Activity Report Lille - Nord Europe 2019

Edition: 2020-03-23

List of Inria's Research Teams

Project-Team BONUS

Big Optimization aNd Ultra-Scale Computing

IN COLLABORATION WITH: Centre de Recherche en Informatique, Signal et Automatique de Lille

IN PARTNERSHIP WITH:
Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Optimization, machine learning and statistical methods

Table of contents

Project-Team BONUS

Creation of the Team: 2017 July 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.)
A1.1.5. - Exascale
A8.2.1. - Operations research
A8.2.2. - Evolutionary algorithms
A9.6. - Decision support
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B3.1. - Sustainable development
B3.1.1. - Resource management
B7. - Transport and logistics
B8.1.1. - Energy for smart buildings

1. Team, Visitors, External Collaborators

Research Scientist

Jan Gmys [University of Mons, Belgium, Researcher]

Faculty Members

Nouredine Melab [Team leader, Université de Lille, Professor, HDR]
Omar Abdelkafi [Université de Lille, Associate Professor]
Bilel Derbel [Université de Lille, Associate Professor, HDR]
Arnaud Liefooghe [Université de Lille, Associate Professor]
El-Ghazali Talbi [Université de Lille, Professor, HDR]

Technical Staff

Nassime Aslimani [Université de Lille, Engineer, from Sep 2019]
Dimitri Delabroye [Inria, Engineer, from Dec 2019]
Jingyu Ji [Inria, Engineer]

PhD Students

Nicolas Berveglieri [Université de Lille]
Alexandre Borges de Jesus [University of Coimbra, Portugal from Sep 2019]
Guillaume Briffoteaux [University of Mons, Belgium]
Sohrab Faramarzi Oghani [Inria, until Jan 2019]
Maxime Gobert [University of Mons, Belgium]
Ali Hebbal [Université de Lille]
Julien Pelamatti [ONERA]
Geoffrey Pruvost [Université de Lille]
Jeremy Sadet [Université de Valenciennes et du Hainaut Cambrésis]

Post-Doctoral Fellows

Tiago Carneiro Pessoa [Inria]
Mohammad Rahimi [Université de Lille, until Jan 2019]

Oumayma Bahri [Université de Lille, ATER]

Visiting Scientists

Martin Gonzalez [University of Cadiz, Spain, from Jul 2019 until Oct 2019]
Renzo Massobrio [Republica University, Uruguay, PhD student, from Jan 2019 to Mar 2019]
Leonard Whitley [Colorado State University, USA, Professor, Jul 2019]
Matthieu Basseur [Université d'Angers, Associate Professor, from Sep 2019 to Aug 2020, HDR]

Administrative Assistant

Karine Lewandowski [Inria]

2. Overall Objectives

2.1. Presentation

In the BONUS project, the context of optimization, where solving a problem consists in optimizing (minimizing or maximizing) one or more objective function(s) under some constraints, is considered. In this context, a problem can be formulated as follows:

$$\begin{aligned} &\text{Min/Max } F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_m(\mathbf{x})) \\ &\text{subject to} \quad \mathbf{x} \in S \end{aligned}$$

where S is the feasible search space and $\mathbf{x}$ is the decision variable vector of dimension n.
Nowadays, in many research and application areas we are witnessing the emergence of the big era (big data, big graphs, etc). In the optimization setting, the problems are increasingly big in practice. Big optimization problems (BOPs) refer to problems composed of a large number of environmental input parameters and/or decision variables (*high dimensionality*), and/or *many objective functions* that may be *computationally expensive*. For instance, in smart grids, there are many optimization problems for which have to be considered a large number of consumers (appliances, electrical vehicles etc.) and multiple suppliers with various energy sources. In the area of engineering design, the optimization process must often take into account a large number of parameters from different disciplines. In addition, the evaluation of the objective function(s) often consist(s) in the execution of an expensive simulation of a black-box complex system. This is for instance typically the case in aerodynamics where a CFD-based simulation may require several hours. On the other hand, to meet the high growing needs of applications in terms of computational power in a wide range of areas including optimization, high-performance computing (HPC) technologies have known a revolution during the last decade (see Top500 [0]). Indeed, HPC is evolving toward *ultra-scale supercomputers composed of millions of cores supplied in heterogeneous devices including multi-core processors with various architectures, GPU accelerators and MIC coprocessors*.
Beyond the "big buzzword", solving BOPs raises at least four major challenges: (1) tackling their high dimensionality; (2) handling many objectives; (3) dealing with computationally expensive objective functions; and (4) scaling on (ultra-scale) modern supercomputers. The overall scientific objectives of the BONUS project consist in addressing efficiently these challenges. On the one hand, the focus will be put on the design, analysis and implementation of optimization algorithms scalable to high-dimensional (in decision variables and/or objectives) and/or expensive problems. On the other hand, the focus will also be put on the design of optimization algorithms able to scale on heterogeneous supercomputers including several millions of processing cores. To achieve these objectives raising the associated challenges a program including three lines of research will be adopted (Fig. 1): *decomposition-based optimization, Machine Learning (ML)-assisted optimization and ultra-scale optimization*. These research lines are developed in the following section.

[0] Top500 international ranking (Edition of November 2018): https://www.top500.org/lists/2018/11/

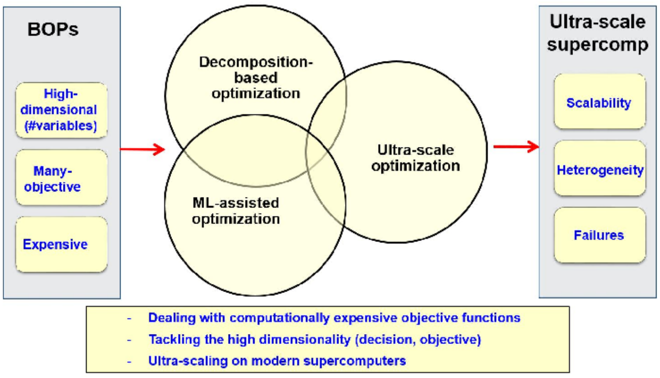

Figure 1. Research challenges/objectives and lines

From the software standpoint, our objective is to integrate the approaches we will develop in our *ParadisEO* [3] [0] framework in order to allow their reuse inside and outside the BONUS team. The major challenge will be to extend ParadisEO in order to make it *more collaborative* with other software including machine learning tools, other (exact) solvers and simulators. *From the application point of view*, the focus will be put on two classes of applications: *complex scheduling and engineering design*.

3. Research Program

3.1. Decomposition-based Optimization

Given the large scale of the targeted optimization problems in terms of the number of variables and objectives, their decomposition into simplified and loosely coupled or independent subproblems is essential to raise the challenge of scalability. The first line of research is to *investigate the decomposition approach in the two spaces and their combination, as well as their implementation on ultra-scale architectures*. The motivation of the decomposition is twofold: first, the decomposition allows the parallel resolution of the resulting subproblems on ultra-scale architectures. Here also several issues will be addressed: the definition of the subproblems, their coding to allow their efficient communication and storage (checkpointing), their assignment to processing cores etc. Second, decomposition is necessary for solving large problems that cannot be solved (efficiently) using traditional algorithms. Indeed, for instance with the popular NSGA-II algorithm the number of non-dominated solutions [0] increases drastically with the number of objectives leading to a very slow convergence to the Pareto Front [0]. Therefore, decomposition-based techniques are gaining a growing interest. The objective of BONUS is to *investigate various decomposition schema and cooperation protocols between the subproblems* resulting from the decomposition to generate efficiently global solutions of good quality.

[0] ParadisEO: http://paradiseo.gforge.inria.fr/

[0] A solution x dominates another solution y if x is better than y for all objectives and there exists at least one objective for which x is strictly better than y.

[0] The Pareto Front is the set of non-dominated solutions.

Several challenges have to be addressed: (1) how to define the subproblems (decomposition strategy), (2) how to solve them to generate local solutions (local rules), and (3) how to combine these latter with those generated by other subproblems and how to generate global solutions (cooperation mechanism), and (4) how to combine decomposition strategies in more than one space (hybridization strategy)? These challenges, which are in the line with the CIS Task Force [0] on decomposition will be addressed in the decision as well as in the objective space.

The *decomposition in the decision space* can be performed following different ways according to the problem at hand. Two major categories of decomposition techniques can be distinguished: the first one consists in *breaking down the high-dimensional decision vector* into lower-dimensional and easier-to-optimize blocks of variables. The major issue is how to define the subproblems (blocks of variables) and their cooperation protocol: randomly *vs.* using some learning (e.g. separability analysis), statically *vs.* adaptively etc. *The decomposition in the decision space can also be guided by the type of variables i.e. discrete vs. continuous*. The discrete and continuous parts are optimized separately using cooperative hybrid algorithms [48]. *The major issue of this kind of decomposition is the presence of categorial variables in the discrete part [44]. The* BONUS *team is addressing this issue, rarely investigated in the literature*, within the context of vehicle aerospace engineering design. The second category consists in the *decomposition according to the ranges of the decision variables*. For continuous problems, the idea consists in iteratively subdividing the search (e.g. design) space into subspaces (hyper-rectangles, intervals etc.) and select those that are most likely to produce the lowest objective function value. *Existing approaches meet increasing difficulty with an increasing number of variables and are often applied to low-dimensional problems. We are investigating this scalability challenge* (e.g. [10]). *For discrete problems, the major challenge is to find a coding (mapping) of the search space to a decomposable entity*. We have proposed an interval-based coding of the permutation space for solving big permutation problems. The approach opens perspectives we are investigating [7], in terms of ultra-scale parallelization, application to multi-permutation problems and hybridization with metaheuristics.

The *decomposition in the objective space* consists in breaking down an original Many-objective problem (MaOP) into a set of cooperative single-objective subproblems (SOPs). The decomposition strategy requires the careful definition of a scalarizing (aggregation) function and its weighting vectors (each of them corresponds to a separate SOP) to guide the search process towards the best regions. Several scalarizing functions have been proposed in the literature including weighted sum, weighted Tchebycheff, vector angle distance scaling etc. These functions are widely used but they have their limitations. For instance, using weighted Tchebycheff might do harm diversity maintenance and weighted sum is inefficient when it comes to deal with nonconvex Pareto Fronts [40]. Defining a scalarizing function well-suited to the MaOP at hand is therefore a difficult and still an open question being investigated in BONUS [6], [5]. Studying/defining various functions and in-depth analyzing them to better understand the differences between them is required. Regarding the weighting vectors that determine the search direction, their efficient setting is also a key and open issue. They dramatically affect in particular the diversity performance. Their setting rises two main issues: how to determine their number according to the available computational ressources? when (statically or adaptively) and how to determine their values? *Weight adaptation is one of our main concerns that we are addressing especially from a distributed perspective.* They correspond to the main scientific objectives targeted by our bilateral ANR-RGC BigMO project with City University (Hong Kong). The other challenges pointed out in the beginning of this section concern the way to solve locally the SOPs resulting from the decomposition of a MaOP and the mechanism used for their cooperation to generate global solutions. To deal with these challenges, our approach is to design the decomposition strategy and cooperation mechanism keeping in mind the parallel and/or distributed solving of the SOPs. Indeed, we favor the local neighborhood-based mating selection and replacement to minimize the network communication cost while allowing an effective resolution [5]. The major issues here are how to define the neighborhood of a subproblem and how to cooperatively update the best-known solution of each subproblem and its neighbors.

To sum up, the objective of the BONUS *team is to come up with scalable decomposition-based approaches in the decision and objective spaces. In the decision space, a particular focus will be put on high dimensionality*

[0] IEEE CIS Task Force, created in 2017 on Decomposition-based Techniques in Evolutionary Computation.

and mixed-continuous variables which have received little interest in the literature. We will particularly continue to investigate at larger scales using ultra-scale computing the interval-based (discrete) and fractal-based (continuous) approaches. We will also deal with the rarely addressed challenge of mixed-continuous including categorial variables (collaboration with ONERA). In the objective space, we will investigate parallel ultra-scale decomposition-based many-objective optimization with ML-based adaptive building of scalarizing functions. A particular focus will be put on the state-of-the-art MOEA/D algorithm. This challenge is rarely addressed in the literature which motivated the collaboration with the designer of MOEA/D (bilateral ANR-RGC BigMO project with City University, Hong Kong). Finally, the joint decision-objective decomposition, which is still in its infancy [50], is another challenge of major interest.

3.2. Machine Learning-assisted Optimization

The Machine Learning (ML) approach based on metamodels (or surrogates) is commonly used, and also adopted in BONUS, to assist optimization in tackling BOPs characterized by time-demanding objective functions. The second line of research of BONUS is focused on ML-aided optimization to raise the challenge of expensive functions of BOPs using surrogates but also to assist the two other research lines (decomposition-based and ultra-scale optimization) in dealing with the other challenges (high dimensionality and scalability). Several issues have been identified to make efficient and effective surrogate-assisted optimization. First, infill criteria have to be carefully defined to adaptively select the adequate sample points (in terms of surrogate precision and solution quality). The challenge is to find the best trade-off between exploration and exploitation to efficiently refine the surrogate and guide the optimization process toward the best solutions. The most popular infill criterion is probably the *Expected Improvement* (EI) [43] which is based on the expected values of sample points but also and importantly on their variance. This latter is inherently determined in the kriging model, this is why it is used in the state-of-the-art *efficient global optimization* (EGO) algorithm [43]. However, such crucial information is not provided in all surrogate models (e.g. Artificial Neural Networks) and needs to be derived. In BONUS, we are currently investigating this issue. Second, it is known that surrogates allow one to reduce the computational burden for solving BOPs with time-demanding function(s). However, using parallel computing as a complementary way is often recommended and cited as a perspective in the conclusions of related publications. Nevertheless, *despite being of critical importance parallel surrogate-assisted optimization is weakly addressed in the literature*. For instance, in the introduction of the survey proposed in [42] it is warned that because the area is not mature yet the paper is more focused on the potential of the surveyed approaches than on their relative efficiency. *Parallel computing is required at different levels that we are investigating*.

Another issue with surrogate-assisted optimization is related to high dimensionality in decision as well as in objective space: it is often applied to low-dimensional problems. *The joint use of decomposition, surrogates and massive parallelism is an efficient approach to deal with high dimensionality. This approach adopted in* BONUS *has received little effort in the literature.* In BONUS, we are considering a generic framework in order to enable a flexible coupling of existing surrogate models within the state-of-the-art decomposition-based algorithm MOEA/D. This is a first step in leveraging the applicability of efficient global optimization into the multi-objective setting through parallel decomposition. Another issue which is a consequence of high dimensionality is the mixed (discrete-continuous) nature of decision variables which is frequent in real-world applications (e.g. engineering design). *While surrogate-assisted optimization is widely applied in the continuous setting it is rarely addressed in the literature in the discrete-continuous framework.* In [44], we have identified different ways to deal with this issue that we are investigating. Non-stationary functions frequent in real-world applications (see Section 4.1) is another major issue we are addressing using the concept of deep GP.

Finally, as quoted in the beginning of this section, ML-assisted optimization is mainly used to deal with BOPs with expensive functions but it will also be investigated for other optimization tasks. Indeed, ML will be useful to assist the decomposition process. In the decision space, it will help to perform the separability analysis (understanding of the interactions between variables) to decompose the vector of variables. In the objective space, ML will be useful to assist a decomposition-based many-objective algorithm in dynamically selecting a scalarizing function or updating the weighting vectors according to their performances in the previous steps of the optimization process [5]. Such a data-driven ML methodology would allow us to understand what makes a problem difficult or an optimization approach efficient, to predict the algorithm performance [4], to select the most appropriate algorithm configuration [8], and to adapt and improve the algorithm design for unknown optimization domains and instances. Such an autonomous optimization approach would adaptively adjust its internal mechanisms in order to tackle cross-domain BOPs.

In a nutshell, to deal with expensive optimization the BONUS *team will investigate the surrogate-based ML approach with the objective to efficiently integrate surrogates in the optimization process. The focus will especially be put on high dimensionality (e.g. using decomposition) with mixed discrete-continuous variables which is rarely investigated. The kriging metamodel (Gaussian Process or GP) will be considered in particular for engineering design (for more reliability) addressing the above issues and other major ones including mainly non stationarity (using emerging deep GP) and ultra-scale parallelization (highly needed by the community). Indeed, a lot of work has been reported on deep neural networks (deep learning) surrogates but not on the others including (deep) GP. On the other hand, ML will be used to assist decomposition: importance/interaction between variables in the decision space, dynamic building (selection of scalarizing functions, weight update etc.) of scalarizing functions in the objective space etc.*

3.3. Ultra-scale Optimization

The third line of our research program that accentuates our difference from other (project-)teams of the related Inria scientific theme is the ultra-scale optimization. *This research line is complementary to the two others, which are sources of massive parallelism* and with which it should be combined to solve BOPs. Indeed, ultra-scale computing is necessary for the effective resolution of the large amount of subproblems generated by decomposition of BOPs, parallel evaluation of simulation-based fitness and metamodels etc. These sources of parallelism are attractive for solving BOPs and are natural candidates for ultra-scale supercomputers [0]. However, their efficient use raises a big challenge consisting in managing efficiently a massive amount of irregular tasks on supercomputers with multiple levels of parallelism and heterogeneous computing resources (GPU, multi-core CPU with various architectures) and networks. Raising such challenge requires to tackle three major issues, scalability, heterogeneity and fault-tolerance, discussed in the following.

The *scalability* issue requires, on the one hand, the definition of scalable data structures for efficient storage and management of the tremendous amount of subproblems generated by decomposition [46]. On the other hand, achieving extreme scalability requires also the optimization of communications (in number of messages, their size and scope) especially at the inter-node level. For that, we target the design of asynchronous locality-aware algorithms as we did in [41], [49]. In addition, efficient mechanisms are needed for granularity management and coding of the work units stored and communicated during the resolution process.

Heterogeneity means harnessing various resources including multi-core processors within different architectures and GPU devices. The challenge is therefore to design and implement hybrid optimization algorithms taking into account the difference in computational power between the various resources as well as the resource-specific issues. On the one hand, to deal with the heterogeneity in terms of computational power, we adopt in BONUS the dynamic load balancing approach based on the Work Stealing (WS) asynchronous paradigm [0] at the inter-node as well as at the intra-node level. We have already investigated such approach, with various victim selection and work sharing strategies in [49], [7]. On the other hand, hardware resource specific-level optimization mechanisms are required to deal with related issues such as thread divergence and memory optimization on GPU, data sharing and synchronization, cache locality, and vectorization on multi-core processors etc. These issues have been considered separately in the literature including our works [9], [1]. Indeed, in most of existing works related to GPU-accelerated optimization only a single CPU core is used. This leads to a huge resource wasting especially with the increase of the number of processing cores integrated into modern

[0] In the context of BONUS, supercomputers are composed of several massively parallel processing nodes (inter-node parallelism) including multi-core processors and GPUs (intra-node parallelism).

