], [65]. 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 [23], [24], [87], [88], [25], [54], [50], [89], [52]. 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 [55] 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.

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) [83]. 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 [85], [71], [72], [78]. 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 [38]. 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 [64], 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 [80], [81], [42], [90]. 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 [44], [43], [76]. 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 [44], [43]. 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 [73], [36] 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 [56], [58]. 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 [30] 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 [53] 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 [51] 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 [35]. 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 [35], [60]. 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 [63] 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 [40].
- 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 [74], 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 [25], 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 [75], [84] 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 [91]. 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 [33], [27], 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 [61] 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 [82] and the proximal point algorithm [41], 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 [29], [31], [68], [79], [47], with successful experimental validations [69], [68], [70], [92]. 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 [30].

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 [46].

- *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 [48], 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 [30], [49], 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. Domain 1

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.* 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 5(a-f). Other instances of nonsmooth dynamical systems can also be found in electrical circuits with ideal components (see Figure 5(g)) and in control theory, mainly with sliding mode control and variable structure systems (see

Figure 5(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 5(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. SICONOS

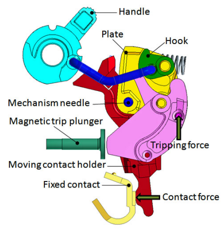
Participants: Vincent Acary, Maurice Brémond, Olivier Huber, Franck Pérignon.

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:

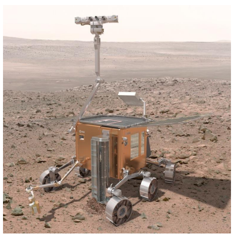
1. SICONOS/NUMERICS This library contains a set of numerical algorithms, already well identified, to solve non smooth dynamical systems. This library is written in low-level languages (C,F77) in order to ensure numerical efficiency and the use of standard libraries (Blas, Lapack, ...)
2. SICONOS/KERNEL 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.

Further informations may be found at http://siconos.gforge.inria.fr/

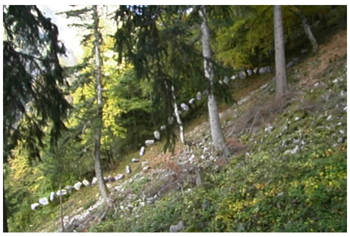


Circuit breakers mechanisms [38]

Granular flows

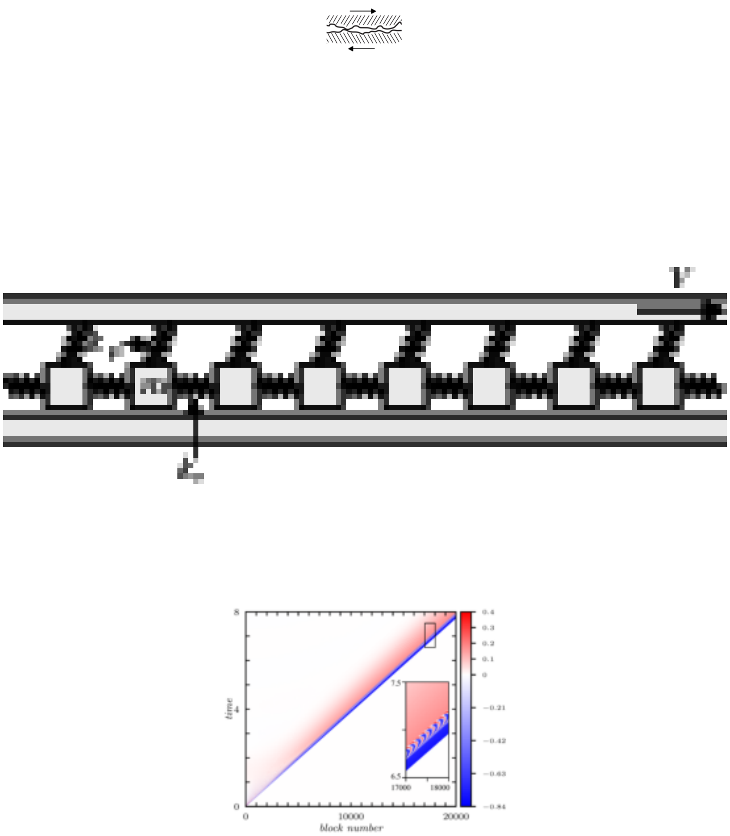


Robots (ESA ExoMars Rover [32])

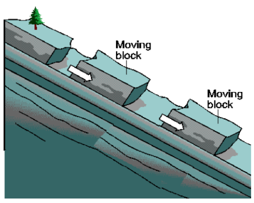


Rockfall [44], [43], [58]

Figure 3. Application fields of nonsmooth dynamics (mechanics)

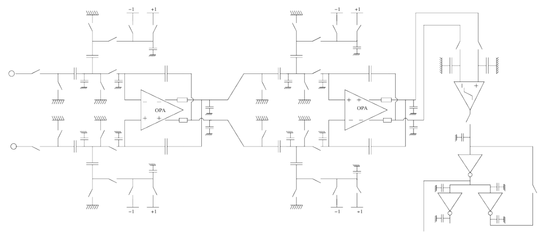


Frictional interface and solitary waves in the Burridge-Knopoff model [80]

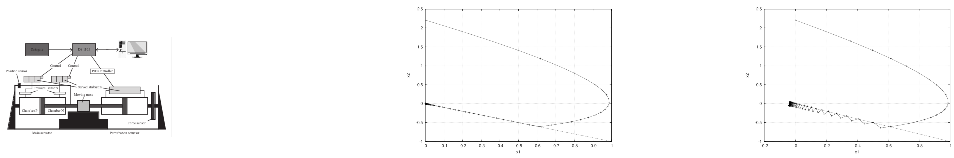


Sliding blocks

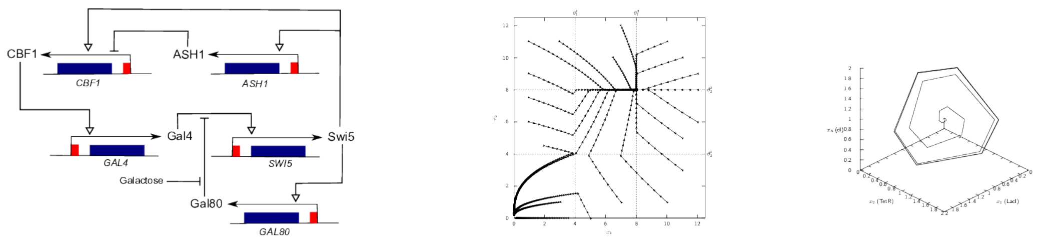
Figure 4. Application fields of nonsmooth dynamics (waves)