[0] A WS mechanism is mainly defined by two components: a victim selection strategy which selects the processing core to be stolen and a work sharing policy which determines the part and amount of the work unit to be given to the thief upon WS request.

processors. Using jointly the two components raises additional issues including data and work partitioning, the optimization of CPU-GPU data transfers etc.

Another issue the scalability induces is the *increasing probability of failures* in modern supercomputers [47]. Indeed, with the increase of their size to millions of processing cores their Mean-Time Between Failures (MTBF) tends to be shorter and shorter [45]. Failures may have different sources including hardware and software faults, silent errors etc. In our context, we consider failures leading to the loss of work unit(s) being processed by some thread(s) during the resolution process. The major issue, which is particularly critical in exact optimization, is how to recover the failed work units to ensure a reliable execution. Such issue is tackled in the literature using different approaches: algorithm-based fault tolerance, checkpoint/restart (CR), message logging and redundancy. The CR approach can be system-level, library/user-level or application-level. Thanks to its efficiency in terms of memory footprint, adopted in BONUS [2], the application-level approach is commonly and widely used in the literature. This approach raises several issues mainly: (1) which critical information defines the state of the work units and allows to resume properly their execution? (2) when, where and how (using which data structures) to store it efficiently? (3) how to deal with the two other issues: scalability and heterogeneity?
The last but not least major issue which is another roadblock to exascale is the programming of massive-scale applications for modern supercomputers. *On the path to exascale, we will investigate the programming environments and execution supports able to deal with exascale challenges: large numbers of threads, heterogeneous resources etc.* Various exascale programming approaches are being investigated by the parallel computing community and HPC builders: extending existing programming languages (e.g. DSL-C++) and environments/libraries (MPI+X etc.), proposing new solutions including mainly Partitioned Global Address Space (PGAS)-based environments (Chapel, UPC, X10 etc.). It is worth noting here that our objective is not to develop a programming environment nor a runtime support for exascale computing. Instead, we aim to collaborate with the research teams (inside or outside Inria) having such objective.

To sum up, we put the focus on the design and implementation of efficient big optimization algorithms dealing jointly (uncommon in parallel optimization) with the major issues of ultra-scale computing mainly the scalability up to millions of cores using scalable data structures and asynchronous locality-aware work stealing, heterogeneity addressing the multi-core and GPU-specific issues and those related to their combination, and scalable GPU-aware fault tolerance. A strong effort will be devoted to this latter challenge, for the first time to the best of our knowledge, using application-level checkpoint/restart approach to deal with failures.

4. Application Domains

4.1. Introduction

For the validation of our findings we obviously use standard benchmarks to facilitate the comparison with related works. In addition, we also target real-world applications in the context of our collaborations and industrial contracts. From the *application* point of view two classes are targeted: *complex scheduling* and *engineering design*. The objective is twofold: proposing new models for complex problems and solving efficiently BOPs using jointly the three lines of our research program. In the following, are given some use cases that are the focus of our current industrial collaborations.

4.2. Big optimization for complex scheduling

Three application domains are targeted: energy, health and transport and logistics. In the **energy** field, with the smart grid revolution (multi-)house energy management is gaining a growing interest. The key challenge is to make elastic with respect to the energy market the (multi-)house energy consumption and management. *This kind of demand-side management will be of strategic importance for energy companies in the near future.* In collaboration with the EDF energy company we are working on the formulation and solving of optimization problems on demand-side management in smart micro-grids for single- and multi-user frameworks. These complex problems require taking into account multiple conflicting objectives and constraints and many (deterministic/uncertain, discrete/continuous) parameters. A representative example of such BOPs that we are addressing is the scheduling of the activation of a large number of electrical and thermal appliances for a set of homes optimizing at least three criteria: maximizing the user's confort, minimizing its energy bill and minimzing peak consumption situations. In the **health** care domain, we are collaborating with the Beckman & Coulter company on the design and planning of large medical laboratories. This is a hot topic resulting from the mutualisation phenomenon which makes these laboratories bigger. As a consequence, being responsible for analyzing medical tests ordered by physicians on patient's samples, these laboratories receive large amounts of prescriptions and tubes making their associated workflow more complex. Our aim is therefore to design and plan any medical laboratory to minimize the costs and time required to perform the tests. More exactly, the focus is put on the multi-objective modeling and solving of large (e.g. dozens of thousands of medical test tubes to be analyzed) strategic, tactical and operational problems such as the layout design, machine selection and confguration, assignment and scheduling. Finally, in **transport and logistics**, within the context of our potential collaboration (being set up) with the EXOTEC company we target the optimization of the robotic logistics of 3D warehouses. More exactly, the problem consists in efficient complex scheduling without collision of thousands of missions realized by a fleet of dozens of robots and several operators in a 3D logistics warehouse. The problem is identified in the literature as the parts-to-picker based order processing in a rack-moving mobile robots environment.

4.3. Big optimization for engineering design

The focus is for now put on the aerospace vehicle design, a complex multidisciplinary optimization process, we are exploring in collaboration with ONERA. The objective is to find the vehicle architecture and characteristics that provide the optimal performance (flight performance, safety, reliability, cost etc.) while satisfying design requirements [39]. A representative topic we are investigating, and will continue to investigate throughout the lifetime of the project given its complexity, is the design of launch vehicles that involves at least four tightly coupled disciplines (aerodynamics, structure, propulsion and trajectory). Each discipline may rely on time-demanding simulations such as Finite Element analyses (structure) and Computational Fluid Dynamics analyses (aerodynamics). Surrogate-assisted optimization is highly required to reduce the time complexity. In addition, the problem is high-dimensional (dozens of parameters and more than three objectives) requiring different decomposition schemas (coupling *vs.* local variables, continuous *vs.* discrete even categorial variables, scalarization of the objectives). Another major issue arising in this area is the non-stationarity of the objective functions which is generally due to the abrupt change of a physical property that often occurs in the design of launch vehicles. In the same spirit than deep learning using neural networks, we use Deep Gaussian Processes to deal with non-stationary multi-objective functions. Finally, the resolution of the problem using only one objective takes one week using a multi-core processor. Therefore, *in addition to surrogates ultra-scale computing is required at different levels to speed up the search and improve the reliability which is a major requirement in aerospace design.* This example shows that we need to use the synergy between the three lines of our research program to tackle such BOPs.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

+ The paper [30] was nominated for the *Best Student Workshop Paper Award* at the 28$^{\text{th}}$ ACM (Companion) Genetic and Evolutionary Computation Conference (GECCO 2019).
+ *USA Patent* with Beckman & Coulter on the optimization of large medical laboratories (Prof. E-G. Talbi, S. Faramarzi Oghani, M. Bué), 2019.

+ The paper [29] has received the *Best Student Paper Award* at the 10th International Conference on Evolutionary Multi-Criterion Optimization (EMO 2019).
  BEST PAPERS AWARDS :
  [29]
  Y. MARCA, H. AGUIRRE, S. Z. MARTINEZ, A. LIEFOOGHE, B. DERBEL, S. VEREL, K. TANAKA.*Approximating Pareto Set Topology by Cubic Interpolation on Bi-objective Problems*, in "EMO 2019 - International Conference on Evolutionary Multi-Criterion Optimization", East Lansing, Michigan, United States, February 2019, p. 386-398 [*DOI :* 10.1007/978-3-030-12598-1_31], https://hal.archives-ouvertes.fr/hal-02064548

6. New Software and Platforms

6.1. Platforms

6.1.1. Grid'5000 testbed: extension with storage and 10G network at Lille

Participants: Nouredine Melab [contact person], Dimitri Delabroye, Thierry Peltier [external collaborator], Lucas Nussbaum [external collaborator].

Keywords: Experimental testbed, large-scale computing, high-performance computing, GPU computing, cloud computing, big data

Functional description: Grid'5000 is a project intiated in 2003 by the French government to promote scientific research on large scale distributed systems. The project is later supported by different research organizations including Inria, CNRS, the french universities, Renater which provides the wide-area network etc. The overall objective of Grid'5000 was to build by 2007 a nation-wide experimental testbed composed of at least 5000 processing units and distributed over several sites in France. From a scientific point of view, the aim was to promote scientific research on large-scale distributed systems.

Grid'5000 was installed at the center of IT resources including supercomputing resources of Université de Lille 1 and opened to users in 2005. Since March 2017, the Grid'5000 site has moved to the premises on Inria Lille within the context of the phase 1 of the CPER data program (see Section 9.1) with a completely new hardware equipement. As a scientific leader of the testbed for the Lille's site, N. Melab has been strongly involved in the extension (phase 2 of CPER data) of the platform with mainly two storage serveurs 200TB and a 10G Internet connexion to the testbed. Grid'5000 at Lille is used by more than 150 users including 100 external ones. The testbed is used for research as well as for teaching allowing a high scientific production (publications, PhD theses etc.) and over 30 master students to get started with parallel and distributed programming.

URL: https://www.grid5000.fr/mediawiki/index.php/Grid5000:Home

7. New Results

7.1. Decomposition-based optimization

During the year 2019, we have investigated decomposition-based optimization in the decision space as well as in the objective space. In the decision space, we have considered discrete as well as continuous problems. For discrete problems, we have reinvestigated our interval-based decomposition approach proposed in [7] as a baseline to explore ultra-scale Branch-and-Bound algorithms using Chapel [13]. The contribution is presented in Section 7.3. For continuous problems, we have extended the geometric fractal decomposition-based approach [10] to multi-objective optimization [32] and importantly to parallel computing [17] to deal with scalability, one of the major scientific challenges as pointed out in Section 3.1. In the objective space, we have deeply studied in [23] the surrogate-assisted multi-objective optimization based on decomposition. The contributions are summarized in the following.

7.1.1. Parallel fractal decomposition for big continuous optimization problems

Participants: El-Ghazali Talbi [contact person], Amir Nakib [Laboratoire Images, Signaux et Systèmes Intelligents (LISSI), Paris], Léo Souquet [Laboratoire Images, Signaux et Systèmes Intelligents (LISSI), Paris].

Fractal Decomposition Algorithm (FDA) is a metaheuristic that was recently proposed to solve high dimensional continuous optimization problems [10]. This approach is based on a geometric fractal decomposition which divides the search space while looking for the optimal solution. While FDA and its fractal decomposition has shown to be an effective optimization algorithm, its running time grows significantly as the problems dimension increases. To deal with this expensive computational time, a parallelized version of FDA, called Parallel Fractal Decomposition Algorithm (PFDA) is proposed in [17]. The focus is on parallelizing the exploration and exploitation phases of the original algorithm in a multi-threaded environment. The performances of PFDA are evaluated on the same Benchmark used to illustrate FDA efficiency, the SOCO 2011. It is composed of 19 functions with *dimensions going from 50 to 5000*. The results show that PFDA allows one to achieve similar performances as the original version with a significantly reduced computational time.

7.1.2. Deterministic multi-objective fractal decomposition

Participants: El-Ghazali Talbi [contact person], Amir Nakib [Laboratoire Images, Signaux et Systèmes Intelligents (LISSI), Paris], Léo Souquet [Laboratoire Images, Signaux et Systèmes Intelligents (LISSI), Paris].

We have proposed in [32] a new deterministic Multi-objective Fractal Decomposition Algorithm (Mo-FDA). The original FDA [10] was designed for single-objective large-scale continuous optimization problems. It is based on a “divide-and-conquer” strategy and a geometric fractal decomposition of the search space using hyperspheres. In this work, a scalarization approach is used to deal with multi-objective problems. The performance of Mo-FDA is compared to state-of-the-art algorithms from the literature on classical benchmarks of multi-objective optimization.

7.1.3. Surrogate-assisted multi-objective optimization based on decomposition

Participants: Nicolas Berveglieri, Bilel Derbel, Arnaud Liefooghe, Hernan Aguirre [Shinshu University, Japan], Kiyoshi Tanaka [Shinshu University, Japan].

A number of surrogate-assisted evolutionary algorithms are being developed for tackling expensive multi-objective optimization problems. On the one hand, a relatively broad range of techniques from both machine learning and multi-objective optimization can be combined for this purpose. Different taxonomies exist in order to better delimit the design choices, advantages and drawbacks of existing approaches. On the other hand, assessing the relative performance of a given approach is a difficult task, since it depends on the characteristics of the problem at hand. In [23], we focus on surrogate-assisted approaches using objective space decomposition as a core component. We propose a refined and fine-grained classification, ranging from EGO-like approaches to filtering or pre-screening. More importantly, we provide a comprehensive comparative study of a representative selection of state-of-the-art methods, together with simple baseline algorithms. We rely on selected benchmark functions taken from the bbob-biobj benchmarking test suite, that provides a variable range of objective function difficulties. Our empirical analysis highlights the effect of the available budget on the relative performance of each approach, and the impact of the training set and of the machine learning model construction on both solution quality and runtime efficiency.

7.2. ML-assisted optimization

As pointed out in our research program 3.2, we investigate the ML-assisted optimization following two directions: building efficiently surrogates to deal with expensive black-box objective functions and automatically building and predicting/improving the performance of metaheuristics through landscape/problem structure analysis. Regarding surrogate-assisted optimization, we put the focus in [31] on mixed discrete-continuous optmization problems, one of the major challenges of this topic. In addition, we focused on the evolution control and batch parallelism to deal with another challenge which consists in efficiently integrating surrogate models (Bayesian neural networks in this case) in evolutionary algorithms [FGCS-Briffoteaux, à compléter]. From the application point of view, we applied our approaches to three different real-world simulation-based problems in the context of three collaborations: multi-stage optimal scheduling of virtual power plants (collaboration with electrical engineering department of UMONS University, Belgium) in [27], aerospace vehicle design (collaboration with ONERA, Paris) in [31], and resource allocation for the Tuberculosis edidemic control (Monash University) in [FGCS-Briffoteax, à compléter]. The two latter ones are summarized in the following sections. Regarding the second direction, we thoroughly study in [16] the impact of landscape characteristics on the performance of search heuristics for black-box multi-objective combinatorial optimization problems. We also introduce in [26] new insightful features for continuous exploratory landscape analysis and algorithm selection. In addition, as pointed out in 3.2, variable selection is highly important to deal with BOPs. In [28], in collaboration with our partners from Shinshu University we come out with an efficient method to classify variables influencing convergence and increase their recombination rate. The contributions are summarized in the following.

7.2.1. Surrogate-assisted optimization of constrained mixed variable problems: application to the design of a launch vehiclethrust frame.

Participants: El-Ghazali Talbi [contact person], Julien Pelamatti, Loïc Brevault [ONERA], Mathieu Balesdent [ONERA], Yannick Guerin [CNES].

Within the framework of complex systems design, such as launch vehicles, numerical optimization is an essential tool as it allows to reduce the design process time and costs. The inclusion of discrete variables in the design process allows to extend the applicability of numerical optimization methods to a broader number of systems and sub-systems. In [31], a recently proposed adaptation of the Efficient Global Optimization method (EGO) for constrained mixed-variable problems is applied to the design optimization of a launch vehicle thrust frame, which depends on both continuous sizing parameters and discrete variables characterizing the number of structural reinforcements. The EGO adaptation that is considered is based on a redefinition of the Gaussian Process kernel as a product between a standard continuous kernel and a second kernel representing the covariance between the discrete variable values. From the results obtained on an analytical test-case as well as on the launch vehicle thrust frame design optimization, it is shown that the use of the mixed-variable EGO algorithm allows to converge towards the neighborhoods of the problems optima with fewer function evaluations when compared to reference optimization algorithms.

7.2.2. Parallel Batched Bayesian Neural Network-assisted GA versus q-EGO

Participants: Guillaume Briffoteaux, Maxime Gobert, Jan Gmys, Nouredine Melab [contact person], Romain Ragonnet [School of Public Health and Preventive Medicine, Monash University, Australia], Mohand Mezmaz [University of Mons, Blegium], Daniel Tuyttens [University of Mons, Blegium].

Surrogate-based optimization has been widely used to deal with expensive black-box simulation-based objective functions. The use of a surrogate model allows to reduce the number of calls to the costly simulator. In (SWEVO, Briffoteaux et al., 2020), the Efficient Global Optimization (EGO) reference framework is challenged by a Bayesian Neural Network-assisted Genetic Algorithm, namely BNN-GA. The Bayesian Neural Network (BNN) surrogate provides an uncertainty measure of the prediction that allows to compute the Expected Improvement of a candidate solution in order to improve the exploration of the objective space. BNN is also more reliable than Kriging models for high-dimensional problems and faster to set up thanks to its incremental training. In addition, we propose a batch-based approach for the parallelization of BNN-GA that is challenged by a parallel version of EGO, called q-EGO. Parallel computing is a complementary way to deal with the computational burden of simulation-based optimization. The comparison of the two parallel approaches is experimentally performed through several benchmark functions and two real-world problems within the scope of Tuberculosis Transmission Control (TBTC). The results demonstrate the efficiency and scalability of the parallel batched BNN-GA for small budgets, outperforming it for larger budgets on the benchmark testbed. A significant improvement of the solutions is obtained for two TBTC problems. Finally, our study proves that parallel batched BNN-GA is a viable alternative to q-EGO approaches being more suitable for high-dimensional problems and parallelization impact.

7.2.3. Landscape-aware performance prediction and algorithm selection for single- and multi-objective evolutionary optimization

Participants: Arnaud Liefooghe, Bilel Derbel, Fabio Daolio [ASOS.com, UK], Sébastien Verel [LISIC, Université du Littoral Côte d'Opale], Hernan Aguirre [Shinshu University, Japan], Kiyoshi Tanaka [Shinshu University, Japan].

Extracting a priori knowledge informing about the landscape underlying an unknown optimization problem has been proved extremely useful for different purposes, such as designing finely-tuned algorithms, predicting algorithm performance and designing automated portfolio-based solving approaches. Considering black-box continuous single-objective optimization problems, in [26], we adopt an exploratory landscape analysis approach providing a unified methodology for integrating landscape features into sophisticated machine learning techniques. More precisely, we consider the design of novel informative and cheap landscape features on the basis of the search tree constructed by the so-called SOO (Simultaneous Optimistic Optimization) algorithm, which is arguably a global optimizer coming from the machine learning community and having its foundations in the multi-armed bandit theory. We thereby provide empirical evidence on the accuracy of the proposed features for both predicting high-level problem properties, and tackling the algorithm selection problem with respect to a given portfolio of available solvers and a broad range of optimisation problems taking from the specialized literature and exposing different degrees of difficulty.

Considering black-box combinatorial multi-objective optimization problems, in [16], we expose and contrast the impact of landscape characteristics on the performance of search heuristics for black-box multi-objective combinatorial optimization problems. A sound and concise summary of features characterizing the structure of an arbitrary problem instance is identified and related to the expected performance of global and local dominance-based multi-objective optimization algorithms. We provide a critical review of existing features tailored to multi-objective combinatorial optimization problems, and we propose additional ones that do not require any global knowledge from the landscape, making them suitable for large-size problem instances. Their intercorrelation and their association with algorithm performance are also analyzed. This allows us to assess the individual and the joint effect of problem features on algorithm performance, and to highlight the main difficulties encountered by such search heuristics. By providing effective tools for multi-objective landscape analysis, we highlight that multiple features are required to capture problem difficulty, and we provide further insights into the importance of ruggedness and multimodality to characterize multi-objective combinatorial landscapes.

7.2.4. Estimating Relevance of Variables for Effective Recombination

Participants: Arnaud Liefooghe, Bilel Derbel, Sébastien Verel [LISIC, Université du Littoral Côte d'Opale], Taishi Ito [Shinshu University, Japan], Hernan Aguirre [Shinshu University, Japan], Kiyoshi Tanaka [Shinshu University, Japan].

Dominance and its extensions, decomposition, and indicator functions are well-known approaches used to design MOEAs. Algorithms based on these approaches have mostly sought to enhance parent selection and survival selection. In addition, several variation operators have been developed for MOEAs. In [28], we focus on the classification and selection of variables to improve the effectiveness of solution search. We propose a method to classify variables that influence convergence and increase their recombination rate, aiming to improve convergence of the approximation found by the algorithm. We incorporate the proposed method into NSGA-II and study its effectiveness using three-objective DTLZ and WFG benchmark functions, including unimodal, multimodal, separable, non-separable, unbiased, and biased functions. We also test the effectiveness of the proposed method on a real-world bi-objective problem. Simulation results verify that the proposed method can contribute to achieving faster and better convergence in several kinds of problems, including the real-world problem.

7.3. Towards ultra-scale Big Optimization

During the year 2019, we have addressed the ultra-scale optimization research line following two main directions: designing efficient optimization algorithms dealing with scalability in terms of number of processing cores and/or in terms of the size of tackled problem instances. For the first direction, the challenge is to take into account in addition to the traditional performance objective the productivity awareness which is highly important to deal with the increasing complexity of medern supercomputers. With the short-term perspective of establishing a collaboration with one of the major HPC builders (Cray Inc.), we have investigated in [13], [24], [25] the exascale-aware Chapel language. Regarding the second direction, we contributed to solve difficult unsolved flow-shop [15] and Quadratic Assignement Problem (QAP) [22] permutation problem instances using moderate-scale parallelism. The contributions are summarized in the following.

7.3.1. Towards ultra-scale Branch-and-Bound using a high-productivity language

Participants: Tiago Carneiro Pessoa, Jan Gmys, Nouredine Melab, Daniel Tuyttens [University of Mons, Blegium].

Productivity is crucial for designing ultra-scale algorithms able to harness modern supercomputers which are increasingly complex, including millions of processing cores and heterogeneous building-block devices. In [13], we investigate the partitioned global address space (PGAS)-based approach using Chapel for the productivity-aware design and implementation of distributed Branc-and-Bound (B&B) for solving large optmization problems. The proposed algorithms are intensively experimented using the Flow-shop scheduling problem as a test-case. The Chapel-based implementation is compared to its MPI+X-based traditionally used counterpart in terms of performance, scalability, and productivity. The results show that Chapel is much more expressive and up to $7.8\times$ more productive than MPI+Pthreads. In addition, the Chapel-based search presents performance equivalent to MPI+Pthreads for its best results on 1024 cores and reaches up to 84% of the linear speedup. However, there are cases where the built-in load balancing provided by Chapel cannot produce regular load among computer nodes. In such cases, the MPI-based search can be up to $4.2\times$ faster and reaches speedups up to $3\times$ higher than its Chapel-based counterpart. Thorough feedback on the experience is given, pointing out the strengths and limitations of the two opposite approaches (Chapel vs. MPI+X). To the best of our knowledge, the present study is pioneering within the context of exact parallel optimization.

7.3.2. A computationally efficient Branch-and-Bound algorithm for the permutation flow-shop scheduling problem

Participants: Jan Gmys, Nouredine Melab, Mohand Mezmaz [University of Mons, Blegium], Daniel Tuyttens [University of Mons, Blegium].

In [15], we propose an efficient Branch-and-Bound (B&B) algorithm for the permutation flow-shop problem (PFSP) with makespan objective. We present a new node decomposition scheme that combines dynamic branching and lower bound refinement strategies in a computationally efficient way. To alleviate the computational burden of the two-machine bound used in the refinement stage, we propose an online learning-inspired mechanism to predict promising couples of bottleneck machines. The algorithm offers multiple choices for branching and bounding operators and can explore the search tree either sequentially or in parallel on multicore CPUs. In order to empirically determine the most efficient combination of these components, a series of computational experiments with 600 benchmark instances is performed. A main insight is that the problem size, as well as interactions between branching and bounding operators substantially modify the trade-off between the computational requirements of a lower bound and the achieved tree size reduction. Moreover, we demonstrate that parallel tree search is a key ingredient for the resolution of large problem instances, a strong super-linear speedups can be observed. An overall evaluation using two well-known benchmarks indicates that the proposed approach is superior to previously published B&B algorithms. For the first benchmark we report the exact resolution – within less than 20 minutes – of two instances defined by 500 jobs and 20 machines that remained open for more than 25 years, and for the second a total of 89 improved best-known upper bounds, including proofs of optimality for 74 of them.

7.3.3. A Parallel Tabu Search for the Large-scale Quadratic Assignment Problem

Participants: Omar Abdelkafi, Bilel Derbel, Arnaud Liefooghe.

Parallelization is an important paradigm for solving massive optimization problems. Understanding how to fully benefit form the aggregated computing power and what makes a parallel strategy successful is a difficult issue. In [22], we propose a simple parallel iterative tabu search (PITS) and study its effectiveness with respect to different experimental settings. Using the quadratic assignment problem (QAP) as a case study, we first consider different small-and medium-size instances from the literature and then tackle a large-size instance that was rarely considered due the its inherent solving difficulty. In particular, we show that a balance between the number of function evaluations each parallel process is allowed to perform before resuming the search is a critical issue to obtain an improved quality.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Our current industrial contracts and granted projects are completely at the heart of the BONUS project. They are summarized in the following.

- *EDF (2015-2019, Paris)*: this project deals with demand-side management in smart grids with EDF, a major electrical power player in France. The Energy Management System (EMS) in the home receives the market and system signals and controls the loads, Heating, Ventilation and Air Conditioning systems (HVAC), storages and local generation units according to the user preferences. A large number of home users and appliances and several conflicting objectives have to be considered.
- *ONERA & CNES (2016-2020, Paris)*: the focus of this project with major European players in vehicle aerospace is put on the design of aerospace vehicles, a high-dimensional expensive multidisciplinary problem. Such problem needs the use of the research lines of BONUS to be tackled effectively and efficiently. Two jointly supervised PhD students (J. Pelamatti and A. Hebbal) are involved in this project.
- *In contact with Decathlon (2019, Lille)*: This project deals with scalable multi-objective optimization for the eco-design of material, clothing and sports shoes.
- *In contact with Vinci Autoroutes (2019, Paris)*: This project deals with the optimization of deep neural networks for computer vision.

9. Partnerships and Cooperations

9.1. Regional Initiatives

- *CPER Data (2015-2020)*: in this project, that promotes research and software development related to advanced data science, the BONUS team is the scientific leader (N. Melab) of one of the three research lines of the project “Optimization and High-Performance Computing”. In this context, a two-year (2018-2019) engineer (J-Y. Ji) is supported to develop a software demonstrator on decomposition-based big optimization. In addition, the team is co-leader of the workpackage/lever “Research infrastructures” related to the Grid’5000 nation-wide experimental testbed. This allowed to extend the testbed at Lille with a GPU-powered cluster highly important for the BONUS project. In addition, two engineers have been hired for the system & network administration of the testbed, user support and development.
- *CPER ELSAT (2015-2020)*: in this project, focused on ecomobility, security and adaptability in transport, the BONUS team is involved in the transversal research line: planning and scheduling of maintenance logistics in transportation. The team got support for a one-year (2017-2018) post-doc position (M. Rahimi) and a one-year (2019-2020) engineer position (N. Aslimani).

9.2. National Initiatives

9.2.1. ANR

- *Bilateral ANR/RGC France/Hong Kong PRCI* (2016-2021), “Big Multi-objective Optimization” in collaboration with City University of Hong Kong

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

Program: H2020

Project acronym: SYNERGY

Project title: Synergy for Smart Multi-Objective Optimisation

Duration: 02 2016 - 03 2019

Coordinator: Jožef Stefan Institute (JSI), Ljubljana, Slovenia

Other partners: University of Lille (France), Cologne University of Applied Sciences (Germany)

Abstract: Many real-world application areas, such as advanced manufacturing, involve optimization of several, often time-consuming and conflicting objectives. For example, they require the maximization of the product quality while minimizing the production cost, and rely on demanding numerical simulations in order to assess the objectives. These, so-called multi-objective optimization problems can be solved more efficiently if parallelization is used to execute the simulations simultaneously and if the simulations are partly replaced by accurate surrogate models.

9.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: COST CA15140

Project acronym: ImAppNIO

Project title: Improving applicability of nature-inspired optimization by joining theory and practice

Duration: 2016-2019

Coordinator: Thomas Jansen

Abstract: The main objective of the COST Action is to bridge this gap and improve the applicability of all kinds of nature-inspired optimisation methods. It aims at making theoretical insights more accessible and practical by creating a platform where theoreticians and practitioners can meet and exchange insights, ideas and needs; by developing robust guidelines and practical support for application development based on theoretical insights; by developing theoretical frameworks driven by actual needs arising from practical applications; by training Early Career Investigators in a theory of nature-inspired optimisation methods that clearly aims at practical applications; by broadening participation in the ongoing research of how to develop and apply robust nature-inspired optimisation methods in different application areas.

9.3.3. Collaborations with Major European Organizations

University of Mons, Belgium, Parallel surrogate-assisted optimization, large-scale exact optimization, two joint PhDs (M. Gobert and G. Briffoteaux).

University of Luxembourg, Q-Learning-based Hyper-Heuristic for Generating UAV Swarming Behaviours.

University of Coimbra and University of Lisbon, Portugal, Exact and heuristic multi-objective search.

University of Manchester, United Kingdom, Local optimality in multi-objective optimization.

University of Elche and University of Murcia, Spain, Matheuristics for DEA.

University of Mohamed V, Morocco, Large scale (multi-objective) optimization.

9.4. International Initiatives

9.4.1. Inria International Labs

9.4.1.1. Other IIL projects

Title: Frontiers in Massive Optimization and Computational Intelligence (MODO)

International Partner (Institution - Laboratory - Researcher): Shinshu University, Japan

Start year: 2017

See also: https://sites.google.com/view/lia-modo/

Abstract: The aim of MODO is to federate French and Japanese researchers interested in the dimensionality, heterogeneity and expensive nature of massive optimization problems. The team receives a yearly support for international exchanges and shared manpower (joint PhD students).

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

Title: Three-fold decomposition in multi-objective optimization (D^3MO)

International Partner (Institution - Laboratory - Researcher): University of Exeter, UK

Start year: 2018

9.4.3. Inria International Partners

9.4.3.1. Informal International Partners

- School of Public Health and Preventive Medicine, Monash University, Australia (ranked 73^{th} over 1000 in the Shangai international ranking).
- Instituto Federal de Educação, Ciência e Tecnologia do Ceará, Maracanaú, Brazil.

9.4.4. Participation in Other International Programs

Title: **Evolutionary many-objective optimization: application to smart cities and engineering design**

International Partner (Institution - Laboratory - Researcher): CINVESTAV-IPN, Mexico

Start year: 2016

Abstract: The project is co-funded by ECOS Nord, France and ANUIES, Mexico. It is focused on evolutionary many-objective optimization and its application to smart cities and engineering design.

Title: **Bridging the gap between exact methods and heuristics for multi-objective search (MOCO-Search)**

International Partner (Institution - Laboratory - Researcher): University of Coimbra and University of Lisbon, Portugal

Start year: 2018

Website: http://sites.google.com/view/moco-search/

Abstract: This international project for scientific cooperation (PICS), funded by CNRS and FCT, aims to fill the gap between exact and heuristic methods for multi-objective optimization. The goal is to establish the link between the design principles of exact and heuristic methods, to identify features that make a problem more difficult to be solved by each method, and to improve their performance by hybridizing search strategies. Special emphasis is given to rigorous performance assessment, benchmarking, and general-purpose guidelines for the design of exact and heuristic multi-objective search.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Luís Paquete, University of Coimbra, Portugal, April 2019
- Darrell Whitley, Colorado State University, USA, Invited Professor, July 2019
- Minami Miyakawa, Shinshu University, Japan, October 2019
- Renzo Massobrio, Republica University, Uruguay, January to March 2019
- Bernabe Dorronsoro, University of Cadiz, Spain, March 2019
- Rachid Ellaia, University of Mohamed V, Morocco, April 2019

9.5.1.1. Internships

- Kazuki Maeda, Shinshu University, Japan, November-December 2019
- Kyo Migishima, Shinshu University, Japan, November-December 2019

9.5.2. Visits to International Teams

9.5.2.1. Explorer programme

- T. Carneiro, Cray Inc., Seattle, WA, USA, December 2019
- E-G. Talbi, University of Elche, Spain, November 2019
- E-G. Talbi, University of Luxembourg, Luxembourg, June 2019
- E-G. Talbi, University of Bangkok, Thailand, January 2019
- E-G. Talbi, University of Colorado, USA, November 2019
- E-G. Talbi, University of Mohamed V, Morocco, April 2019
- A. Liefooghe, Shinshu University, Japan, May 2019
- B. Derbel, University of Coimbra, Portugal, October 2019
- A. Liefooghe, University of Coimbra, Portugal, October 2019
- N. Melab, University of Mons, Belgium, working meetings throughout the year

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- E-G. Talbi (Conference chair): Intl. Conf. on Multiple Objective Programming and Goal Programming (MOPGP'2019), Marrakech, Morocco, Oct 28-31, 2019.
- N. Melab (Workshop co-chair): Intl. Workshop on the Synergy of Parallel Computing, Optimization and Simulation (HPCS/PaCOS'2019), Dublin, Ireland, Jul 15-19, 2019.
- E-G. Talbi (Steering committee Chair): Intl. Conf. on Optimization and Learning (OLA'2019), Bangkok, Thailand, Jan 29-31, 2019.
- E-G. Talbi (Steering committee): 9^{th} IEEE Workshop Parallel Distributed Computing and Optimization (IPDPS/PDCO'2019), Rio de Janeiro, Brazil, May 20-24, 2019.
- O. Abdelkafi, B. Derbel and A. Liefooghe (workshop co-chairs): 2^{nd} Intl. Workshop on Computational Intelligence for Massive Optimization (CIMO 2019), Lille, France, July 2019 (with H. Aguirre, K. Tanaka and S. Verel).
- B. Derbel (workshop co-chair): Decomposition Techniques in Evolutionary Optimization (DTEO), workshop at GECCO 2019, Prague, Czech Republic, July 2019 (with K. Li, X. Li, S. Zapotecas, and Q. Zhang).

- A. Liefooghe (workshop co-chair): Landscape-aware heuristic search (LAHS), workshop at GECCO 2019, Prague, Czech Republic, July 2019 (with N. Veerapen, S. Verel and G. Ochoa).
- B. Derbel (special session co-chair): Advances in Decomposition-based Evolutionary Multi-objective Optimization (ADEMO), special session at CEC 2019, Wellington, New Zealand, June 2019 (with S. Zapotecas, K. Li and Q. Zhang).
- A. Liefooghe (track chair): GECCO 2019: Genetic and Evolutionary Computation Conference, Evolutionary multi-objective Optimization (EMO) track (Prague, Czech Republic, 2019).
- N. Melab: Chair of 4 simulation and HPC-related seminars at Université de Lille: The Arctic University of Norway (HPC Group, UiT), IBM, DSI (Supercomputing division) Ulille, Inria Lille), Oct-Dec 2019.

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- A. Liefooghe (program co-chair): EvoCOP 2019: 19th European Conference on Evolutionary Computation in Combinatorial Optimisation (Leipzig, Germany, 2019).

10.1.2.2. Member of the Conference Program Committees

- IEEE Congress on Evolutionary Computation (CEC), Wellington, New Zealand, Jun 10-13, 2019.
- IEEE Workshop on Understanding of Evolutionary Optimization Behavior, Wellington, New Zealand, Jun 10-13, 2019.
- The ACM Genetic and Evolutionary Computation Conference (GECCO), Prague, Czech Republic, July 13-17, 2019.
- 13th Metaheuristics International Conference (MIC), Cartagena, Colombia, July 28-31, 2019.
- The 2018 International Conference on High Performance Computing & Simulation (HPCS), Dublin, Ireland, July 15–19, 2019.
- IEEE Intl. Workshop on Parallel/Distributed Computing and Optimization (IPDPS/PDCO), Rio de Janeiro, Brazil, May 20-24, 2019.
- EvoCOP'2019, 19th European Conference on Evolutionary Computation in Combinatorial Optimization, Leipzig, Germany, Apr 24-26, 2019.
- Int. Conf. on Multiple objective Programming and Goal Programming (MOPGP'2019), Marrakech, Morocco, Oct 28-31, 2019.
- Intl. Conf. on Optimization and Learning (OLA'2019), Bangkok, Thailand, Jan 29-31, 2019.
- Int. Conf. on Computing (ICC 2019), Springer's CCIS Book Series, Riyadh, Saudi Arabia, Oct 22-24, 2019.
- 15th Workshop on Foundations of Genetic Algorithms (FOGA), Potsdam, Germany, 2019.
- 10th Intl. Conference on Evolutionary Multi-criterion Optimization (EMO), Michigan, USA, 2019.
- ICANN'2019 28th Intl. Conf. on Artificial Neural Networks, Munich, Germany, Sept 2019.
- Metaheuristics Intl. Conference (MIC'2019), Cartagena, Colombia, July 2019.
- Intl. Conference on Industrial Engineering and Systems Management (IESM 2019), Shanghai, China, Sept 2019.
- 8th Intl. Conf. on Modeling, Simulation and Applied Optimizaiton (ICMSAO'2019), Bahrin, April 2019.
- Colloque sur l'Optimisation et les Systèmes d'information (COSI), Tizi-Ouzou, Algérie, Jun 24-26, 2019.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- N. Melab: Associate Editor of ACM Computing Surveys (IF: 6.131) since 2019.
- N. Melab: Guest and Managing Editor (in collaboration with P. Korosec, J. Gmys and I. Chakroun) of a special on *Synergy between Parallel Computing, Optimization and Simulation* in *Journal of Computational Science (JoCS)*, 2019.
- E-G. Talbi: Guest Editor (with L. Amodeo and F. Yalaoui) of a special issue on *Metaheuristics in Industry 4.0* in *Swarm and evolutionary computation*, 2019.
- E-G. Talbi. Guest Editor (with H. Masri) on *Multiple criteria decision making models for economic developement* in *JMCDA Journal of Multi-Criteria Decision Analysis*, 2019.
- B. Derbel: Associate Editor, IEEE Transactions on Systems, Man and Cybernetics: Systems (IEEE).