Switched electrical circuits (delta-sigma converter) [26]



Sliding mode control [29], [31], [69], [70], [79]



Gene regulatory networks [35]

Figure 5. Application fields of nonsmooth dynamics

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 compression waves (solitary waves or fronts) or localized 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 [8], 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-Lipschitz 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 addition, G.J. has developed a theoretical method to analyze impacts in homogeneous granular chains with KK dissipation. The idea is to use the exponent α of the contact force as a parameter and derive simpler dynamical equations through an asymptotic analysis, in the limit when α approaches unity and long waves are considered. In that case, different continuum limits of the granular chain can be obtained. When the contact damping constant remains of order unity, wave profiles are well approximated by solutions of a viscous Burgers equation with logarithmic nonlinearity. For small contact damping, dispersive effects must be included and the continuum limit corresponds to a KdV-Burgers equation with logarithmic nonlinearity. By studying traveling wave solutions to these partial differential equations, we obtain analytical approximations of wave profiles such as compression fronts. We observe that these approximations remain meaningful for the classical exponent $\alpha = 3/2$. Indeed, they are close to exact wave profiles computed numerically for the KK model, using both dynamical simulations (response of the chain to a compression by a piston) and the Newton method (computation of exact traveling waves by a shooting method). In addition, in analogy with the Rankine-Hugoniot conditions for hyperbolic systems, we relate the asymptotic states of the KK model (for an infinite granular chain) to the velocity of a propagating front. These results are described in an article in preparation.

6.2. 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.

6.3. Hybrid Differential Algebraic equations

Participants: Vincent Acary, Bernard Brogliato, Alexandre Rocca.

In [18], [21], we study differential algebraic equations with constraints defined in a piecewise manner using a conditional statement. Such models classically appear in systems where constraints can evolve in a very small time frame compared to the observed time scale. The use of conditional statements or hybrid automata are a powerful way to describe such systems and are, in general, well suited to simulation with event driven numerical schemes. However, such methods are often subject to chattering at mode switch in presence of sliding modes, or can result in Zeno behaviours. In contrast, the representation of such systems using differential inclusions and method from non-smooth dynamics are often closer to the physical theory but may be harder to interpret. Associated time-stepping numerical methods have been extensively used in mechanical modelling with success and then extended to other fields such as electronics and system biology. In a similar manner to the previous application of non-smooth methods to the simulation of piecewise linear ODEs, non-smooth event-capturing numerical scheme are applied to piecewise linear DAEs. In particular, the study of a 2-D dynamical system of index-2 with a switching constraint using set-valued operators, is presented.

6.4. 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 [57].

6.4.1. Numerical solvers for frictional contact problems.

Participants: Vincent Acary, Maurice Brémond, Paul Armand.

In [34], 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.

Recently, new developments have been carried out on two new applications of well-known numerical methods in Optimization:

- *Interior point methods* With the visit of Paul Armand, Université de Limoges, we co-supervise a M2 internship, Maksym Shpakovych on the application of interior point methods for quadratic problem with second-order cone constraints. The results are encouraging and a publication in computational mechanics is in progress.
- *Alternating Direction Method of Multipliers*. In collaboration with Yoshihiro Kanno, University of Tokyo, the use of the Alternating Direction Method of Multipliers (ADMM) has been adapted to the discrete frictional contact problems. With the help of some acceleration and restart techniques for first-order optimization methods and a residual balancing technique for adapting the proximal penalty parameter, the method proved to be efficient and robust on our test bench examples. A publication is also in preparation on this subject.

6.4.2. Modeling and numerical methods for frictional contact problems with rolling resistance

Participants: Vincent Acary, Franck Bourrier.

In [19], the Coulomb friction model is enriched to take into account the resistance to rolling, also known as rolling friction. Introducing the rolling friction cone, an extended Coulomb's cone and its dual, a formulation of the Coulomb friction with rolling resistance as a cone complementarity problem is shown to be equivalent to the standard formulation of the Coulomb friction with rolling resistance. Based on this complementarity formulation, the maximum dissipation principle and the bi-potential function are derived. Several iterative numerical methods based on projected fixed point iterations for variational inequalities and block-splitting techniques are given. The efficiency of these method strongly relies on the computation of the projection onto the rolling friction cone. In this article, an original closed-form formulae for the projection on the rolling friction cone is derived. The abilities of the model and the numerical methods are illustrated on the examples of a single sphere sliding and rolling on a plane, and of the evolution of spheres piles under gravity.

6.4.3. Finite element modeling of cable structures

Participants: Vincent Acary, Charlélie Bertrand.

Standard finite element discretization for cable structures suffer from several drawbacks. The first one is related to the mechanical assumption that the cable can not support compression. Standard formulations do not take into account this assumption. The second drawback comes from the high stiffness of the cable model when we deal with large lengths with high Young modulus such as cable ropeways installations. In this context, standard finite element applications cannot avoid compressive solutions and have huge difficulties to converge. In a forthcoming paper, we propose to a formulation based on a piecewise linear modeling of the cable constitutive behavior where the elasticity in compression is canceled. Furthermore, a dimensional analysis help us to formulate a problem that is well-balanced and the conditioning of the problem is diminished. The finite element discretization of this problem yields a robust method where convergence is observed with the number of elements and the nonlinear solver based on nonsmooth Newton strategy is converging up to tight tolerances. The convergence with the number of element allows one to refine the mesh as much as we want that will be of utmost importance for applications with contact and friction. Indeed, a fine discretization with respect to the whole length of the cable will be possible in the contact zone.

6.4.4. Well-posedness of the contact problem

We continue in [3] the analysis of the so-called contact problem for Lagrangian systems with bilateral and unilateral constraints, with set-valued Coulomb's friction. The problem that is analysed this time concerns sticking contacts (in both the normal and the tangential directions), *i.e.*, does there exist a solution (possibly unique) to the contact problem (that takes the form of a complementarity problem) when all contacts are sticking ? An algorithm is proposed that allows in principle to compute solutions. We rely strongly on results of existence and uniqueness of solutions to variational inequality of the second kind, obtained in the team some years ago. Let us note also the erratum/addendum of the monograph [45] in [17], which is regularly updated.

6.5. Analysis and Control of Set-Valued Systems

Participants: Bernard Brogliato, Christophe Prieur, Vincent Acary.

6.5.1. Robust sliding-mode control: continuous and discrete-time

The implicit method for the time-discretization of set-valued sliding-mode controllers was introduced in [29], [31]. The backstepping approach is used in [9] to design a continuous-time and a discrete-time nested set-valued controller that is able to reject unmatched disturbances (a problem that is known to be tough in the sliding-mdoe control community). In [13], [10] we continue 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). The discrete-time analysis of the twisting and the super-twisting algorithms are tackled in [7], [4].

6.5.2. Analysis of set-valued Lur'e dynamical systems

Lur'e systems are very popular in the Automatic Control field since their introduction by Lur'e in 1944. In [5] we propose a very complete survey/tutorial on the set-valued version of such dynamical systems (in finite dimension) which mainly consist of the negative feedback interconnection of an ODE with a maximal monotone set-valued operator. The first studies can be traced back to Yakubpovich in 1963 who analysed the stability of a linear time invariant system with positive real constraints, in negative feedback connection with a hysteresis operator. About 600 references are analysed from the point of view of the mathematical formalisms (Moreau's sweeping process, evolution variational inequalities, projected dynamical systems, complementarity dynamical systems, maximal monotone differential inclusions, differential variational inequalities), the relationships between these formalisms, the numerous fields of application, the well-posedness issues (existence, uniqueness and continuous dependence of solutions), and the stability issues (generalized equations for fixed points, Lyapunov stability, invariance principles).

6.5.3. Optimal control of LCS

The quadratic and minimum time optimal control of LCS as in (6) is tackled in [14], [12]. 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).

6.6. Dissipative systems

Participant: Bernard Brogliato.

The third edition of the book Dissipative Systems Analysis and Control has been released https://www.springer.com/gp/book/9783030194192. Also a short proof of equivalence of side conditions for strictly positive real (SPR) transfer functions is done in [6], closing a long debate in the Automatic Control community about the characterization of SPR transfer matrices.

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. Regional Initiatives

The SMART PROTECT project (2019–2022) is a R&D booster project granted by the Région Auvergne Rhône–Alpes. The project is coordinated by Géolithe Innov, a French company specialized in the innovation in Geotechnics. The partners are Géolithe, Irstea and Myotis. The aim of the project is to design and validate a new type of protection structures against rockfall and avalanches. The role of the TRIPOP team is to propose a numerical modeling of the structure and to improve the link between simulations and wireless sensors, which will equip the structure.

8.2. National Initiatives

8.2.1. ANR project Digitslid

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).

8.2.2. 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.

8.2.3. 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.

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

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.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Selection

9.1.1.1. Member of the Conference Program Committees

1. Vincent Acary is co-organisor (with O. Bruls and R. Leine) of the mini-symposium Nonsmooth Dynamics, 10th Nonlinear Dynamics Conference (ENOC 2020), Lyon, July 2020.
2. Vincent Acary was co-organisor (with O. Bruls and R. Leine) of the Eighth Symposium of the European Network for Nonsmooth Dynamics at Inria Grenoble Rhone–Alpes. 17–18 September 2019 http://ennsd.gforge.inria.fr/eighthsymposium.html
3. Bernard Brogliato is member of the National Organization Committee of the 10th Euromech Nonlinear Dynamics Conference (ENOC 2020), Lyon, 5-10 July 2020. https://enoc2020.sciencesconf.org/resource/page/id/3

4. Bernard Brogliato is co-organisor (with N. van de Wouw and A. Pavlov) of the mini-symposium Control and Synchronization in Nonlinear Systems, 10th Nonlinear Dynamics Conference (ENOC 2020), Lyon, July 2020. https://enoc2020.sciencesconf.org/resource/page/id/8
5. Bernard Brogliato was co-organisor (with M. Reichhartinger and A. Polyakov) of the invited session Approximation-Methods, Discrete-time Equivalents and Analysis of Sliding Mode Control Systems, conférence NOLCOS 2019, Wien.

9.1.1.2. Reviewer

1. Vincent Acary was reviewer for IFAC World Congress, American Control Conference.
2. Bernard Brogliato was reviewer for IFAC NOLCOS, IFAC World Congress, IEEE International Conference on Decision and Control.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

1. Bernard Brogliato was Associate Editor for ASME Journal of Nonlinear and Computational Dynamics (ending December 2019).
2. Bernard Brogliato was Associate Editor for Nonoinear Analysis: Hybrid Systems (ending December 2019).