10.1.3.2. Reviewer - Reviewing Activities

- Transactions on Evolutionary Computation (IEEE TEC, IF: 8.508), IEEE.
- ACM Computing Surveys (IF: 6.131), ACM.
- Swarm and Evolutionary Computation (SWEVO, IF: 6.330), Elsevier.
- Future Generation Computer Systems (FGCS, IF: 5.7), Elsevier.
- European Journal of Operational Research [15], Elsevier.
- IEEE Transactions on Parallel and Distributed Systems (IEEE TPDS, IF: 3.4), IEEE.
- Journal of Computational Science (JoCS, IF: 2.5), Elsevier.
- Annals of Operations Research (IF: 2.284), Springer.
- Journal of Heuristics (JoH, IF: 1.392), Springer.
- Intl. Journal on Artificial Intelligence Tools (IJAIT, IF: 0.8), World Scientific.

10.1.4. Invited Talks

- E-G. Talbi. Learning-based metaheuristics, Distinguished seminar, University of Michigan, USA, Jan 2019.
- E-G. Talbi. Machine learning and optimization: inseparable disciplines, Keynote speaker, KST'2019 11^{th} Intl. Conf. on Knowledge and Smart Technology, Phuket, Thailand, Jan 2019.
- E-G. Talbi. How machine learning can help metaheuristics?, Invited seminar, American University of Sharjah, Emirates Arabs United, Feb 2019.
- E-G. Talbi. Machine learning into optimization, Invited seminar, University of Elche, Spain, Nov 2019.

10.1.5. Leadership within the Scientific Community

- N. Melab: scientific leader of Grid'5000 (https://www.grid5000.fr) at Lille, since 2004.
- E-G. Talbi: Co-president of the working group "META: Metaheuristics - Theory and applications", GDR RO and GDR MACS.
- E-G. Talbi: Co-Chair of the IEEE Task Force on Cloud Computing within the IEEE Computational Intelligence Society.
- A. Liefooghe: co-secretary of the association "Artificial Evolution" (EA).

10.1.6. Scientific Expertise

- N. Melab: Member of the advisory committee for the IT and maganement engineer training at Faculté Polytechnique de Mons, BELGIUM.

10.1.7. Research Administration

- N. Melab: Member of the Scientific Board (Bureau Scientfique du Centre) for the Inria Lille - Nord Europe research center.
- N. Melab: Member of the steering committee of "Maison de la Simulation" at Université de Lille.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- International Master lecture: N. Melab, Supercomputing, 24h ETD, M2, Université de Lille, FRANCE.
- Master lecture: N. Melab, Operations Research, 72h ETD, M1, Université de Lille, FRANCE.
- Master leading: N. Melab, Co-head (with B. Merlet) of the international Master 2 of High-perforpmance Computing and Simulation, Université de Lille, FRANCE.
- Licence: A. Liefooghe, Algorithmic and Data structure, 36h ETD, L2, Université de Lille, FRANCE.
- Licence: A. Liefooghe, Algorithmic for Operations Research, 36h ETD, L3, Université de Lille, FRANCE.
- Master: A. Liefooghe, Databases, 30h ETD, M1, Université de Lille, FRANCE.
- Master: A. Liefooghe, Advanced Object-oriented Programming, 53h ETD, M2, Université de Lille, FRANCE.
- Master: A. Liefooghe, Combinatorial Optimization, 10h ETD, M2, Université de Lille, FRANCE
- Master: A. Liefooghe, Multi-criteria Decision Aid and Optimization, 25h ETD, M2, Université de Lille, FRANCE.
- Master leading: A. Liefooghe, superviser of the Master 2 MIAGE IPI-NT.
- Master: Bilel Derbel, Combinatorial Optimization, 35h, M2, Université de Lille, FRANCE.
- Master: Bilel Derbel, Grid Computing, 16h, M2, Université de Lille, FRANCE.
- Master: Bilel Derbel, Parallel and Distributed Programming, 35h, M1, Université de Lille, FRANCE.
- Master: Bilel Derbel, Algorithms and Applications, 28h, M1, Université de Lille, FRANCE.
- Engineering school: El-Ghazali Talbi, Advanced optimization, 36h, Polytech'Lille, Université de Lille, FRANCE.
- Engineering school: El-Ghazali Talbi, Data mining, 36h, Polytech'Lille, Université de Lille, FRANCE.
- Engineering school: El-Ghazali Talbi, Operations research, 60h, Polytech'Lille, Université de Lille, FRANCE.
- Engineering school: El-Ghazali Talbi, Graphs, 25h, Polytech'Lille, Université de Lille, FRANCE.
- Master leading: B. Derbel, head of the Master MIAGE, Université de Lille, FRANCE.
- Licence: O. Abdelkafi, Computer Science, 46.5 ETD, L1, Université de Lille, FRANCE.
- Licence: O. Abdelkafi, Web Technologies, 36 ETD, L1, Université de Lille, FRANCE.
- Licence: O. Abdelkafi, Unix system introduction, 6 ETD, L2, Université de Lille, FRANCE.
- Licence: O. Abdelkafi, Web Technologies, 24 ETD, L2 S3H, Université de Lille, FRANCE.
- Licence: O. Abdelkafi, object-oriented programming, 36 ETD, L2, Université de Lille, FRANCE.
- Licence: O. Abdelkafi, Relational Databases, 36h ETD, L3, Université de Lille, FRANCE.
- Licence: O. Abdelkafi, Algorithmic for Operations Research, 36h ETD, L3, Université de Lille, FRANCE.
- Master leading: O. Abdelkafi, superviser of the Master MIAGE alternance, Université de Lille, FRANCE.

10.2.2. Supervision

- PhD: Z. Garroussi, Multi-objective matheuristics for demand-side management in smart grids, To be defended on December 21th, 2019. PhD supervised by El-Ghazali Talbi and Rachid Ellaia (EMI, MOROCCO).
- PhD in progress: J. Pelamatti, Multi-disciplinary design of aerospace vehicles, Jan 2017, El-Ghazali Talbi.
- PhD in progress: Ali Hebbal, Deep Gaussian processes and Bayesian optimization for non-stationary, multi-objective and multi-fidelity problems, Oct 2017, El-Ghazali Talbi and Nouredine Melab.
- PhD in progress (cotutelle): Maxime Gobert, Parallel multi-objective global optimization with applications to several simulation-based exlporation parameter, Oct 2018, Nouredine Melab (Université de Lille) and Daniel Tuyttens (Université de Mons, BELGIUM).
- PhD in progress (cotutelle): Guillaume Briffoteaux, Bayesian Neural Networks-assisted multi-objective evolutionary algorithms: Application to the Tuberculosis Transmission Control, Oct 2017, Nouredine Melab (Université de Lille) and Daniel Tuyttens (Université de Mons, BELGIUM).
- PhD in progress: Jeremy Sadet, Surrogate-based optimization in automotive brake design, El-Ghazali Talbi (Université de Lille), Thierry Tison (Université Polytechnique Hauts-de-France, FRANCE).
- PhD in progress: Geoffrey Pruvost, Machine learning and decomposition techniques for large-scale multi-objective optimization, Oct 2018, Bilel Derbel and Arnaud Liefooghe.
- PhD in progress: Nicolas Berveglieri, Meta-models and machine learning for massive expensive optimization, Oct 2018, Bilel Derbel and Arnaud Liefooghe.
- PhD in progress (cotutelle): Alexandre Jesus, Algorithm selection in multi-objective optimization, Bilel Derbel and Arnaud Liefooghe (Université de Lille), Luís Paquete (University of Coimbra, PORTUGAL).

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

- N. Melab: Chargé de Mission of High Performance Computing and Simulation at Université de Lille, since 2010.

10.3.2. Internal action

- A. Liefooghe: 30 minutes de sciences, Inria Lille-Nord Europe, Lille, France, June 2019.
- O. Abdelkafi: Intervention auprès des inspecteurs d'académie, Inria Lille-Nord Europe, Lille, France, Mars 2019.

11. Bibliography

Major publications by the team in recent years

[1] O. ABDELKAFI, L. IDOUMGHAR, J. LEPAGNOT.*A Survey on the Metaheuristics Applied to QAP for the Graphics Processing Units*, in "Parallel Processing Letters", 2016, vol. 26, n^{o} 3, p. 1–20

[2] A. BENDJOUDI, N. MELAB, E. TALBI.*FTH-B&B: A Fault-Tolerant HierarchicalBranch and Bound for Large ScaleUnreliable Environments*, in "IEEE Trans. Computers", 2014, vol. 63, n^{o} 9, p. 2302–2315

[3] S. CAHON, N. MELAB, E. TALBI.*ParadisEO: A Framework for the Reusable Design of Parallel and Distributed Metaheuristics*, in "J. Heuristics", 2004, vol. 10, n^{o} 3, p. 357–380

[4] F. DAOLIO, A. LIEFOOGHE, S. VEREL, H. AGUIRRE, K. TANAKA.*Problem Features versus Algorithm Performance on Rugged Multiobjective Combinatorial Fitness Landscapes*, in "Evolutionary Computation", 2017, vol. 25, n^{o} 4

[5] B. DERBEL.*Contributions to single- and multi- objective optimization: towards distributed and autonomous massive optimization*, Université de Lille, 2017, HDR dissertation

[6] B. DERBEL, A. LIEFOOGHE, Q. ZHANG, H. AGUIRRE, K. TANAKA.*Multi-objective Local Search Based on Decomposition*, in "Parallel Problem Solving from Nature - PPSN XIV - 14th International Conference, Edinburgh, UK, September 17-21, 2016, Proceedings", 2016, p. 431–441

[7] J. GMYS, M. MEZMAZ, N. MELAB, D. TUYTTENS.*IVM-based parallel branch-and-bound using hierarchical work stealing on multi-GPU systems*, in "Concurrency and Computation: Practice and Experience", 2017, vol. 29, n^{o} 9

[8] A. LIEFOOGHE, B. DERBEL, S. VEREL, H. AGUIRRE, K. TANAKA.*Towards Landscape-Aware Automatic Algorithm Configuration: Preliminary Experiments on Neutral and Rugged Landscapes*, in "Evolutionary Computation in Combinatorial Optimization - 17th European Conference, EvoCOP 2017, Amsterdam, The Netherlands, April 19-21, 2017, Proceedings", 2017, p. 215–232

[9] T. V. LUONG, N. MELAB, E. TALBI.*GPU Computing for Parallel Local Search Metaheuristic Algorithms*, in "IEEE Trans. Computers", 2013, vol. 62, n^{o} 1, p. 173–185

[10] A. NAKIB, S. OUCHRAA, N. SHVAI, L. SOUQUET, E. TALBI.*Deterministic metaheuristic based on fractal decomposition for large-scale optimization*, in "Appl. Soft Comput.", 2017, vol. 61, p. 468–485

Publications of the year

Articles in International Peer-Reviewed Journal

[11] J. S. ALMEIDA, P. P. REBOUÇAS FILHO, T. CARNEIRO, W. WEI, R. DAMAŠEVIČIUS, R. MASKELIŪNAS, V. H. C. DE ALBUQUERQUE.*Detecting Parkinson's Disease with Sustained Phonation and Speech Signals using Machine Learning Techniques*, in "Pattern Recognition Letters", July 2019, vol. 125, p. 55-62 [*DOI :* 10.1016/J.PATREC.2019.04.005], https://hal.archives-ouvertes.fr/hal-02380596

[12] L. ASLI, M. AÏDER, E.-G. TALBI.*Solving a dynamic combinatorial auctions problem by a hybrid metaheuristic based on a fuzzy dominance relation*, in "RAIRO - Operations Research", January 2019, vol. 53, n^{o} 1, p. 207-221 [*DOI :* 10.1051/RO/2018051], https://hal.archives-ouvertes.fr/hal-02304722

[13] T. CARNEIRO, J. GMYS, N. MELAB, D. TUYTTENS.*Towards ultra-scale Branch-and-Bound using a high-productivity language*, in "Future Generation Computer Systems", November 2019 [*DOI :* 10.1016/J.FUTURE.2019.11.011], https://hal.archives-ouvertes.fr/hal-02371238

[14] N. DUPIN, E.-G. TALBI.*Parallel matheuristics for the discrete unit commitment problem with min-stop ramping constraints*, in "International Transactions in Operational Research", January 2020, vol. 27, n^{o} 1, p. 219-244 [*DOI :* 10.1111/ITOR.12557], https://hal.archives-ouvertes.fr/hal-02304758

[15] J. GMYS, M. MEZMAZ, N. MELAB, D. TUYTTENS.*A computationally efficient Branch-and-Bound algorithm for the permutation flow-shop scheduling problem*, in "European Journal of Operational Research", 2020, forthcoming, https://hal.inria.fr/hal-02421229

[16] A. LIEFOOGHE, F. DAOLIO, S. VEREL, B. DERBEL, H. AGUIRRE, K. TANAKA.*Landscape-aware performance prediction for evolutionary multi-objective optimization*, in "IEEE Transactions on Evolutionary Computation", 2019, forthcoming [*DOI :* 10.1109/TEVC.2019.2940828], https://hal.archives-ouvertes.fr/hal-02294201

[17] A. NAKIB, L. SOUQUET, E.-G. TALBI.*Parallel fractal decomposition based algorithm for big continuous optimization problems*, in "Journal of Parallel and Distributed Computing", November 2019, vol. 133, p. 297-306 [*DOI :* 10.1016/J.JPDC.2018.06.002], https://hal.archives-ouvertes.fr/hal-02304882

[18] J. PELAMATTI, L. BRÉVAULT, M. BALESDENT, E.-G. TALBI, Y. GUERIN.*Efficient global optimization of constrained mixed variable problems*, in "Journal of Global Optimization", March 2019, vol. 73, n^o 3, p. 583-613 [*DOI :* 10.1007/S10898-018-0715-1], https://hal.archives-ouvertes.fr/hal-02304730

[19] O. SCHUTZE, C. HERNANDEZ, E.-G. TALBI, J.-Q. SUN, Y. NARANJANI, F.-R. XIONG.*Archivers for the representation of the set of approximate solutions for MOPs*, in "Journal of Heuristics", February 2019, vol. 25, n^o 1, p. 71-105 [*DOI :* 10.1007/S10732-018-9383-Z], https://hal.archives-ouvertes.fr/hal-02304717

[20] E.-G. TALBI.*A unified view of parallel multi-objective evolutionary algorithms*, in "Journal of Parallel and Distributed Computing", November 2019, vol. 133, p. 349-358 [*DOI :* 10.1016/J.JPDC.2018.04.012], https://hal.archives-ouvertes.fr/hal-02304734

[21] A. TCHERNYKH, U. SCHWIEGELSOHN, E.-G. TALBI, M. BABENKO.*Towards understanding uncertainty in cloud computing with risks of confidentiality, integrity, and availability*, in "Journal of computational science", September 2019, vol. 36, 100581 [*DOI :* 10.1016/J.JOCS.2016.11.011], https://hal.archives-ouvertes.fr/hal-02304771

International Conferences with Proceedings

[22] O. ABDELKAFI, B. DERBEL, A. LIEFOOGHE.*A Parallel Tabu Search for the Large-scale Quadratic Assignment Problem*, in "IEEE CEC 2019 - IEEE Congress on Evolutionary Computation", Wellington, New Zealand, June 2019, https://hal.archives-ouvertes.fr/hal-02179193

[23] N. BERVEGLIERI, B. DERBEL, A. LIEFOOGHE, H. AGUIRRE, K. TANAKA.*Surrogate-assisted multiobjective optimization based on decomposition*, in "GECCO '19 - Proceedings of the Genetic and Evolutionary Computation Conference", Prague, Czech Republic, ACM Press, July 2019, p. 507-515 [*DOI :* 10.1145/3321707.3321836], https://hal.archives-ouvertes.fr/hal-02292851

[24] T. CARNEIRO, N. MELAB.*An Incremental Parallel PGAS-based Tree Search Algorithm*, in "HPCS 2019 - International Conference on High Performance Computing & Simulation", Dublin, Ireland, July 2019, https://hal.archives-ouvertes.fr/hal-02170842

[25] T. CARNEIRO, N. MELAB.*Productivity-aware Design and Implementation of Distributed Tree-based Search Algorithms*, in "ICCS 2019 - International Conference on Computational Science", Faro, Portugal, June 2019, https://hal.archives-ouvertes.fr/hal-02139177

[26] B. DERBEL, A. LIEFOOGHE, S. VEREL, H. AGUIRRE, K. TANAKA.*New Features for Continuous Exploratory Landscape Analysis based on the SOO Tree*, in "FOGA 2019 - 15th ACM/SIGEVO Workshop on Foundations of Genetic Algorithms", Potsdam, Germany, ACM Press, August 2019, p. 72-86, https://hal.inria.fr/hal-02282986

[27] M. GOBERT, J. GMYS, J.-F. TOUBEAU, F. VALLEE, N. MELAB, D. TUYTTENS.*Surrogate-Assisted Optimization for Multi-stage Optimal Scheduling of Virtual Power Plants*, in "PaCOS 2019 - International Workshop on the Synergy of Parallel Computing, Optimization and Simulation (part of HPCS 2019)", Dublin, Ireland, July 2019, https://hal.inria.fr/hal-02178314

[28] T. ITO, H. AGUIRRE, K. TANAKA, A. LIEFOOGHE, B. DERBEL, S. VEREL.*Estimating Relevance of Variables for Effective Recombination*, in "EMO 2019 - International Conference on Evolutionary Multi-Criterion Optimization", East Lansing, Michigan, United States, February 2019, p. 411-423 [*DOI :* 10.1007/978-3-030-12598-1_33], https://hal.archives-ouvertes.fr/hal-02064547

[29] *Best Paper*
Y. MARCA, H. AGUIRRE, S. Z. MARTINEZ, A. LIEFOOGHE, B. DERBEL, S. VEREL, K. TANAKA.*Approximating Pareto Set Topology by Cubic Interpolation on Bi-objective Problems*, in "EMO 2019 - International Conference on Evolutionary Multi-Criterion Optimization", East Lansing, Michigan, United States, February 2019, p. 386-398 [*DOI :* 10.1007/978-3-030-12598-1_31], https://hal.archives-ouvertes.fr/hal-02064548.

[30] H. MONZÓN, H. AGUIRRE, S. VEREL, A. LIEFOOGHE, B. DERBEL, K. TANAKA.*Dynamic compartmental models for algorithm analysis and population size estimation*, in "Genetic and Evolutionary Computation Conference Companion", Prague, Czech Republic, Proceedings of the Genetic and Evolutionary Computation Conference Companion (GECCO '19), ACM Press, July 2019, p. 2044-2047 [*DOI :* 10.1145/3319619.3326912], https://hal.archives-ouvertes.fr/hal-02436226

[31] J. PELAMATTI, L. BRÉVAULT, M. BALESDENT, E.-G. TALBI, Y. GUERIN.*Surrogate model based optimization of constrained mixed variable problems: application to the design of a launch vehicle thrust frame*, in "SciTech 2019 - AIAA Science and Technology Forum and Exposition", San Diego, United States, American Institute of Aeronautics and Astronautics, January 2019 [*DOI :* 10.2514/6.2019-1971], https://hal.archives-ouvertes.fr/hal-02304816

[32] L. SOUQUET, A. NAKIB, E.-G. TALBI.*Deterministic multi-objective fractal decomposition algorithm*, in "MIC 2019 - 13th Metaheuristics International Conference", Cartagena, Colombia, July 2019, https://hal.archives-ouvertes.fr/hal-02304975

National Conferences with Proceeding

[33] 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

Scientific Books (or Scientific Book chapters)

[34] T. BARTZ-BEIELSTEIN, B. FILIPIČ, P. KOROŠEC, E.-G. TALBI.*High-Performance Simulation-Based Optimization*, Springer, 2020 [*DOI :* 10.1007/978-3-030-18764-4], https://hal.archives-ouvertes.fr/hal-02304686

[35] N. DUPIN, F. NIELSEN, E.-G. TALBI.*K-Medoids Clustering Is Solvable in Polynomial Time for a 2d Pareto Front*, in "Optimization of Complex Systems: Theory, Models, Algorithms and Applications", Springer, June 2020, p. 790-799 [*DOI :* 10.1007/978-3-030-21803-4_79], https://hal.archives-ouvertes.fr/hal-02304806

[36] A. LIEFOOGHE, L. PAQUETE.*Proceedings of the 19th European conference on evolutionary computation in combinatorial optimization (EvoCOP 2019)*, Lecture Notes in Computer Science, Springer, 2019, vol. 11452 [*DOI :* 10.1007/978-3-030-16711-0], https://hal.archives-ouvertes.fr/hal-02292912

[37] N. MELAB, J. GMYS, M. MEZMAZ, D. TUYTTENS.*Many-core Branch-and-Bound for GPU accelerators and MIC coprocessors*, in "High-Performance Simulation-Based Optimization", T. BARTZ-BEIELSTEIN, B. FILIPIČ, P. KOROŠEC, E.-G. TALBI (editors), Studies in Computational Intelligence, Springer, June 2019, vol. 833, 16, https://hal.inria.fr/hal-01924766

[38] J. PELAMATTI, L. BRÉVAULT, M. BALESDENT, E.-G. TALBI, Y. GUERIN.*Overview and Comparison of Gaussian Process-Based Surrogate Models for Mixed Continuous and Discrete Variables: Application on Aerospace Design Problems*, in "High-Performance Simulation-Based Optimization", Springer, June 2020, p. 189-224 [*DOI :* 10.1007/978-3-030-18764-4_9], https://hal.archives-ouvertes.fr/hal-02304707

References in notes

[39] M. BALESDENT, L. BRÉVAULT, N. B. PRICE, S. DEFOORT, R. LE RICHE, N.-H. KIM, R. T. HAFTKA, N. BÉREND.*Advanced Space Vehicle Design Taking into Account Multidisciplinary Couplings and Mixed Epistemic/Aleatory Uncertainties*, in "Space Engineering: Modeling and Optimization with Case Studies", G. FASANO, J. D. PINTÉR (editors), Springer International Publishing, 2016, p. 1–48, http://dx.doi.org/10.1007/978-3-319-41508-6_1

[40] B. DERBEL, D. BROCKHOFF, A. LIEFOOGHE, S. VEREL.*On the Impact of Multiobjective Scalarizing Functions*, in "Parallel Problem Solving from Nature - PPSN XIII - 13th International Conference, Ljubljana, Slovenia, September 13-17, 2014. Proceedings", 2014, p. 548–558

[41] B. DERBEL, A. LIEFOOGHE, G. MARQUET, E. TALBI.*A fine-grained message passing MOEA/D*, in "IEEE Congress on Evolutionary Computation, CEC 2015, Sendai, Japan, May 25-28, 2015", 2015, p. 1837–1844

[42] R. HAFTKA, D. VILLANUEVA, A. CHAUDHURI.*Parallel surrogate-assisted global optimization with expensive functions – a survey*, in "Structural and Multidisciplinary Optimization", 2016, vol. 54(1), p. 3–13

[43] D. JONES, M. SCHONLAU, W. WELCH.*Efficient Global Optimization of Expensive Black-Box Functions*, in "Journal of Global Optimization", 1998, vol. 13(4), p. 455–492

[44] J. PELAMATTI, L. BREVAULT, M. BALESDENT, E.-G. TALBI, Y. GUERIN.*How to deal with mixed-variable optimization problems: An overview of algorithms and formulations*, in "Advances in Structural and Multidisciplinary Optimization, Proc. of the 12th World Congress of Structural and Multidisciplinary Optimization (WCSMO12)", Springer, 2018, p. 64–82, http://dx.doi.org/10.1007/978-3-319-67988-4_5

[45] F. SHAHZAD, J. THIES, M. KREUTZER, T. ZEISER, G. HAGER, G. WELLEIN.*CRAFT: A library for easier application-level Checkpoint/Restart and Automatic Fault Tolerance*, in "CoRR", 2017, vol. abs/1708.02030, http://arxiv.org/abs/1708.02030

[46] N. SHAVIT.*Data Structures in the Multicore Age*, in "Communications of the ACM", 2011, vol. 54, n^{o} 3, p. 76–84

[47] M. SNIR, AL..*Addressing Failures in Exascale Computing*, in "Int. J. High Perform. Comput. Appl.", May 2014, vol. 28, n^{o} 2, p. 129–173

[48] E.-G. TALBI.*Combining metaheuristics with mathematical programming, constraint programming and machine learning*, in "Annals OR", 2016, vol. 240, n^{o} 1, p. 171–215

[49] T. VU, B. DERBEL.*Parallel Branch-and-Bound in multi-core multi-CPU multi-GPU heterogeneous environments*, in "Future Generation Comp. Syst.", 2016, vol. 56, p. 95–109

[50] X. ZHANG, Y. TIAN, R. CHENG, Y. JIN.*A Decision Variable Clustering-Based Evolutionary Algorithm for Large-Scale Many-Objective Optimization*, in "IEEE Trans. Evol. Computation", 2018, vol. 22, n^{o} 1, p. 97–112

Project-Team DEFROST

DEFormable RObotics SofTware

IN COLLABORATION WITH: Centre de Recherche en Informatique, Signal et Automatique de Lille

IN PARTNERSHIP WITH:
CNRS

Ecole Centrale de Lille

Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Robotics and Smart environments

Table of contents

Project-Team DEFROST

Creation of the Team: 2015 January 01, updated into Project-Team: 2017 November 01

Keywords:

Computer Science and Digital Science:

A2.3.3. - Real-time systems
A3.1.1. - Modeling, representation
A5.10. - Robotics
A6.2.1. - Numerical analysis of PDE and ODE
A6.2.6. - Optimization
A6.4.3. - Observability and Controlability
A6.4.4. - Stability and Stabilization

Other Research Topics and Application Domains:

B2.5.1. - Sensorimotor disabilities
B2.5.3. - Assistance for elderly
B2.7. - Medical devices
B5.1. - Factory of the future
B5.5. - Materials
B5.6. - Robotic systems
B5.7. - 3D printing
B9.2. - Art

1. Team, Visitors, External Collaborators

Research Scientists

Christian Duriez [Team leader, Inria, Senior Researcher, HDR]
Olivier Goury [Inria, Researcher]
Gang Zheng [Inria, Researcher, HDR]

Faculty Members

Jeremie Dequidt [Université de Lille, Associate Professor]
Alexandre Kruszewski [Ecole centrale de Lille, Associate Professor]

Technical Staff

Yinoussa Adagolodjo [Inria, Engineer, from Mar 2019]
Eulalie Coevoet [Inria, Engineer, until May 2019]
Stefan Escaida Navarro [Inria, Engineer, from Dec 2019]
Meichun Lin [Université de Lille, Engineer]
Bruno Carrez [Inria, Engineer]
Damien Marchal [CNRS, Engineer]
Bruno Marques [Inria, Engineer]
Thor Morales Bieze [Inria, Engineer]
Zhongkai Zhang [Inria, Engineer, until Aug 2019]

PhD Students

Walid Amehri [Inria, PhD Student]
Pierre Schegg [Robocath, PhD Student, granted by CIFRE]

Maxime Thieffry [Université de Valenciennes et du Hainaut Cambrésis, PhD Student, until Oct 2019]
Félix Vanneste [Inria, PhD Student]
Ke Wu [Inria, PhD Student, from Oct 2019]

Post-Doctoral Fellow

Stefan Escaida Navarro [Inria, Post-Doctoral Fellow, until Nov 2019]

Visiting Scientists

Margaret Koehler [Université de Lille, Jul 2019]
Van Pho Nguyen [JAIST Japan, from Apr 2019 until Sep 2019]
Federico Renda [Khalifa University, Abu Dhabi, Jun 2019]

Administrative Assistant

Anne Rejl [Inria]

2. Overall Objectives

2.1. Overall Objectives

The team DEFROST aims to address the open problem of control and modelling methods for deformable robots by answering the following challenges:

- Providing numerical methods and software support to reach the real-time constraint needed by robotic systems: the numerical solutions for the differential equations governing the deformation generate tens of thousands degrees of freedom, which is three orders of magnitude of what is frequently considered in classical methods of robotic modelling and control.
- Integrating deformation models in the control methods of soft robot: In soft-robotics, sensing, actuation and motion are coupled by the deformations. Deformable models must be placed at the heart of the control algorithm design.
- Investigating predictable interaction models with soft-tissues and parameter estimation by visual feedback from medical imaging: On the contrary too many cases in surgical robotics, the contact of the soft robot with the anatomy is permitted and it creates additional deformations on the robot.

3. Research Program

3.1. Introduction

Our research crosses different disciplines: numerical mechanics, control design, robotics, optimisation methods and clinical applications. Our organisation aims at facilitating the team work and cross-fertilisation of research results in the group. We have three objectives (1, 2 and 3) that correspond to the main scientific challenges. In addition, we have two transverse objectives that are also highly challenging: the development of a high performance software support for the project (objective 4) and the validation tools and protocols for the models and methods (objective 5).

3.2. Objective 1: Accurate model of soft robot deformation computed in finite time

The objective is to find concrete numerical solutions to the challenge of modelling soft robots with strong real-time constraints. To solve continuum mechanics equations, we will start our research with real-time FEM or equivalent methods that were developed for soft-tissue simulation. We will extend the functionalities to account for the needs of a soft-robotic system:

- Coupling with other physical phenomenons that govern the activity of sensors and actuators (hydraulic, pneumatic, electro-active polymers, shape-memory alloys...).
- Fulfilling the new computational time constraints (harder than surgical simulation for training) and find better tradeoff between cost and precision of numerical solvers using reduced-order modelling techniques with error control.
- Exploring interactive and semi-automatic optimisation methods for design based on obtained solution for fast computation on soft robot models.

3.3. Objective 2: Model based control of soft robot behavior

The focus of this objective is on obtaining a generic methodology for soft robot feedback control. Several steps are needed to design a model based control from FEM approach:

- The fundamental question of the kinematic link between actuators, sensors, effectors and contacts using the most reduced mathematical space must be carefully addressed. We need to find efficient algorithms for real-time projection of non-linear FEM models in order to pose the control problem using the only relevant parameters of the motion control.
- Intuitive remote control is obtained when the user directly controls the effector motion. To add this functionality, we need to obtain real-time inverse models of the soft robots by optimisation. Several criteria will be combined in this optimisation: effector motion control, structural stiffness of the robot, reduce intensity of the contact with the environment...
- Investigating closed-loop approaches using sensor feedback: as sensors cannot monitor all points of the deformable structure, the information provided will only be partial. We will need additional algorithms based on the FEM model to obtain the best possible treatment of the information. The final objective of these models and algorithms is to have robust and efficient feedback control strategies for soft robots. One of the main challenge here is to ensure / prove stability in closed-loop.

3.4. Objective 3: Modeling the interaction with a complex environment

Even if the inherent mechanical compliance of soft robots makes them safer, more robust and particularly adapted to interaction with fragile environments, the contact forces need to be controlled by:

- Setting up real-time modelling and the control methods needed to pilot the forces that the robot imposes on its environment and to control the robot deformations imposed by its environment. Note that if an operative task requires to apply forces on the surrounding structures, the robot must be anchored to other structures or structurally rigidified.
- Providing mechanics models of the environment that include the uncertainties on the geometry and on the mechanical properties, and are capable of being readjusted in real-time.
- Using the visual feedback of the robot behavior to adapt dynamically the models. The observation provided in the image coupled with an inverse accurate model of the robot could transform the soft robot into sensor: as the robot deforms with the contact of the surroundings, we could retrieve some missing parameters of the environment by a smart monitoring of the robot deformations.

3.5. Objective 4: Soft Robotics Software

Expected research results of this project are numerical methods and algorithms that require high-performance computing and suitability with robotic applications. There is no existing software support for such development. We propose to develop our own software, in a suite split into three applications:

- The first one will facilitate the design of deformable robots by an easy passage from CAD software (for the design of the robot) to the FEM based simulation.
- The second one is an anticipative clinical simulator. The aim is to co-design the robotic assistance with the physicians, thanks to a realistic simulation of the procedure or the robotic assistance. This will facilitate the work of reflection on new clinical approaches prior any manufacturing.
- The third one is the control design software. It will provide the real-time solutions for soft robot control developed in the project.

3.6. Objective 5: Validation and application demonstrations

The implementation of experimental validation is a key challenge for the project. On one side, we need to validate the model and control algorithms using concrete test case example in order to improve the modelling and to demonstrate the concrete feasibility of our methods. On the other side, concrete applications will also feed the reflexions on the objectives of the scientific program.

We will build our own experimental soft robots for the validation of objectives 2 and 3 when there is no existing “turn-key” solution. Designing and making our own soft robots, even if only for validation, will help the setting-up of adequate models.

For the validation of objective 4, we will develop “anatomical soft robot”: soft robot with the shape of organs, equipped with sensors (to measure the contact forces) and actuators (to be able to stiffen the walls and recreate natural motion of soft-tissues). We will progressively increase the level of realism of this novel validation set-up to come closer to the anatomical properties.

4. Application Domains

4.1. Industry

Robotics in the manufacturing industry is already widespread and is one of the strategies put in place to maintain the level of competitiveness of companies based in France and to avoid relocation to cheap labor countries. Yet, in France, it is considered that the level of robotization is insufficient, compared to Germany for instance. One of the challenges is the high investment cost for the acquisition of robotic arms. In recent years, this challenge has led to the development of “generic” and “flexible” (but rigid) robotic solutions that can be mass produced. But their applicability to specific tasks is still challenging or too costly. With the development of 3D printing, we can imagine the development of a complete opposite strategy: a “task-specific” design of robots. Given a task that needs to be performed by a deformable robot, we could optimize its shape and its structure to create the set of desired motions. A second important aspect is the reduction of the manufacturing cost: it is often predicted that the cost of deformable robots will be low compared to classical rigid robots. The robot could be built on one piece using rapid prototyping or 3D printers and be more adapted for collaborative work with operators. In this area, using soft materials is particularly convenient as they provide a mass/carried load ratio several orders of magnitude higher than traditional robots, highly decreasing the kinetic energy thus increasing the motion speed allowed in presence of humans. Moreover, the technology allows more efficient and ergonomic wearable robotic devices, opening the option for exo-skeletons to be used by human operators inside the factories and distribution centers. This remains to be put in place, but it can open new perspectives in robotic applications. A last remarkable property of soft robots is their adaptability to fragile or tortuous environments. For some particular industry fields (chemistry, food industry...) this could also be an advantage compared to existing rigid solutions. For instance, the German company http://www.festo.com, key player in the industrial robotics field, is experimenting with deformable trunk robots that exhibit great compliance and adaptability, and we are working on their accurate control.

4.2. Personal and service robotics

The personal and service robotics are considered an important source of economic expansion in the coming years. The potential applications are numerous and in particular include the challenge of finding robotic solutions for active and healthy aging at home. We plan to develop functional orthosis for which it is better not to have a rigid exoskeleton that is particularly uncomfortable. These orthosis will be ideally personalized for each patient and built using rapid prototyping. On this topic, the place of our team will be to provide algorithms for controlling the robots. We will find some partners to build these robots that would fall in the category of "wearable robots". With this thematic we also connect with a strong pole of excellence of the region on intelligent textiles (see Up-Tex) and with the strategic plan of Inria (Improving Rehabilitation and Autonomy).

4.3. Entertainment industry and arts

Robots have a long history with entertainment and arts where animatronics have been used for decades for cinematographic shootings, theater, amusement parks (Disney's audio-animatronic) and performing arts. We believe that soft robots could be a good support for art. As an example, last year we collaborated with the artist Dewi Brunet in the creation of animated origami structures (see https://dewiorigami.com/).

Figure 1. Dewi Brunet and his "Origami 2.0" exhibit at the AMV in Trélon

4.4. Medical Applications

Soft robots have many medical applications as their natural compliance makes them safer than traditional robots when interacting with humans. Such robots can be used for minimally invasive surgery, to access and act on remote parts of the body through minimal incisions in the patient. Applications include laparascopic and

brain surgery, treatment of several cancers including prostate cancer, and cardiology, for example percutaneous coronary interventions.
As an example, we received an industry grant (CIFRE) with Robocath to work on autonomous catheter navigation. See section 8.2. Another application is cochlear implant surgery.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Fondamental results

Three PhD students have defended excellent thesis in 2019:

- **Eulalie Coevoet**: Optimization-Based Inverse Model of Soft Robots With Contact Handling
- **Zhongkai Zhang**: Vision-based Calibration, Position Control and Force Sensing for Soft Robots
- **Maxime Thieffry**: Dynamic control of soft robots

In each of these thesis we presented fundamental results on the team's roadmap. Eulalie Coevoet presented the first algorithms that allow inversing the robot model in contact situations. It can be used for planning, manipulation and locomotion. Zhongkai Zhang's results allow the use of the robot as a generalized force sensor thanks to vision. We use that for feedback control for both position and force. Maxime Thieffry developped the first method for dynamic control based on model order reduction. The method is very generic and significantly improves the precision of soft robots.