9.1.2.2. Reviewer - Reviewing Activities

1. Vincent Acary was reviewer for IEEE Transactions on Automatic Control, Multibody Systems Dynamics, Applied Mathematics and Computation, ASME Journal of Nonlinear and Computational Dynamics, Nonlinear Dynamics, International Journal for Numerical Methods in Engineering, SIAM Journal on Control and Optimization.
2. Bernard Brogliato is a regular reviewer for Automatica, IEEE Transactions on Automatic Control, SIAM Journal on Control and Optimization, Multibody Systems Dynamics, European Journal of Mechanics A/Solids, Set-Valued and Variational Analysis, *etc*.
3. Franck Bourrier was reviewer for International Journal for Numerical and Analytical Methods in Geomechanics, International Journal of Solids and Structures, Computers and Geotechnics, Soils and Foundation, Rock Mechanics and Rock Engineering, European Journal of Environmental and Civil Engineering, Landslides.
4. Guillaume James was reviewer for Multibody System Dynamics, Communications in Mathematical Physics, Physica D.

9.1.3. Invited Talks

1. Vincent Acary gave a keynote lecture at Multibody Dynamics Workshop 2019 2nd International Multibody Summer School, 20-24 May 2019, Parma, Italy http://www.multibodysummerschool.eu/index.html
2. Bernard Brogliato gave a plenary talk at the workshop Control of State-Constrained Dynamical Systems in Valparaiso, September 2019 (http://coscds2019.mat.utfsm.cl/).
3. Guillaume James gave a talk at the conference Equadiff 2019, July 2019, Leiden, in the minisymposium *Coherent structures of nonlinear evolution and lattice equations*.

9.1.4. Leadership within the Scientific Community

- V. Acary is coordinator with R. Leine of the Europe Network for Nonsmooth Dynamics http://ennsd.gforge.inria.fr/.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

- Licence : G. James, Introduction to Numerical Methods, 31 hETD, L3, Grenoble INP - Pagora (1st year).
- Licence : G. James, Normed Vector Spaces, 26 hETD, L2, Prépa INP, Grenoble.
- Master : G. James, Numerical Methods, 91 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 : Franck Bourrier, 5H éq TD Modélisation des chutes de blocs, Master GAIA, Université Savoie Mont-Blanc.

9.2.2. Supervision

- PhD in progress : Rami Sayoud, Analyse vibratoire des armoires électriques, January 2018, université Grenoble Alpes, Vincent Acary and Bernard Brogliato
- PhD in progress : Benoit Viano, Nonsmooth modelling of impacted elastoplastic beams, November 2019, université Grenoble Alpes, Vincent Acary and Franck Bourrier.
- PhD in progress : Charlélie Bertrand, September 2018, Mechanical model for cable vibrations, ENTPE, Claude Lamarque and Vincent Acary.
- PhD in progress : Christelle Kozaily. Structural analysis for multi-mode DAE systems, Octobre 2018, V. Acary and B. Caillaud.
- PhD in progress : Vivien Cros, Analyse de la réponse vibratoire d'arbres sous sollicitations dynamiques, November 2018, université Grenoble Alpes, Franck Bourrier.

9.2.3. Juries

1. Bernard Brogliato was referee for the Ph.D. thesis of Arindam Bhattacharjee (April 2019), IIT Kanpur, India, (supervisor Prof. Anindya Chaterjee, Mechanical Engineering Dept. IIT Kanpur), *New approximations in vibroimpact problems*.

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] B. Brogliato, J. Kövecses, V. Acary.*The contact problem in Lagrangian systems with redundant frictional bilateral and unilateral constraints and singular mass matrix. The all-sticking contacts problem*, in "Multibody System Dynamics", 2020, vol. 48, n^{o} 2, p. 151-192 [*DOI :* 10.1007/s11044-019-09712-1], https://hal.inria.fr/hal-02315547

[4] B. Brogliato, A. Polyakov, D. Efimov.*The implicit discretization of the super-twisting sliding-mode control algorithm*, in "IEEE Transactions on Automatic Control", 2019, p. 1-8, forthcoming [*DOI :* 10.1109/TAC.2019.2953091], https://hal.inria.fr/hal-02336599

[5] B. BROGLIATO, A. TANWANI.*Dynamical systems coupled with monotone set-valued operators: Formalisms, applications, well-posedness, and stability*, in "SIAM Review", 2019, p. 1-125, forthcoming, https://hal.inria.fr/hal-02379498

[6] A. FERRANTE, A. LANZON, B. BROGLIATO.*A direct proof of the equivalence of side conditions for strictly positive real matrix transfer functions*, in "IEEE Transactions on Automatic Control", 2020, vol. 65, n^{o} 1, p. 450-452 [*DOI :* 10.1109/TAC.2019.2918123], https://hal.inria.fr/hal-01947938

[7] O. HUBER, V. ACARY, B. BROGLIATO.*Lyapunov stability analysis of the implicit discrete-time twisting control algorithm*, in "IEEE Transactions on Automatic Control", 2019, p. 1-8, forthcoming [*DOI :* 10.1109/TAC.2019.2940323], https://hal.inria.fr/hal-01622092

[8] G. JAMES, K. VOROTNIKOV, B. BROGLIATO.*Kuwabara-Kono numerical dissipation: a new method to simulate granular matter*, in "IMA Journal of Applied Mathematics", 2019, p. 1-29, forthcoming, https://hal.archives-ouvertes.fr/hal-01878973

[9] F. MIRANDA-VILLATORO, F. CASTAÑOS, B. BROGLIATO.*Continuous and discrete-time stability of a robust set-valued nested controller*, in "Automatica", September 2019, vol. 107, p. 406-417 [*DOI :* 10.1016/J.AUTOMATICA.2019.06.003], https://hal.inria.fr/hal-01944920

[10] A. POLYAKOV, D. EFIMOV, B. BROGLIATO.*Consistent Discretization of Finite-time and Fixed-time Stable Systems*, in "SIAM Journal on Control and Optimization", 2019, vol. 57, n^{o} 1, p. 78-103 [*DOI :* 10.1137/18M1197345], https://hal.inria.fr/hal-01838712

[11] A. TONNELIER.*Signal propagations along excitable chains*, in "SIAM Journal on Applied Dynamical Systems", 2019, vol. 18, n^{o} 3, p. 1391-1419 [*DOI :* 10.1137/18M1234229], https://hal.archives-ouvertes.fr/hal-02180588