5.1.2. Awards for software development

DEFROST actively contributed to the open source community by developing plugins for the SOFA framework. The team participated in the SofaWeek2019, during which the SOFA consortium organized the "Open-Source SOFA awards". One prize was offered to the candidate who developed the best open source plugin for SOFA. Another prize was offered to the best open source plugin according to the public (conference attendants). **Both prizes were won by the DEFROST team**, for the Model Order Reduction plugin and the SofaPython3 plugin respectively: link.

5.1.3. Organization of workshops and tutorials

This year, special effort was expended on the promotion of our tools through the organization of workshops and tutorials. A full tutorial day about our tools was organized at the IEEE International Conference on Soft Robotics (RobotSoft019). We then organized the first Journée de Robotique Souple in Lille with 70 participants from 9 countries. The team also participated in the organization of the (2nd Workshop on Proximity Perception in Robotics) at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2019).

5.1.4. ANR project ROBOCOP

The project ROBOCOP (ROBOtization of COchlear ImPlant) was funded by the ANR (Agence nationale de la recherche) for the development of cochlear implants for the future that are activated by electroactive polymers. The project is in collaboration with the IEMN, the LPPI, Inserm and Oticon Medical. This project will allow us to fund a PhD Student and a postdoctoral fellow for 2 years.

6. New Software and Platforms

6.1. SOFA

Simulation Open Framework Architecture

KEYWORDS: Real time - Multi-physics simulation - Medical applications

FUNCTIONAL DESCRIPTION: SOFA is an Open Source framework primarily targeted at real-time simulation, with an emphasis on medical simulation. It is mostly intended for the research community to help develop new algorithms, but can also be used as an efficient prototyping tool. Based on an advanced software architecture, it allows : the creation of complex and evolving simulations by combining new algorithms with algorithms already included in SOFA, the modification of most parameters of the simulation (deformable behavior, surface representation, solver, constraints, collision algorithm, etc.) by simply editing an XML file, the building of complex models from simpler ones using a scene-graph description, the efficient simulation of the dynamics of interacting objects using abstract equation solvers, the reuse and easy comparison of a variety of available methods.

- Participants: Christian Duriez, François Faure, Hervé Delingette and Stéphane Cotin
- Partner: IGG
- Contact: Hugo Talbot
- URL: http://www.sofa-framework.org

6.2. SoftRobots

SoftRobots plugin for Sofa

KEYWORDS: Numerical simulations - Problem inverse - Soft robotics

FUNCTIONAL DESCRIPTION: This plugin allows the modeling of deformable robots in the Sofa platform. It allows the modeling of different actuators, such as cable, pneumatic pressure, hydraulics and other simpler types of actuation. It also contains useful tools for animation design or communication with the robot. Coupled with the SoftRobots.Inverse plugin, it also allows the control of these robots. More information can be found on the dedicated website.

- Participants: Christian Duriez, Olivier Goury, Jérémie Dequidt, Damien Marchal, Eulalie Coevoet and Félix Vanneste
- Contact: Christian Duriez
- URL: https://project.inria.fr/softrobot/

6.3. Model Order Reduction Plugin for SOFA

KEYWORDS: Model Order Reduction - Sofa - Finite element modelling

SCIENTIFIC DESCRIPTION: This plugin allows speed-up of SOFA simulations by providing tools to create a reduced version of the SOFA simulation that runs at much higher rates but remains accurate. Starting with a snapshot of the object deformations on a high-dimensional Finite Element mesh, Proper Orthogonal Decomposition (POD) is used to compute a reduced basis of small dimension representing correctly all the possible deformations of the object. The original system describing the object motion is then greatly reduced. To keep numerical efficiency, a hyper-reduction method is used to speed-up the construction of the reduced system.

FUNCTIONAL DESCRIPTION: This plugin allows to dramatically reduce computational time in mechanical simulation in the SOFA framework. A reduced simulation, of much smaller dimension but still accurate is created in an automatic way by the plugin. Building the reduced model may take time, but this operation is made once only. The user can then benefit from a reduced and interactive version of his/her simulation without significant loss of accuracy.

RELEASE FUNCTIONAL DESCRIPTION: This is the first version of the plugin.

NEWS OF THE YEAR: Publication using this plugin accepted dans IEEE Transactions on Robotics

- Participants: Olivier Goury, Félix Vanneste, Christian Duriez and Eulalie Coevoet
- Contact: Olivier Goury
- Publication: Fast, generic and reliable control and simulation of soft robots using model order reduction
- URL: https://project.inria.fr/modelorderreduction/

6.4. SoftRobots.Inverse

KEYWORDS: Sofa - SoftRobots

FUNCTIONAL DESCRIPTION: This plugin builds on the plugin SoftRobots. Inside the plugin, there is some constraint components that are used to describe the robot (effectors, actuators, sensors). An optimisation algorithm is provided to find the efforts to put on actuators in order to place the robot in a the closest possible configuration than the one described by "effectors", or to a state described by "sensors". This method used to control the soft-robots in the task space is patented.

- Partners: CNRS - Université de Lille - Ecole Centrale de Lille
- Contact: Christian Duriez
- URL: https://project.inria.fr/softrobot.inverse

6.5. SofaPython3

KEYWORDS: Python - Numerical simulations - Sofa

FUNCTIONAL DESCRIPTION: This plugin allows to use Sofa as a library from any python3 program. It also allows to write new mechanical component for a Sofa simulation in python3.

- Contact: Christian Duriez
- URL: https://github.com/SofaDefrost/plugin.SofaPython3/

6.6. SofaQtQuick

runSofa2

KEYWORDS: Sofa - GUI (Graphical User Interface) - Modeling - Physical simulation

FUNCTIONAL DESCRIPTION: Smooth the user experience with Sofa By integrating authoring features into runSofa so we can design simulation in an integrated environment. We should be able to model scenes, simulate & debug them.

This tool replaces the old "runSofa" interface, today deprecated but still in use by most SOFA users.

SofaQtQuick provides a fluid and dynamic user experience for SOFA, thanks to the integration of authoring tools to design complex simulations directly in the 3D environment, rather that scripting them as it is done today.

FEATURES:

*Scene graph editing

*Interactive modeling

*Project oriented approach

*Prefab as reusable and parametric object

*2D Canvas

*Custom widgets per component

*Live coding

*Node base interface for data link debugging

*Everything with a non-linear workflow

Based on a code gift from Anatoscope, stringly inspired by Blender & Unity's workflow.

RELEASE FUNCTIONAL DESCRIPTION: 1st Beta version, unstable, but testable.

- Contact: Christian Duriez
- URL: http://github.com/SofaDefrost/SofaQtQuick

7. New Results

7.1. Soft robots locomotion and manipulation control using FEM simulation and quadratic programming

In this work, we proposed a method to control the motion of soft robots able to manipulate objects or roll from one place to another. We used the Finite Element Method (FEM) to simulate the deformations of the soft robot, its actuators, and its environment. To find the inverse model of the robot interacting with obstacles, and with constraints on its actuators, we wrote the problem as a quadratic program with complementarity constraints. The novelty of this work was that friction contacts (sticking contact only) is taken into account in the optimization process, allowing the control of these specific tasks that are locomotion and manipulation. We proposed a formulation that simplifies the optimization problem, together with a dedicated solver [22]. The algorithm had real-time performance and handles evolving environments as long as we know them. To show the effectiveness of the method, we presented several numerical examples, and a demonstration on a real robot (see Figure 2 and 3).

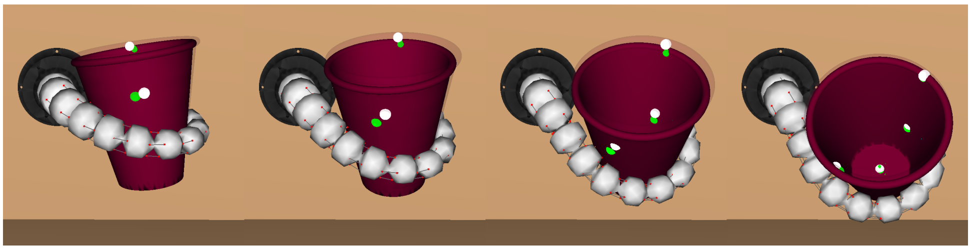

Figure 2. Simulation of a soft gripper holding a deformable cup subject to gravity. Here we optimize the cables displacements to control the position/orientation of the cup. A phantom of the cup target is shown in transparency. The cup have four controlled points represented by the green spheres. The corresponding targets are represented by the white spheres.

7.2. Toward Shape Optimization of Soft Robots

This year, we obtained new results on shape optimization for soft robotics where the shape is optimized for a given soft robot usage. To obtain a parametric optimization with a reduced number of parameters, we relied on an approach where the designer progressively refines the parameter space and the fitness function until a satisfactory design is obtained. In our approach, we automatically generate FEM simulations of the soft robot and its environment to evaluate a fitness function while checking the consistency of the solution. Finally, we have coupled our framework to an evolutionary optimization algorithm, and demonstrated its use for optimizing the design of a deformable leg of a locomotive robot. A paper presenting the approach was accepted at IEEE/International Conference On Soft-Robotics2019 [29].

Figure 3. Real soft robot actuated online using the output of the simulation. In this scenario, using our control framework, we are able to control the orientation of the real plastic cup (see the attached video).

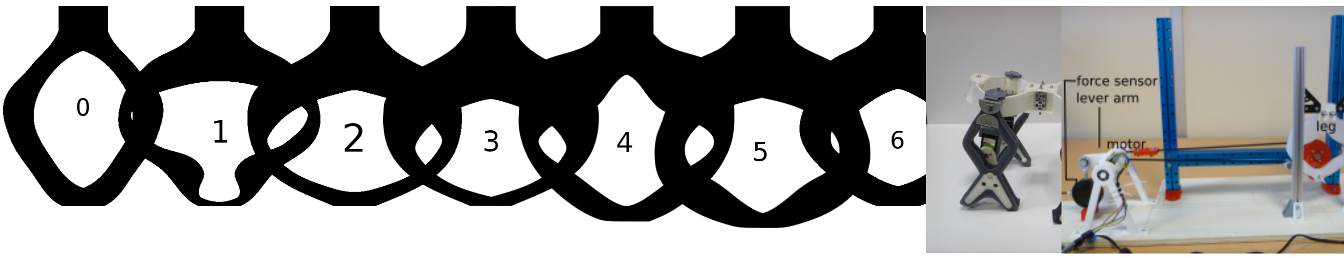

Figure 4. Initial shape (labelled 0) for the optimization and different shapes (labelled 1 to 6) obtained by numerical experiments. Pictures represents the robot with the legs and the test best that has been used to verify the accuracy of the simulation

7.3. Modeling Novel Soft Mechanosensors based on Air-Flow Measurements

In this work, we introduce a new pneumatic mechanosensor dedicated to Soft Robotics and propose a generic method to reconstruct the magnitude of a contact-force acting on it. This is illustrated by Fig. 5. Changes in cavity volumes inside a soft silicon pad are measured by air-flow sensors. The resulting mechanosensor is characterized by its high sensitivity, repeatability, dynamic range and accurate localization capability in 2D. Using a regression found by machine learning techniques we can predict the contact location and force magnitude accurately when the force magnitudes are within the range of the training data. To be able to provide a more general model, a novel approach based on a Finite Element Method (FEM) is introduced. We formulate an optimization problem, which yields the contact load that best explains the observed changes in cavity volumes. This method makes no assumptions on the force range, the shape of the soft pad or the shape of its cavities. The prediction of the force also results in a model for the deformation of the soft pad. We characterize our sensor and evaluate two designs, a soft pad and a kidney-shaped sensor, in different scenarios. A paper was accepted for the journal *Robotics and Automation Letters (RA-L)* [5]. Furthermore, an extended abstract was accepted at the *RoboTac 2019 Workshop* at *IROS 2019*, leading to a presentation of a demo of the proposed technology.

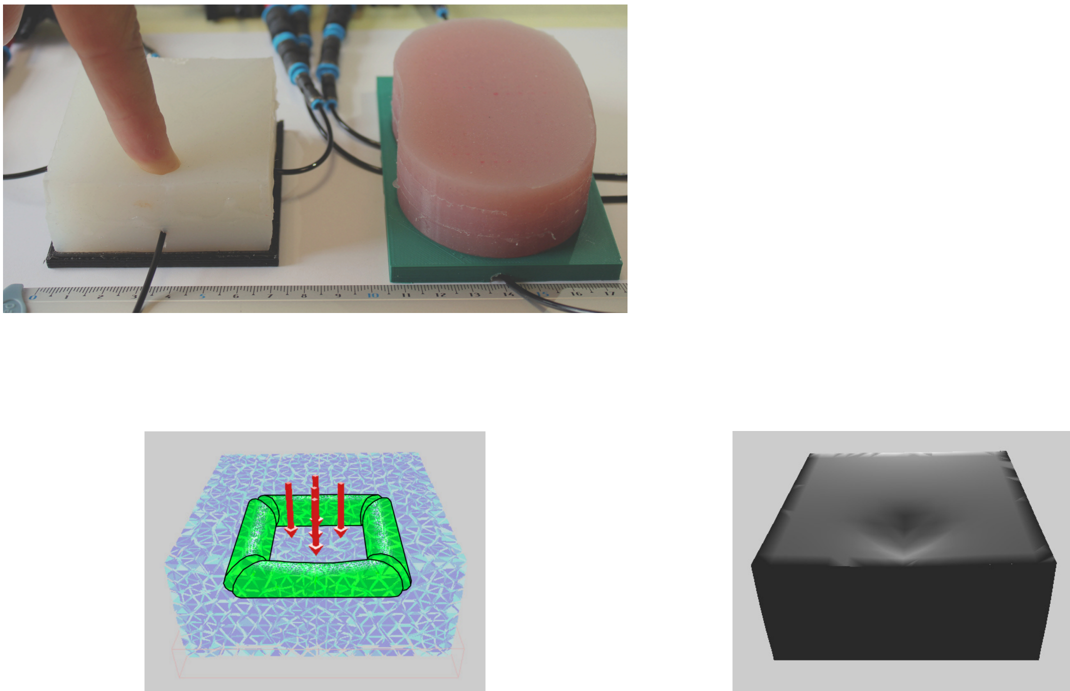

Figure 5. In this work, we show two designs of a novel soft mechanosensor made out of silicone (top, a soft pad and a kidney). When an external force is applied, the volume of cavities embedded in the silicone changes (left). This change in volume is registered through air-flow sensors. Using machine learning and FEM-based techniques, we show that it is possible to estimate the location and magnitude of an external force on the mechanosensor. Using the FEM also yields an estimation of the deformation of the sensor (left and right).

7.4. Calibration and External Force Sensing for Soft Robots using an RGB-D Camera

Benefiting from the deformability of soft robots, calibration and force sensing for soft robots are possible using an external vision-based system, instead of embedded mechatronic force sensors. In this work, we first propose a calibration method to calibrate both the sensor-robot coordinate system and the actuator inputs. This task is addressed through a sequential optimization problem for both variables. We also introduce an external force sensing system based on a real-time Finite Element (FE) model with the assumption of static configurations, and which consists of two steps: force location detection and force intensity computation. The algorithm that estimates force location relies on the segmentation of the point cloud acquired by an RGB-D camera. Then, the force intensities can be computed by solving an inverse quasi-static problem based on matching the FE model with the point cloud of the soft robot. As for validation, the proposed strategies for calibration and force sensing have been tested using a parallel soft robot driven by four cables (see figure 6 and reference [18]).

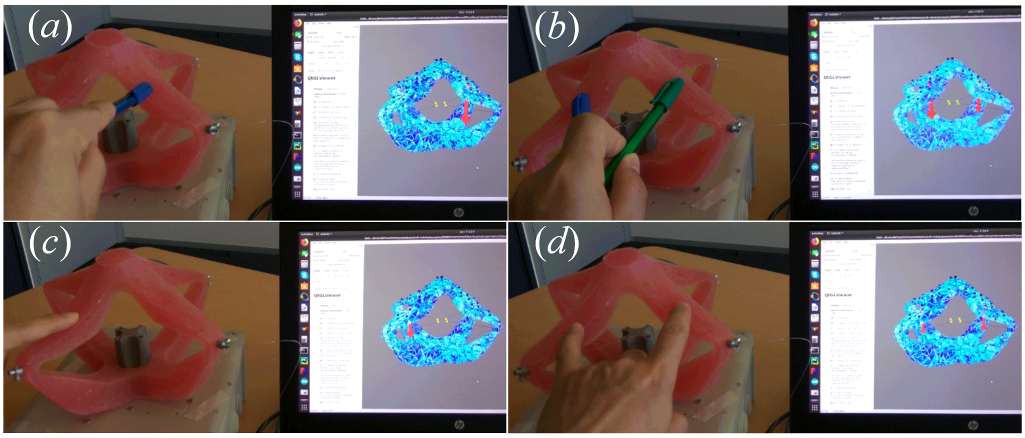

Figure 6. Screenshot of external force sensing. The robot has four cables with constant length for the experiment. (a) and (c) show one external force on the actuated soft robot. (b) and (d) show case with two external forces.

7.5. Motion Control of Cable-Driven Continuum Catheter Robot through Contacts

Catheter-based intervention plays an important role in minimally invasive surgery. For the closed-loop control of catheter robot through contacts, the loss of contact sensing along the entire catheter might result in task failure. To deal with this problem, we propose a decoupled motion control strategy which allows to control insertion and bending independently. We model the catheter robot and the contacts using the Finite Element Method. Then, we combine the simulated system and the real system for the closed-loop motion control. The control inputs are computed by solving a quadratic programming (QP) problem with a linear complementarity problem (LCP). A simplified method is proposed to solve this optimization problem by converting it into a standard QP problem. Using the proposed strategy, not only the control inputs but also the contact forces along the entire catheter can be computed without using force sensors. Finally, we validate the proposed methods using both simulation and experiments on a cable-driven continuum catheter robot for the real-time motion control through contacts [17].