[12] A. VIEIRA, B. BROGLIATO, C. PRIEUR.*Quadratic Optimal Control of Linear Complementarity Systems: First order necessary conditions and numerical analysis*, in "IEEE Transactions on Automatic Control", 2020, p. 1-8, forthcoming [*DOI :* 10.1109/TAC.2019.2945878], https://hal.inria.fr/hal-01690400

International Conferences with Proceedings

[13] A. POLYAKOV, D. EFIMOV, B. BROGLIATO, M. REICHHARTINGER.*Consistent Discretization of Locally Homogeneous Finite-time Stable Control Systems*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, IEEE, June 2019, p. 1616-1621 [*DOI :* 10.23919/ECC.2019.8795633], https://hal.inria.fr/hal-02069717

[14] A. VIEIRA, B. BROGLIATO, C. PRIEUR.*Optimality conditions for the minimal time problem for Complementarity Systems*, in "Joint 8th IFAC Symposium on Mechatronic Systems (MECHATRONICS'19) and 11th IFAC Symposium on Nonlinear Control Systems (NOLCOS'19)", Vienne, Austria, IFAC, September 2019, p. 325-330, https://hal.inria.fr/hal-01856054

Scientific Books (or Scientific Book chapters)

[15] B. BROGLIATO, R. LOZANO, B. MASCHKE, O. EGELAND.*Dissipative Systems Analysis and Control : Theory and Application*, Communication and Control Engineering, Springer International Publishing, 2020 [*DOI :* 10.1007/978-3-030-19420-8], https://hal.archives-ouvertes.fr/hal-02407669

Research Reports

[16] V. ACARY, O. BONNEFON, M. BRÉMOND, O. HUBER, F. PÉRIGNON, S. SINCLAIR.*An introduction to Siconos*, Inria, November 2019, n^o RT-0340, 97, https://hal.inria.fr/inria-00162911

[17] B. BROGLIATO.*Nonsmooth Mechanics. Models, Dynamics and Control : Erratum/Addendum*, Inria Grenoble - Rhone-Alpes, September 2019, p. 1-15, https://hal.inria.fr/hal-01331565

[18] A. ROCCA, V. ACARY, B. BROGLIATO.*Index-2 hybrid DAE: a case study with well-posedness and numerical analysis*, Inria - Research Centre Grenoble – Rhône-Alpes, November 2019, https://hal.inria.fr/hal-02381489

Other Publications

[19] V. ACARY, F. BOURRIER.*Coulomb friction with rolling resistance as a cone complementarity problem*, November 2019, working paper or preprint, https://hal.inria.fr/hal-02382213

[20] C. BERTRAND, V. ACARY, C. H. LAMARQUE, A. TURE SAVADKOOHI.*A robust and efficient numerical finite element method for cables*, January 2020, working paper or preprint, https://hal.inria.fr/hal-02439982

[21] A. ROCCA, V. ACARY, B. BROGLIATO.*Index-2 hybrid DAE: a case study with well-posedness and numerical analysis*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02391311

[22] V. N. VO.*Trajectory tracking design for linear complementarity systems with continuous solutions*, university of Limoges, France, July 2019, INTERNSHIP REPORT for the degree of Master in Applied Mathematics, https://hal.inria.fr/hal-02267750

References in notes

[23] 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

[24] 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

[25] 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

[26] 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

[27] 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

[28] 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

[29] 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

[30] 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

[31] 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

[32] 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

[33] 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

[34] 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", S. I. PUBLISHING (editor), June 2018, p. 375-457 [*DOI :* 10.1007/978-3-319-75972-2_10], https://hal.inria.fr/hal-01878539

[35] 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

[36] 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/

[37] S. ADLY.*A Variational Approach to Nonsmooth Dynamics. Applications in Unilateral Mechanics and Electronics*, SpringerBriefs in Mathematics, Springer Verlag, 2017

[38] 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

[39] 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

[40] 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

[41] 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 [*DOI :* 10.1137/080716542]

[42] 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

[43] 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

[44] 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

[45] 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

[46] 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

[47] 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

[48] 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

[49] 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

[50] 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

[51] B. CAILLAUD.*Hybrid vs. nonsmooth dynamical systems*, 2014, synchron, http://synchron2014.inria.fr/wp-content/uploads/sites/13/2014/12/Caillaud-nsds.pdf

[52] 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

[53] 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

[54] 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

[55] R. W. COTTLE, J. PANG, R. E. STONE.*The Linear Complementarity Problem*, Academic Press, Inc., Boston, MA, 1992

[56] 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^{o} 6, p. 128-141, https://doi.org/10.3188/szf.2007.0128

[57] F. DUBOIS, V. ACARY, M. JEAN.*The Contact Dynamics method: A nonsmooth story* , in "Comptes Rendus Mécanique", March 2018, vol. 346, n^{o} 3, p. 247-262 [*DOI :* 10.1016/J.CRME.2017.12.009], https://hal.archives-ouvertes.fr/hal-01676287

[58] 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

[59] 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

[60] A. F. FILIPPOV.*Differential Equations with Discontinuous Right Hand Sides*, Kluwer, Dordrecht, the Netherlands, 1988

[61] M. P. FRIEDLANDER, D. ORBAN.*A primal–dual regularized interior-point method for convex quadratic programs*, in "Mathematical Programming Computation", 2012, vol. 4, n^{o} 1, p. 71–107, http://dx.doi.org/10.1007/s12532-012-0035-2

[62] D. GOELEVEN.*Complementarity and Variational Inequalities in Electronics*, Academic Press, 2017

[63] 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

[64] F. GÉNOT, B. BROGLIATO.*New results on Painlevé Paradoxes*, in "European Journal of Mechanics - A. solids", 1999, vol. 18, p. 653-677

[65] 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

[66] J. HIRIART-URRUTY, C. LEMARÉCHAL.*Fundamentals of Convex Analysis*, Springer Verlag, 2001

[67] J. HIRIART-URRUTY, C. LEMARÉCHAL.*Convex Analysis and Minimization Algorithms*, Springer Verlag, Berlin, 1993, vol. I and II

[68] 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^{o} 10, p. 3016-3030, In press

[69] 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, p. 129–141

[70] 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

[71] G. JAMES.*Periodic travelling waves and compactons in granular chains*, in "J. Nonlinear Sci.", 2012, vol. 22, p. 813-848

[72] G. JAMES, P. G. KEVREKIDIS, J. CUEVAS.*Breathers in oscillator chains with Hertzian interactions*, in "Physica D", 2013, vol. 251, p. 39–59

[73] 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

[74] 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

[75] R. LEINE.*Bifurcation in Discontinuous Mechanical Systems of Filippov-Type*, Technische Universiteit Eindhoven, 2000

[76] 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

[77] 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

[78] L. LIU, G. JAMES, P. KEVREKIDIS, A. VAINCHTEIN.*Nonlinear waves in a strongly resonant granular chain*, in "Nonlinearity", 2016, p. 3496-3527

[79] F. A. MIRANDA-VILLATORO, B. BROGLIATO, F. CASTANOS.*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

[80] 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

[81] J. E. 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

[82] 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° 2, p. 372–376

[83] 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.

[84] F. Z. NQI, M. SCHATZMAN.*Computation of Lyapunov Exponents for dynamical system with impact*, in "Applied Mathematial Sciences", 2010, vol. 4, n° 5, p. 237–252

[85] M. PORTER, P. KEVREKIDIS, C. DARAIO.*Granular crystals: Nonlinear dynamics meets materials engineering*, in "Physics Today", 2015, vol. 68, n° 44

[86] R. ROCKAFELLAR.*Convex Analysis*, Princeton University Press, 1970

[87] 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

[88] 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° 15, p. 250–276

[89] 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

[90] A. TONNELIER.*McKean caricature of the FitzHugh-Nagumo model: traveling pulses in a discrete diffusive medium*, in "Phys. Rev. E", 2003, vol. 67, n° 036105

[91] A. VIEIRA, B. BROGLIATO, C. PRIEUR.*Optimal control of linear complementarity systems*, in "IFAC World Congress on Automatic Control", Toulouse, France, 2017

[92] 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° 5, p. 2071–2075

[93] 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 Claude Bernard, Professor, HDR]
Nils Gesbert [Institut polytechnique de Grenoble, Associate Professor]
Cécile Roisin [Univ Grenoble Alpes, Professor, HDR]

Technical Staff
Thomas Calmant [Inria, Engineer]

PhD Students
Fateh Boulmaiz [Qualifret SAS, PhD Student]
Sarah Chlyah [Inria, PhD Student]
Amela Fejza [Univ Grenoble Alpes, PhD Student]
Muideen Lawal [Univ Grenoble Alpes, PhD Student]

Administrative Assistant
Helen Pouchot-Rouge-Blanc [Inria, Administrative Assistant]

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.

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: Application to Graph Queries

Graph databases have received a lot of attention as they are particularly useful in many applications such as social networks, life sciences and the semantic web. Various languages have emerged to query graph databases, many of which embed forms of recursion which reveal essential for navigating in graphs. The relational model has benefited from a huge body of research in the last half century and that is why many graph databases rely on techniques of relational 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. We propose μ-RA, a variation of the Relational Algebra equipped with a fixpoint operator for expressing recursive relational queries. μ-RA can notably express unions of conjunctive regular path queries. Leveraging the fact that this fixpoint operator makes recursive terms more amenable to algebraic transformations, we propose new rewrite rules. These rules make it possible to generate new query execution plans, that cannot be obtained with previous approaches. We have defined the syntax and semantics of μ-RA, together with the rewriting rules that we specifically devised to tackle the optimization of recursive queries. We have also conducted practical experiments that show that the newly generated plans can provide significant performance improvements for evaluating recursive queries over graphs.

These results will be presented at the SIGMOD 2020 conference [9].

6.2. An Algebra with a Fixpoint Operator for Distributed Data Collections

We propose an algebra with a fixpoint operator which is suitable for modeling recursive computations with distributed data collections. We show that under reasonable conditions this fixpoint can be evaluated by parallel loops with one final merge rather than by a global loop requiring network overhead after each iteration. We also propose rewrite rules, showing when and how filters can be pushed through recursive terms, and how to filter inside a fixpoint before a join. Experiments with the Spark platform illustrate performance gains brought by these systematic optimizations [10].

6.3. 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 only describe the subtree structure of the given node. In 2015, Giuseppe Castagna and our team independently proposed 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. Recently, 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.

6.4. 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 are developping distributed models that are scalable and interpretable. Key insights include analyzing 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.

Preliminary results were presented at the 2018 International Conference on Data Science and Applications, and extended results have been submitted for publication consideration.

6.5. What can millions of laboratory test results tell us about the temporal aspect of data quality? Study of data spanning 17 years in a clinical data warehouse.

In this work, our objective is to identify common temporal evolution profiles in biological data and to propose a semi-automated method to these patterns in a clinical data warehouse (CDW). We leveraged the CDW of the European Hospital Georges Pompidou and tracked the evolution of 192 biological parameters over a period of 17 years (for 445,0 0 0 + patients, and 131 million laboratory test results). We have identified three common profiles of evolution: discretization, breakpoints, and trends. We developed computational and statistical methods to identify these profiles in the CDW. Overall, of the 192 observed biological parameters (87,814,136 values), 135 presented at least one evolution. We identified breakpoints in 30 distinct parameters, discretizations in 32, and trends in 79. As a conclusion, we can say that our method allows the identification of several temporal events in the data. Considering the distribution over time of these events, we identified probable causes for the observed profiles: instruments or software upgrades and changes in computation formulas. We evaluated the potential impact for data reuse. Finally, we formulated recommendations to enable safe use and sharing of biological data collection to limit the impact of data evolution in retrospective and federated studies (e.g. the annotation of laboratory parameters presenting breakpoints or trends) [4].

6.6. Interactive Mapping Specification with Exemplar Tuples

While schema mapping specification is a cumbersome task for data curation specialists, it becomes unfeasible for non-expert users, who are unacquainted with the semantics and languages of the involved transformations.

In this work, we propose an interactive framework for schema mapping specification suited for non-expert users. The underlying key intuition is to leverage a few exemplar tuples to infer the underlying mappings and iterate the inference process via simple user interactions under the form of Boolean queries on the validity of the initial exemplar tuples. The approaches available so far are mainly assuming pairs of complete universal data examples, which can be solely provided by data curation experts, or are limited to poorly expressive mappings.

We present a quasi-lattice-based exploration of the space of all possible mappings that satisfy arbitrary user exemplar tuples. Along the exploration, we challenge the user to retain the mappings that fit the user's requirements at best and to dynamically prune the exploration space, thus reducing the number of user interactions. We prove that after the refinement process, the obtained mappings are correct and complete. We present an extensive experimental analysis devoted to measure the feasibility of our interactive mapping strategies and the inherent quality of the obtained mappings [2].

6.7. Schema Validation and Evolution for Graph Databases

Despite the maturity of commercial graph databases, little consensus has been reached so far on the standardization of data definition languages (DDLs) for property graphs (PG). Discussion on the characteristics of PG schemas is ongoing in many standardization and community groups. Although some basic aspects of a schema are already present in most commercial graph databases, full support is missing allowing to constraint property graphs with more or less flexibility. In this work, we show how schema validation can be enforced through homomorphisms between PG schemas and PG instances by leveraging a concise schema DDL inspired by Cypher syntax. We also briefly discuss PG schema evolution that relies on graph rewriting operations allowing to consider both prescriptive and descriptive schemas [6].

6.8. MapRepair: Mapping and Repairing under Policy Views

Mapping design is overwhelming for end users, who have to check at par the correctness of the mappings and the possible information disclosure over the exported source instance. In our tool MapRepair, we focus on the latter problem by proposing a novel practical solution to ensure that a mapping faithfully complies with a set of privacy restrictions specified as source policy views. We showcase MapRepair, that guides the user through the tasks of visualizing the results of the data exchange process with and without the privacy restrictions. MapRepair leverages formal privacy guarantees and is inherently data-independent, i.e. if a set of criteria are satisfied by the mapping statement, then it guarantees that both the mapping and the underlying instances do not leak sensitive information. Furthermore, MapRepair also allows to automatically repair an input mapping w.r.t. a set of policy views in case of information leakage. We build on various demonstration scenarios, including synthetic and real-world instances and mappings [5].

6.9. Approximate Querying on Property Graphs

Property graphs are becoming widespread when modeling data with complex structural characteristics and enhancing edges and nodes with a list of properties. We worked on the approximate evaluation of counting queries involving recursive paths on property graphs. As such queries are already difficult to evaluate over pure RDF graphs, they require an ad-hoc graph summary for their approximate evaluation on property graphs. We prove the intractability of the optimal graph summarization problem, under our algorithm's conditions. We design and implement a novel property graph summary suitable for the above queries, along with an approximate query evaluation module. Finally, we show the compactness of the obtained summaries as well as the accuracy of answering counting recursive queries on them [8].

6.10. RDF Graph Anonymization Robust to Data Linkage

Privacy is a major concern when publishing new datasets in the context of Linked Open Data (LOD). A new dataset published in the LOD is indeed exposed to privacy breaches due to the linkage to objects already present in the other datasets of the LOD. In this work, we focus on the problem of building safe anonymizations of an RDF graph to guarantee that linking the anonymized graph with any external RDF graph will not cause privacy breaches. Given a set of privacy queries as input, we study the data-independent safety problem and the sequence of anonymization operations necessary to enforce it. We provide sufficient conditions under which an anonymization instance is safe given a set of privacy queries. Additionally, we show that our algorithms for RDF data anonymization are robust in the presence of sameAs links that can be explicit or inferred by additional knowledge.

6.11. Navigating the Maze of Wikidata Query Logs

We propose an in-depth and diversified analysis of the Wikidata query logs, recently made publicly available. Although the usage of Wikidata queries has been the object of recent studies, our analysis of the query traffic reveals interesting and unforeseen findings concerning the usage, types of recursion, and the shape classification of complex recursive queries. Wikidata specific features combined with recursion let us identify a significant subset of the entire corpus that can be used by the community for further assessment. We consider and analyze the queries across many different dimensions, such as the robotic and organic queries, the presence/absence of constants along with the correctly executed and timed out queries. A further investigation that we pursue is to find, given a query, a number of queries structurally similar to the given query. We provide a thorough characterization of the queries in terms of their expressive power, their topological structure and shape, along with a deeper understanding of the usage of recursion in these logs. We make the code for the analysis available as open source [7].

6.12. Graph Generators: State of the Art and Open Challenges

The abundance of interconnected data has fueled the design and implementation of graph generators reproducing real-world linking properties, or gauging the effectiveness of graph algorithms, techniques and applications manipulating these data. We consider graph generation across multiple subfields, such as Semantic Web, graph databases, social networks, and community detection, along with general graphs. Despite the disparate requirements of modern graph generators throughout these communities, we analyze them under a common umbrella, reaching out the functionalities, the practical usage, and their supported operations. We argue that this classification is serving the need of providing scientists, researchers and practitioners with the right data generator at hand for their work. This survey provides a comprehensive overview of the state-of-the-art graph generators by focusing on those that are pertinent and suitable for several data-intensive tasks. Finally, we discuss open challenges and missing requirements of current graph generators along with their future extensions to new emerging fields [3].

6.13. A trichotomy for regular simple path queries on graphs

We focus on the computational complexity of regular simple path queries (RSPQs). We consider the following problem RSPQ(L) for a regular language L: given an edge-labeled digraph Gand two nodes xand y, is there a simple path from x to y that forms a word belonging to L? We fully characterize the frontier between tractability and intractability for RSPQ(L). More precisely, we prove RSPQ(L)is either AC0, NL-complete or NP-complete depending on the language L. We also provide a simple characterization of the tractable fragment in terms of regular expressions. Finally, we also discuss the complexity of deciding whether a language L belongs to the fragment above. We consider several alternative representations of L: DFAs, NFAs or regular expressions, and prove that this problem is NL-complete for the first representation and PSpace-complete for the other two [1].

7. Partnerships and Cooperations

7.1. Regional Initiatives

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.

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific Events: Organisation

8.1.1.1. General Chair, Scientific Chair

- Angela Bonifati is program chair of EDBT 2020, and co-chair of the Workshops of SIGMOD 2019.

8.1.1.2. 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), and co-chair and organizer of the EDBT 2019 summer school.

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 IJCAI'19, AAAI'20, IJCAI-PRICAI'20.
- A. Bonifati has been program committee member of VLDB 2019, PODS 2019, AAAI 2019, ICDE 2019, EDBT 2019, SIGMOD 2019, DEBS 2019, ICDT 2020.

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.

8.1.4. Scientific Expertise

- P. Genevès has been a scientific expert at ANRT for the CIFRE funding process.

8.1.5. Research Administration

- P. Genevès is responsible for the Computer Science Specialty at the Doctoral School MSTII (ED 217)
- C. Roisin is a member of the CNU (Conseil National des Universités).
- C. Roisin is a member of the Inria Grenoble Inria-Hub committee.

- C. Roisin has been president of a Committee of Selection for an assistant position at university Grenoble-Alpes.
- 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-Rhône-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
- Master : N. Gesbert, 'Principes des systèmes de gestion de bases de données', 42 h eq TD, M1, 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
- Master : N. Gesbert, 'Introduction to lambda-calculus', 5 h eq TD, M2, UGA-Grenoble INP (MOSIG)
- 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 M2-level course 'Semantic Web: from XML to OWL' of the MOSIG program at UGA (36h)
- P. Genevès is responsible and teacher in 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 in progress: Muideen Lawal, Cost models for optimizing compilers based on mu-terms, PhD started in October 2017, co-supervised by Pierre Genevès and Nabil Layaïda.
- 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.

9. Bibliography

Publications of the year

Articles in International Peer-Reviewed Journal

[1] G. BAGAN, A. BONIFATI, B. GROZ.*A trichotomy for regular simple path queries on graphs*, in "Journal of Computer and System Sciences", December 2019, vol. 108, p. 29-48, forthcoming [*DOI :* 10.1016/J.JCSS.2019.08.006], https://hal.inria.fr/hal-02435355

[2] A. BONIFATI, U. COMIGNANI, E. COQUERY, R. THION.*Interactive Mapping Specification with Exemplar Tuples*, in "ACM Transactions on Database Systems", June 2019, vol. 44, n^{o} 3, 44, https://hal.inria.fr/hal-02096764

[3] A. BONIFATI, I. HOLUBOVÀ, A. PRAT-PÉREZ.*Graph Generators: State of the Art and Open Challenges*, in "ACM Computing Surveys", December 2019, forthcoming, https://hal.inria.fr/hal-02435371

[4] V. LOOTEN, L. KONG WIN CHANG, A. NEURAZ, M.-A. LANDAU-LORIOT, B. VEDIE, J.-L. PAUL, L. MAUGE, N. RIVET, A. BONIFATI, G. CHATELLIER, A. BURGUN, B. RANCE.*What can millions of laboratory test results tell us about the temporal aspect of data quality? Study of data spanning 17 years in a clinical data warehouse*, in "Computer Methods and Programs in Biomedicine", November 2019, vol. 181, p. 1-20, https://github.com/equipe22 [*DOI :* 10.1016/J.CMPB.2018.12.030], https://hal.archives-ouvertes.fr/hal-01978796

International Conferences with Proceedings

[5] A. BONIFATI, U. COMIGNANI, E. TSAMOURA.*Repairing mappings under policy views*, in "SIGMOD 2019 - ACM SIGMOD/PODS International Conference on Management of Data", Amsterdam, Netherlands, ACM, June 2019, p. 1873-1876, Demonstration, https://hal.inria.fr/hal-02096750

[6] A. BONIFATI, P. FURNISS, A. GREEN, R. HARMER, E. OSHURKO, H. VOIGT.*Schema validation and evolution for graph databases*, in "ER 2019 - 38th International Conference on Conceptual Modeling", Salvador, Brazil, Springer, November 2019, p. 448-456 [*DOI :* 10.1007/978-3-030-33223-5_37], https://hal.archives-ouvertes.fr/hal-02138771

[7] A. BONIFATI, W. MARTENS, T. TIMM.*Navigating the Maze of Wikidata Query Logs*, in "WWW 2019 - The World Wide Web Conference", San Francisco, United States, ACM, May 2019, p. 127-138 [*DOI :* 10.1145/3308558.3313472], https://hal.inria.fr/hal-02096714

[8] S. DUMBRAVA, A. BONIFATI, A. NAZABAL RUIZ DIAZ, R. VUILLEMOT.*Approximate Querying on Property Graphs*, in "SUM 2019 - 13th international conference on Scalable Uncertainty Management", Compiègne, France, December 2019, p. 250-265 [*DOI :* 10.1007/978-3-030-35514-2_19], https://hal.archives-ouvertes.fr/hal-02417259

[9] L. JACHIET, P. GENEVÈS, N. GESBERT, N. LAYAÏDA.*On the Optimization of Recursive Relational Queries: Application to Graph Queries*, in "SIGMOD 2020 - ACM International Conference on Management of Data", Portland, United States, June 2020, p. 1-23, https://hal.inria.fr/hal-01673025

Other Publications

[10] S. CHLYAH, N. GESBERT, P. GENEVÈS, N. LAYAÏDA.*An Algebra with a Fixpoint Operator for Distributed Data Collections*, March 2019, working paper or preprint, https://hal.inria.fr/hal-02066649