7.6. Control Design for Soft Robots based on Reduced Order Model

Inspired by nature, soft robots promise disruptive advances in robotics. Soft robots are naturally compliant and exhibit nonlinear behavior, which makes their study challenging. No unified framework exists to control these

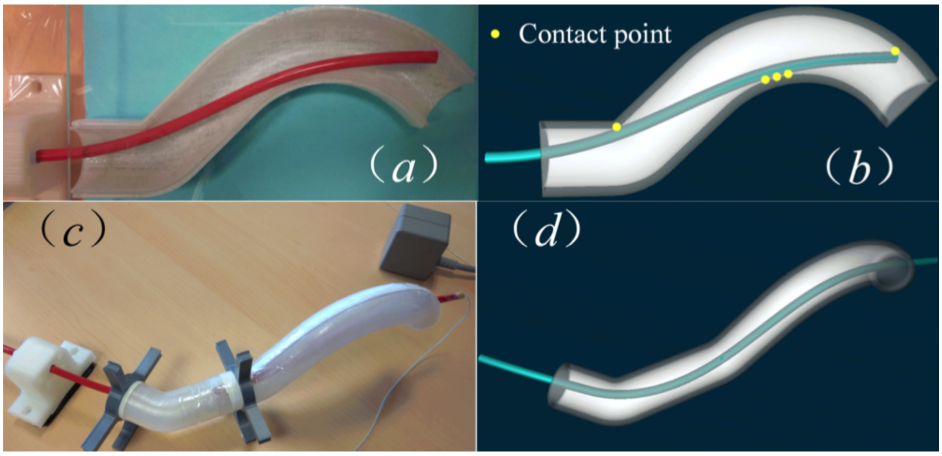

Figure 7. (a) and (c) present experimental setups for the validation. (b) and (d) show that we take into account the contact and provide accurate results in the simulation

robots, especially when considering their dynamics. This work proposes a methodology to study this type of robots around a stable equilibrium point. It can make the robot converge faster and with reduced oscillations to a desired equilibrium state. Using computational mechanics, a large-scale dynamic model of the robot is obtained and model reduction algorithms enable the design of low order controller and observer. A real robot is used to demonstrate the interest of the results [14].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

We would like to acknowledge FACEBOOK company for the donation of $25,000 for our research (department FACEBOOK Reality Labs).

8.2. Bilateral Grants with Industry

We received an industry grant (CIFRE) with Robocath to work on autonomous catheter navigation. This grant will fund a PhD student for 3 years, starting in February 2019.

We have an ongoing bilateral project with the company InSimo on the simulation of suture.

9. Partnerships and Cooperations

9.1. Regional Initiatives

- **INVENTOR** Innovative tool for soft robot design and its application for surgery. This project is financed by **I-Site ULNE EXPAND**, supported by “le programme d’Investissements d’Avenir” and “la Métropole Européenne de Lille”. The objective of this project is to develop an innovative tool for the facilitation of soft robot design.

- **COMOROS** Control of deformable robots for surgery Duration april 2017 to march 2020 Program: FEDER Coordinator: C. Duriez Abstract: Surgical procedures are often carried out using instruments made of stiff materials that interact with delicate biological tissues such as internal organs, blood vessel walls and small cavities. This incompatibility of stiffness is one of the sources of danger in many surgical procedures. The use of robots made of soft materials, also called soft robots, would limit such risks by reducing contact pressures and stress concentrations. Their intrinsic deformability would also increase the ability to manoeuvre in confined spaces. However, the promising concept of using soft robots for surgical procedures cannot be practically implemented, due to the lack of precise modelling and control methods for soft robots. This scientific obstacle, identified as a pending issue by major surveys in this field, becomes particularly challenging when interacting with an environment as complex as the human anatomy. Drawing on our background in soft tissue simulation, contact models, surgical applications and soft robotics, our ambition in this project is to:
  - Develop accurate and generic numerical methods for continuum mechanics, adapted to strong real-time constraints in order to demonstrate the ability to model soft mechatronics systems.
  - Reconsider parametrization methodologies of digital models of the patient anatomy through the observation of mechanical interactions with soft robots via embedded sensors and medical imaging
  - Rethink motion generation and teleoperation control with force feedback so as to be compatible with the large number of degrees of freedom of soft robots and be based on accurate, rapidly-computed deformable models and interaction models.

  The project also targets the development of software with the required performance and features, as well as the experimental validation of models and methods using prototypes in realistic environments.
- The PhD Thesis of Félix Vanneste is half-funded by the Hauts-de-France region.

9.2. National Initiatives

- **ROBOCOP**: Robotization of Cochlear implant. This is a 4-year project, supported by the ANR (French National Agency for Research) in the framework of PRCE, starting from 1 October 2019 until 30 September 2023. ROBOCOP aims at creating a new prototype of cochlear implant, and robotize (i.e. actuate and control) its insertion process to facilitate the work of surgeon, to increase the success ratio, and to decrease the probability of trauma.
- **SIMILAR** Soft robotIcs framework for modeling, simulation and control. This project is supported by **Inria ADT**, and the objective is to design new 3D interactive software to design soft-robots. This new software will be on the top of our existing software stack relying on SOFA for all numerical simulation aspects and 3D rendering aspects.
- **Tremplin ERC** Christian Duriez received a **ANR** grant "tremplin ERC" (150k€) given the result obtained last year on the ERC proposal (evaluated at "grade A"). The project has allowed to allocate new resources on the developments that were presented in this ERC.

9.3. European Initiatives

9.3.1. Collaborations in European Programs, Except FP7 & H2020

Meichun Lin was doing a project belonged to Interreg - 2 Seas Mers Zeeën on Cooperate Brachytherapy(CoBra), it is a 4 years project which gathers the experts from the countries between English Channel and southern North Sea aiming on finding an advance method for curing prostate cancer. (see more details on https://cobra-2seas.eu/) The project is divided by several fields which are - MR compatible robot design, radiation dose measurement, steerable needle design, mimic soft-tissue (phantom) design and virtual reality real-time training tool development etc. Meichun was working on developing virtual reality real-time training tool with Defrost team. The aim is to have a interactive platform for human and the robot. By using SOFA framework to simulate the soft tissue's deformation and the interaction with needle insertion under the real-time, also with the Image modelling of MRI and soft-tissue modelling and so on and so forth.

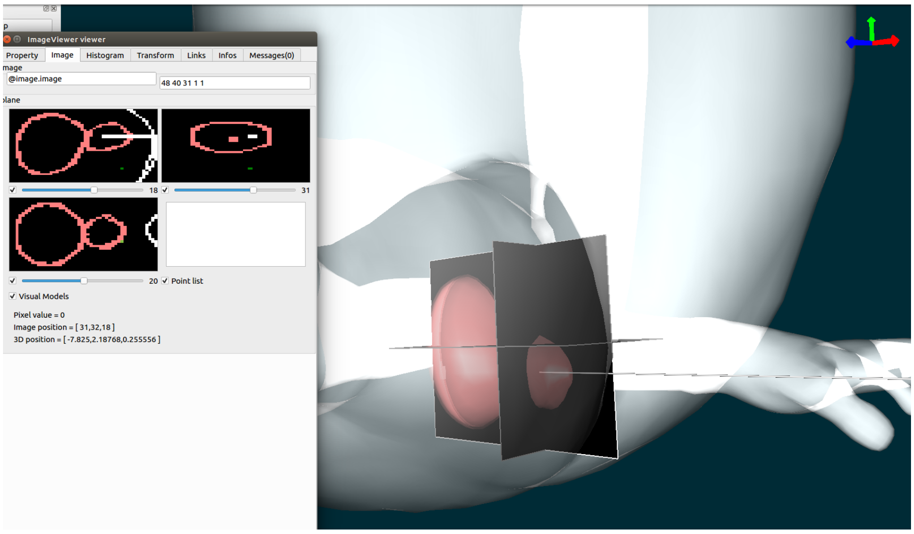

Figure 8. The virtual reality real-time simulation of the CoBra project

9.4. International Initiatives

9.4.1. Inria International Labs

9.4.1.1. AC/DC: A Charm lab / Defrost team Collaboration

Inria@SiliconValley Associate Team

Defrost team (Deformable Robotic Software, Inria Lille – Nord Europe) and the Charm Lab (Collaborative HAptics and Robotics in Medicine Lab, Stanford University, USA) on the topic of soft robots. On this topic, these two entities are very complementary because the Charm Lab is interested in the new design, the realization, the planning and the experimentation and the Defrost team is more centered on mechanical modeling, simulation and the algorithms of control. The collaboration is based on two axes: (1) the creation of flexible robots whose position and rigidity can be controlled, (2) the mechanical modeling and simulation of a robot that navigates in an environment through growth.

- Partner: Allison Okamura at the Department of Mechanical Engineering of Stanford University, USA
- Start year: 2019
- See also: https://team.inria.fr/defrost/collaboration-with-charm-lab-stanford/

9.5. International Research Visitors

9.5.1. Visits of International Scientists

Federico Renda from Khalifa University of Abu Dhabi visited the DEFROST team for a month to work on the implementation of a Cosserat Implementation for Beam simulation in the SOFA framework.

9.5.2. Internships

- Van Pho Nguyen, PhD Candidate from Japan Advanced Institute of Science and Technology (JAIST), visited the team for 6 months to work on the topic of underwater robots.
- Margaret Koehler from the Charm Lab, Stanford University, USA, visited the team for a month to work on the simulation of a soft haptic device.

9.5.3. Visits to International Teams

9.5.3.1. Research Stays Abroad

- Gang Zheng has visited Nanjing University of Science and Technology (China) for 1 month in July 2019.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- Gang Zheng is member of Bureau ED (Ecole Doctorale) SPI 072 in the domain of AGITSI (Automatique, Génie Informatique Traitement du Signal et des Images), 2020-2024.
- Gang Zheng is member of Bureau Scientifique (ex-BCP) of Inria Lille – Nord Europe, from February 2019.
- Gang Zheng was a vice-chair of the IFAC Technical Committee “Social Impact of Automation”, International Federation of Automatic Control, TC9.2, till 2019
- Gang Zheng is co-chair of the working group “Commande et pilotage en environnement incertain” of GRAISYHM
- Christian Duriez was chair of the workshop “ Modeling, Simulation and Control of Deformable Robots on SOFA Framework” organized during Robosoft Conference 2019 https://team.inria.fr/defrost/software/defrost-platform-modeling-simulation-and-control-of-deformable-robots-on-sofa-framework/

10.1.1.2. Member of the Organizing Committees

- This year, Stefan Escaida Navarro was a Co-Organizer of the *2nd Workshop on Proximity Perception in Robotics* at *IROS 2019* in Macao, China. [0]
- Gang Zheng, Associate Editor, SIAM CT19, Chengdu, China (SIAM Conference of Control & Its Applications 2019)
- Christian Duriez is member of the organizing committee of ICRA 2020 (in Paris), chair of Social Media and Community.

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

- Gang Zheng is IPC member of SIAM19, ICSRT19, ICFCTA19, ICRAI19, AITC19, APCRAS20, ICSRT20, ETFA20, ICSC20.
- Jeremie Dequidt is IPC member of International Symposium on Visual Computing 2019
- Christian Duriez is IPC of Robosoft 2019 and Robosoft 2020

10.1.2.2. Reviewer

Alexandre Kruszewski was revewier for:

- 2020 21st IFAC World Congress
- 2020 3rd IEEE International Conference on Soft Robotics
- 2020 American Control Conference
- 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems
- 2019 IEEE Conference on Decision and Control
- 2019 IEEE International Conference on Fuzzy Systems
- 2019 Chinese Control and Decison Conference

[0] https://www.proxelsandtaxels.org/en/

Olivier Goury was reviewer for:

- 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
- IEEE International Conference on Soft Robotics (ROBOSOFT 2020)

Stefan Escaida Navarro was reviewer for:

- 2020 IEEE International Conference on Robotics and Automation (ICRA 2020)
- 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2019)

Jeremie Dequidt was reviewer for:

- 2019 International Symposium on Visual Computing
- 2020 IEEE Conference on Virtual Reality and 3D User Interfaces
- 2019 IEEE International Conference on Robotics and Automation (ICRA 2019)
- 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2019)

Christian Duriez was reviewer for:

- 2019 IEEE International Conference on Robotics and Automation (ICRA 2019)
- 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2019)
- Conference Robotics Science and System (RSS 2019)
- Siggraph Conference 2019

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Christian Duriez is associate editor of:

- IEEE Transactions on Haptics
- IEEE Robotics and Automation Letters

10.1.3.2. Reviewer - Reviewing Activities

Olivier Goury was reviewer for:

- ACM CHI Conference on Human Factors in Computing Systems
- International Journal of Robotics Research (IJRR)
- IEEE Robotics and Automation Letters

Stefan Escaida Navarro was reviewer for:

- IEEE Robotics and Automation Letters (RA-L)
- IEEE Sensors Journal

Alexandre Kruszewski was revewier for:

- IEEE Transaction on Fuzzy Systems
- IEEE Robotics and Automation Letters
- IEEE Transactions on Vehicular Technology
- International Journal of Robotics Research
- Systems & Control Letters
- Fuzzy Sets and Systems
- IET Control Theory & Applications

Jeremie Dequidt was reviewer for:

- IEEE Transaction on Haptics
- IEEE Robotics and Automation
- IEEE Robotics and Automation Letters

Christian Duriez was reviewer for:

- Nature Robotics
- International Journal of Robotic Research
- IEEE Robotics and Automation Letters
- IEEE Transactions on Graphics

10.1.4. Invited Talks

Christian Duriez was invited for:

- Keynote at IEEE Robosoft Conference
- Keynote during scientific day of FEMTO (Besançon)
- Keynote & Practice during the 9th Summer School on Surgical Robotics
- Keynote during the FOOR 2019 (Forum Ouvert Oeuvre et Recherche) in Lille

10.1.5. Scientific Expertise

Christian Duriez is expert of the European Community for monitoring the FET project HybridHeart

10.1.6. Research Administration

- Christian Duriez has been nominated Director of the Inria Lille – Nord Europe center for an interim of 3 months (July to September). He is also the president of the "Commission des Emplois de Recherche" (Research Jobs Commission)
- Olivier Goury is an elected member of the "Comité de centre" and a member of the "Comité de développement technologique" (CDT) at Inria Lille – Nord Europe.
- Alexandre Kruszewski is member of the Laboratory council (CRIStAL).
- Damien Marchal is lead of the "Pôle d'Appui au Développement et à la Recherche" of the CRIStAL Laboratory.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Engineering cycle: Walid Amehri, Start & Go Arts et Sciences, 24h, level L3, Centrale Lille.

Engineering cycle: Walid Amehri, Start & Go Conception et Environement, 10h, level L3, Centrale Lille.

Engineering cycle: Walid Amehri, Automatique IE3, 20h, level L3, Centrale Lille.

Engineering cycle: Alexandre Kruszewski, 14 modules (automatic control, numeric control, embeded systems, robotics etc.), ≈250h, level (L3, M1, M2), Centrale Lille

Engineering cycle: Jeremie Dequidt 6 modules (Programming, Software Engineering, Embedded Systems, Databases, Medical Simulation etc.) ≈350h, level (L3, M1, M2), Polytech Lille

Master: Christian Duriez, Soft robotics, 24h, M2, Graduate degree en intelligence artificielle à l'Ecole Polytechnique (Palaiseau)

Master: Christian Duriez, Interactive simulation, 20h, M2, Master Image Visualisation Interaction, Université de Lille

10.2.2. Supervision

PhD: Eulalie Coevoet, Optimization Based Inverse Model of Soft Robots With Contact Handling, Université de Lille, 9/01/2019, C. Duriez

PhD: Zhongkai Zhang, Vision-based calibration, position control and force sensing for soft robots, Université de Lille, 10/01/2019, J. Dequidt, C. Duriez

PhD: Maxime Thieffry, Modélisation et contrôle de robots déformables à grande vitesse, UPHF, 16/10/2019, A. Kruszewski, C. Duriez, T.M. Guerra

PhD in progress: Walid Amehri, Workspace analysis of soft robots, G. Zheng, A. Kruszewski

PhD in progress: Ke Wu, Control of soft robot under constraints, G. Zheng

PhD in progress: Pierre Schegg, Catheter Navigation using Reinforcement Learning, J. Dequidt and C. Duriez

PhD in progress: Félix Vanneste, Design and simulation of Soft Robots made of mesostructured materials, 01/12/2018, C Duriez, O. Goury

10.2.3. Juries

Gang Zheng has participated the following juries of thesis:

- Imen Mrad, Observabilité et inversion à gauche des systèmes dynamiques hybrides, defended at 16/12/2019, Université de Cergy-Pontoise (Reviewer);
- Yanqiao Wei, Non-asymptotic and robust fractional order differentiators using generalized modulating functions, defended at 15/11/2019, INSA Val de Loire (Examinator);
- Saber Laamiri, Commande des systèmes électriques: Machines synchrones et convertisseurs multiniveaux, defended at 27/09/2019, Ecole centrale de Nantes (Examinator).

Christian Duriez was member of the following juries of PhD thesis:

- Margaret Koehler, Model-Based Design And Control Of Deformable Robots And Haptic Devices, Stanford University, USA, defended on the 14th of November 2019 (Reviewer)
- François Schmitt, Méthodes et procédés pour l'assistance à la chirurgie laparoscopique par comanipulation, University of Strasbourg, defended the 30th of September 2019 (Reviewer)
- Quentin Peyron, Concept de robot à tube concentrique magnétique: introduction et analyse, defended the 5th of December 2019 (President of the Jury)

Christian Duriez was member of the following juries for Habilitation thesis:

- Sinan Haliyo, Interactions Multi-Echelles, defended the 6th of December 2019 (Reviewer)
- Guillaume Caron, Vision Robotique Directe, defended the 10th of December 2019 (President of the Jury)

10.3. Popularization

- http://handsonsoftrobotics.lille.inria.fr/
- SIDO 2019: SIDO is leading European Solutions and Technologies event for IoT, Artificial Intelligence and Robotics for strategic decision-makers in innovation and business operations. Jeremie Dequidt has presented current research works arund Soft Roboctics.

10.3.1. Interventions

- Fête de la science 2019: Bruno Carrez and Olivier Goury took part in the " village des sciences" at the "Gare Saint-Sauveur" in Lille city center to present soft robots to several groups of students from middle school and high school which represents about 150 children over 2 days between October 10th and 11th 2019.
- FOSDEM 2019 Brussels: its an open-software manifestation which happen every years during 2 days. Some people of our team (Damiens Marchal, Eulalie Coevoet, Bruno Marques, Bruno Carrez and Félix Vanneste) went there to help the SOFA consortium to make some advertisement of our simulation tools.
- ISN intervention: Olivier Goury presented the field of soft robotics to high school students taking a Computer Science option (Informatiques sciences du numérique) and exchanged about the school projects on March the 20th 2019.

- Défi Robotique at Lycée Baggio: Olivier Goury took part in the Défi Robotique at Lycée Baggio in which teams of middle school, high school, preparatory classes and engineering students have 24 hours to solve a robotics challenge. Olivier Goury gave a conference on Soft Robotics and was a member of jury.
- Journée de la recherche Centrale Lille: Alexandre Kruszewski, Thor Morales-Bieze and Bruno Carrez presented the research activities (with live demos) in a showroom dedicated to the promotion of the scientific activities of the Laboratory of the campus.

11. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] E. COEVOET.*Optimization-based inverse model of soft robots, with contact handling*, Université de Lille 1, Sciences et Technologies, January 2019, https://hal.archives-ouvertes.fr/tel-02446416

[2] M. THIEFFRY.*Model-Based Dynamic Control of Soft Robots*, Université Polytechnique des Hauts-de-France, October 2019, https://hal.archives-ouvertes.fr/tel-02363267

[3] Z. ZHANG.*Vision-based calibration, position control and force sensing for soft robots*, Université de Lille, January 2019, https://hal.archives-ouvertes.fr/tel-01990867

Articles in International Peer-Reviewed Journal

[4] B.-H. DU, A. POLYAKOV, G. ZHENG, Q. QUAN.*Quadrotor trajectory tracking by using fixed-time differentiator*, in "International Journal of Control", 2019, vol. 2019 - Issue 12 [*DOI :* 10.1080/00207179.2018.1462534], https://hal.inria.fr/hal-01947365

[5] S. ESCAIDA NAVARRO, O. GOURY, G. ZHENG, T. MORALES BIEZE, C. DURIEZ.*Modeling Novel Soft Mechanosensors based on Air-Flow Measurements*, in "IEEE Robotics and Automation Letters", 2019, vol. 4, n^{o} 4, p. 4338 - 4345 [*DOI :* 10.1109/LRA.2019.2932604], https://hal.inria.fr/hal-02239080

[6] Z. JIANG, O. MAYEUR, J.-F. WITZ, P. LECOMTE-GROSBRAS, J. DEQUIDT, M. COSSON, C. DURIEZ, M. BRIEU.*Virtual image correlation of magnetic resonance images for 3D geometric modelling of pelvic organs*, in "Strain", January 2019, vol. 55, n^{o} 3, e12305 [*DOI :* 10.1111/STR.12305], https://hal.archives-ouvertes.fr/hal-02403391

[7] M. KOEHLER, A. M. OKAMURA, C. DURIEZ.*Stiffness Control of Deformable Robots Using Finite Element Modeling*, in "IEEE Robotics and Automation Letters", April 2019, vol. 4, n^{o} 2, p. 469-476 [*DOI :* 10.1109/LRA.2019.2890897], https://hal.inria.fr/hal-02079146

[8] S. MAALEJ, A. KRUSZEWSKI, L. BELKOURA.*Derivative-Based Sampled Data Control for Continuous Linear Parameter Varying System With Unknown Parameters*, in "Journal of Dynamic Systems, Measurement, and Control", August 2019, vol. 141, n^{o} 8 [*DOI :* 10.1115/1.4042947], https://hal.archives-ouvertes.fr/hal-02403902

[9] M. MBOUP, C. JOIN, M. FLIESS, Y. WANG, G. ZHENG, D. EFIMOV, W. PERRUQUETTI.*Comments on "Differentiator application in altitude control for an indoor blimp robot"*, in "International Journal of Control", 2020, forthcoming [*DOI :* 10.1080/00207179.2018.1500040], https://hal.inria.fr/hal-01832874

[10] J. REN, L. YU, C. LYU, G. ZHENG, H. SUN, J.-P. BARBOT.*Dynamical Sparse Signal Recovery with Fixed-Time Convergence*, in "Signal Processing", 2019, vol. 162, p. 65-74, https://hal.inria.fr/hal-02406715

[11] C. ROGNON, M. KOEHLER, C. DURIEZ, D. FLOREANO, A. M. OKAMURA.*Soft Haptic Device to Render the Sensation of Flying Like a Drone*, in "IEEE Robotics and Automation Letters", July 2019, vol. 4, n^{o} 3, p. 2524-2531 [*DOI :* 10.1109/LRA.2019.2907432], https://hal.archives-ouvertes.fr/hal-02400542

[12] W. SAADI, D. BOUTAT, G. ZHENG, L. SBITA, L. YU.*Algorithm to compute nonlinear partial observer normal form with multiple outputs*, in "IEEE Transactions on Automatic Control", 2019, forthcoming, https://hal.inria.fr/hal-02406825

[13] C. SONG, H. WANG, Y. TIAN, G. ZHENG.*Event-triggered observer design for delayed output-sampled systems*, in "IEEE Transactions on Automatic Control", 2019, forthcoming, https://hal.inria.fr/hal-02406816

[14] M. THIEFFRY, A. KRUSZEWSKI, C. DURIEZ, T.-M. GUERRA.*Control Design for Soft Robots based on Reduced Order Model*, in "IEEE Robotics and Automation Letters", January 2019, vol. 4, n^{o} 1, p. 25-32 [*DOI :* 10.1109/LRA.2018.2876734], https://hal.archives-ouvertes.fr/hal-01901031

[15] M. THIEFFRY, A. KRUSZEWSKI, T.-M. GUERRA, C. DURIEZ.*Trajectory Tracking Control Design for Large-Scale Linear Dynamical Systems With Applications to Soft Robotics*, in "IEEE Transactions on Control Systems Technology", 2019, p. 1-11, forthcoming [*DOI :* 10.1109/TCST.2019.2953624], https://hal.archives-ouvertes.fr/hal-02404003

[16] Y. WANG, G. ZHENG, D. EFIMOV, W. PERRUQUETTI.*Disturbance Compensation Based Controller for an Indoor Blimp Robot*, in "Robotics and Autonomous Systems", 2019, forthcoming, https://hal.inria.fr/hal-02406793

[17] Z. ZHANG, J. DEQUIDT, J. BACK, H. LIU, C. DURIEZ.*Motion Control of Cable-Driven Continuum Catheter Robot through Contacts*, in "IEEE Robotics and Automation Letters", February 2019, vol. 4, n^{o} 2, p. 1852-1859 [*DOI :* 10.1109/LRA.2019.2898047], https://hal.archives-ouvertes.fr/hal-02052637

[18] Z. ZHANG, A. PETIT, J. DEQUIDT, C. DURIEZ.*Calibration and External Force Sensing for Soft Robots using an RGB-D Camera*, in "IEEE Robotics and Automation Letters", March 2019, vol. 4, n^{o} 3, p. 2356 - 2363 [*DOI :* 10.1109/LRA.2019.2903356], https://hal.archives-ouvertes.fr/hal-02060976

[19] G. ZHENG, H. WANG.*Finite-time estimation for linear time-delay systems via homogeneous method*, in "International Journal of Control", 2019, vol. 92, n^{o} 6, p. 1252–1263 [*DOI :* 10.1080/00207179.2017.1390255], https://hal.inria.fr/hal-01649434

[20] G. ZHENG, Y. ZHOU, M. JU.*Robust control of a silicone soft robot using neural networks*, in "ISA Transactions", 2019, forthcoming, https://hal.inria.fr/hal-02406765

International Conferences with Proceedings

[21] F. J. BEJARANO, R. USHIROBIRA, G. ZHENG.*On the observability and detectability of non-commensurate time-delay linear systems*, in "2019 18th European Control Conference (ECC)", Naples, Italy, IEEE, June 2019, p. 2921-2925 [*DOI :* 10.23919/ECC.2019.8796285], https://hal.inria.fr/hal-02406860

[22] E. COEVOET, A. ESCANDE, C. DURIEZ.*Soft robots locomotion and manipulation control using FEM simulation and quadratic programming*, in "RoboSoft 2019 - IEEE International Conference on Soft Robotics", Seoul, South Korea, April 2019, https://hal.inria.fr/hal-02079151

[23] S. WANG, A. POLYAKOV, G. ZHENG.*Quadrotor Control Design under Time and State Constraints: Implicit Lyapunov Function Approach*, in "ECC'19 - European Control Conference", Naples, Italy, June 2019, https://hal.inria.fr/hal-02093641

[24] Y. WANG, G. ZHENG, D. EFIMOV, W. PERRUQUETTI.*Disturbance Compensation Based Control for an Indoor Blimp Robot*, in "2019 International Conference on Robotics and Automation (ICRA)", Montreal, Canada, IEEE, May 2019, p. 2040-2046 [*DOI :* 10.1109/ICRA.2019.8793535], https://hal.inria.fr/hal-02406847

[25] Y. ZHOU, M. JU, G. ZHENG.*Closed-loop control of soft robot based on machine learning*, in "2019 Chinese Control Conference (CCC)", Guangzhou, France, IEEE, July 2019, p. 4543-4547 [*DOI :* 10.23919/CHICC.2019.8866257], https://hal.inria.fr/hal-02406854

Conferences without Proceedings

[26] A. BERNARDIN, C. DURIEZ, M. MARCHAL.*An Interactive Physically-based Model for Active Suction Phenomenon Simulation*, in "SWS19 - SOFA Week Symposium", Paris, France, November 2019, https://hal.inria.fr/hal-02419381

[27] R. K. KATZSCHMANN, M. THIEFFRY, O. GOURY, A. KRUSZEWSKI, T.-M. GUERRA, C. DURIEZ, D. RUS.*Dynamically Closed-Loop Controlled Soft Robotic Arm using a Reduced Order Finite Element Model with State Observer*, in "IEEE 2019 International Conference on Soft Robotics", Séoul, South Korea, April 2019, https://hal.archives-ouvertes.fr/hal-02078809

[28] S. MAALEJ, A. KRUSZEWSKI, L. BELKOURA.*Model-free control of LTI systems: An algebraic state estimation approach*, in "2017 18th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)", Monastir, Tunisia, IEEE, August 2019, vol. 141, n^{o} 8, p. 30-35 [*DOI :* 10.1109/STA.2017.8314951], https://hal.archives-ouvertes.fr/hal-02403910

[29] T. MORZADEC, D. MARCHAL, C. DURIEZ.*Toward Shape Optimization of Soft Robots*, in "RoboSoft 2019 - IEEE International Conference on Soft Robotics", Séoul, South Korea, April 2019, https://hal.archives-ouvertes.fr/hal-02078776

[30] G. ZHENG, O. GOURY, M. THIEFFRY, A. KRUSZEWSKI, C. DURIEZ.*Controllability pre-verification of silicone soft robots based on finite-element method*, in "2019 International Conference on Robotics and Automation (ICRA)", Montreal, Canada, IEEE, May 2019, p. 7395-7400 [*DOI :* 10.1109/ICRA.2019.8794370], https://hal.archives-ouvertes.fr/hal-02403899

Other Publications

[31] R. DAGHER, G. ZHENG, A. QUADRAT.*General closed-form solutions of the position self-calibration problem*, October 2019, Paper under submission, https://hal.inria.fr/hal-02419854

Project-Team FUN

self-organizing Future Ubiquitous Network

RESEARCH CENTER
Lille - Nord Europe

THEME
Networks and Telecommunications

Table of contents

Project-Team FUN

Creation of the Team: 2012 January 01, updated into Project-Team: 2013 July 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
A1.2.7. - Cyber-physical systems
A1.2.8. - Network security
A1.4. - Ubiquitous Systems
A5.10.6. - Swarm robotics

Other Research Topics and Application Domains:

B5.1. - Factory of the future
B5.6. - Robotic systems
B5.9. - Industrial maintenance
B6.4. - Internet of things
B7. - Transport and logistics
B8. - Smart Cities and Territories

1. Team, Visitors, External Collaborators

Research Scientists

Nathalie Mitton [Team leader, Inria, Senior Researcher, HDR]
Valeria Loscri [Inria, Researcher, HDR]

Faculty Member

Antoine Gallais [Université de Strasbourg, Associate Professor, until Aug 2019, HDR]

Technical Staff

Ibrahim Amadou [Inria, Engineer]
Rehan Malak [Inria, Engineer, until Mar 2019]
Matthieu Berthome [Inria, Engineer, until Aug 2019]
Julien Vandaele [Inria]

PhD Students

Emilie Bout [Inria, from Oct 2019]
Brandon Foubert [Inria]
Mohammad Hussein Ghosn [Lebanese International University, from Mar 2019]
Meysam Mayahi [Inria, from Oct 2019]
Carola Rizza [Inria, from Oct 2019]
Edward Staddon [Inria, from Oct 2019]

Post-Doctoral Fellows

Allan Blanchard [Inria, until Jun 2019]
Antonio Costanzo [Inria, from Dec 2019]

Christian Salim [Inria, from Sep 2019]

Visiting Scientist

Noura Mares [Ecole Nationale des Ingénieurs de Tunis, from Jun 2019 until Jul 2019]

Administrative Assistant

Anne Rejl [Inria]

2. Overall Objectives

2.1. Overall Objectives

Context.

The Internet of Things [41] is a large concept with multiple definitions. However, the main definitions are the same in every vision and could be summed up as follows: *Imagine a world where every object has the capacity to communicate with its environment. Everything can be both analogue and digitally approached - reformulates our relationship with objects - things - as well as the objects themselves. Any object relates not only to you, but also to other objects, relations or values in a database. In this world, you are no longer alone, anywhere.* (Internet of Things council).

Future Ubiquitous Networks (FUN) are part of the Internet of Things. They are composed of tens to thousands heterogeneous hardware-constrained devices that interact with our environment and the physical world. These devices have limited resources in terms of storage and computing capacities and energy. They communicate through unreliable and unpredictable short-range wireless links and run on batteries that are not envisaged to be changed in current systems since generally deployed in hostile environments. Providing FUNs with energy saving protocols is thus a key issue. Due to these specific features, any centralized control is not conceivable, the new generation of FUNs must be autonomous, self-organized and dynamically adapting to their environment. The devices that compose CPNs can be sensors, small robots, RFID readers or tags.

Objects or things can now communicate with their environment through the use for instance of an RFID (Radio Frequency IDentification) tag that provides them a unique identifier (ID) and a way to communicate through radio waves.

In the case of a simple passive **RFID tag**, the thing only embeds a tag equipped with an antenna and some memory. To communicate, it needs to be powered by the electromagnetic field of an RFID reader. This reader may then broadcast the information read on tag over a network.

When this tag is equipped with a battery, it is now able to communicate with nearby things similar to itself that may relay its message. Tags can also be equipped with additional capacity and sensors (for light, temperature, etc.). The Internet of Things can thus now refer to a **wireless sensor** network in which each sensor sends the data it collects over its environment and then sends it to a sink, *i.e.* a special sensor node able to analyze those data. In every case, RFID tags or sensor nodes can **be moved unexpectedly** like hold by moving things or animals. We speak then about **'*undergone* mobility'**.

So far, things can thus communicate information about their environment. But when the capacity of sensors is extended even further, they can also act on their environment (for instance, the detection of an event (fire) may trigger an action like switching the light or fire hoses on). Sensor nodes become **actuators**. When this extended capacity is the faculty to move, actuators are also referred as actors or robots. In this latter case, the mobility is computed on purpose, we then speak about **'*controlled* mobility'**. Actuators are not moved but move by themselves.

The FUN research group aims to focus on self-organizing techniques for these heterogeneous Future Ubiquitous Networks (FUNs). FUNs need various self-organization techniques to work properly. Self-organization encompasses neighbor discovery (which what other devices a sensor/actuator can communicate directly?), communication, self-deployment, self-localization, activity scheduling (when to wake up, when to send data to save energy without being detrimental to the well behavior of the network, etc.)...

Solutions provided by FUN should facilitate the use of FUNs and rub away heterogeneity and difficulties. These techniques should be **scalable**, **energy-aware**, **standard-compliant**, should manage undergone **mobility** and take advantage of controlled mobility when available [53].

Solutions provided by FUN will consider vagaries of the realistic wireless environment by integrating cross-layer techniques in their design.

Motivation.

To date, many self-organizing techniques for wireless sensor networks and mobile ad hoc networks are proposed in the literature and also by the FUN research group. Some of them are very efficient for routing [46], [40], discovering neighborhood [51], [50], scheduling activity and coverage [48], localization [54], [38], [29], etc. Nevertheless, to the best of our knowledge, most of them **have not been validated by experimentation**, only by simulation and thus cannot consider the real impact of the wireless links and real **node mobility** in different environments. In addition, some of them rely on assumptions that are known not to be true in realistic networks such as the fact that the transmission range of a node is a perfect disk. Other may perform well only when nodes are static. None of them considers to **take advantage of controlled mobility** to enhance performances. Similarly, many propositions arise regarding self-organization in RFID networks, mainly at the middleware level [58], [47] and at the MAC layer level [52]. Although these latter propositions are generally experimented, they are validated only in static environments with very few tags and readers. To fit realistic features, such algorithms should also be evaluated with regards to scalability and mobility.

RFID and sensor/actor technologies **have not been merged**. Though, RFID readers may now be mobile and communicate in a wireless peer-to-peer manner either with other RFID readers or wireless sensor nodes and all belong to the same network. This implies a study of the standards to allow inter-dependencies in a transparent manner. Although such works have been initiated inside EPC Global working groups, research actions remain scarce.

FUN research group aims at **filling this scientific gap** by proposing self-stabilizing solutions, considering vagaries of wireless links, node mobility and heterogeneity of nodes in compliance with current standards. Validation by experimentation is mandatory to prove the effectiveness of proposed techniques in realistic environments.

FUN will investigate new protocols and communication paradigms that allow the **transparent merging** of technologies. Objects and events might interconnect while **respecting on-going standards** and building an autonomic and smart network being compliant with hardware resources and environment. FUN expects to rub away the difficulty of use and programmability of such networks by unifying the different technologies. In addition, FUN does not only expect to validate the proposed solutions through experimentation, but also to learn from these experiments and from the observation of the impact of the wireless environment, to take these features into consideration in the design of future solutions.

3. Research Program

3.1. Introduction

We will focus on wireless ubiquitous networks that rely on constrained devices, i.e. with limited resources in terms of storage and computing capacities. They can be sensors, small robots, RFID readers or tags. A wireless sensor retrieves a physical measure such as light. A wireless robot is a wireless sensor that in addition has the ability to move by itself in a controlled way. A drone is a robot with the ability to manoeuvre in 3D (in the air or in the water). RFID tags are passive items that embed a unique identifier for a place or an object allowing accurate traceability. They can communicate only in the vicinity of an RFID reader. An RFID reader can be seen as a special kind of sensor in the network which data is the one read on tags. These devices may run on batteries that are not envisaged to be changed or recharged. These networks may be composed of ten to thousands of such heterogeneous devices for which energy is a key issue.

Today, most of these networks are homogeneous, i.e. composed of only one kind of devices. They have mainly been studied in application and technology silos. Because of this, they are approaching fundamental limitations especially in terms of topology deployment, management and communications, while exploiting the complementarity of heterogeneous devices and communication technologies would enlarge their capacities and the set of applications. Finally, these networks must work efficiently even in dynamic and realistic situations, i.e. they must consider by design the different dynamic parameters and automatically self-adapt to their variations.

Our overall goal is represented by Figure 1. We will investigate wireless ubiquitous IoT services for constrained devices by smartly combining **different frequency bands** and **different medium access and routing techniques** over **heterogeneous devices** in a **distributed** and **opportunistic** fashion. Our approach will always deal with **hardware constraints** and take care of **security** and **energy** issues to provide protocols that ride on **synergy** and **self-organization** between devices.

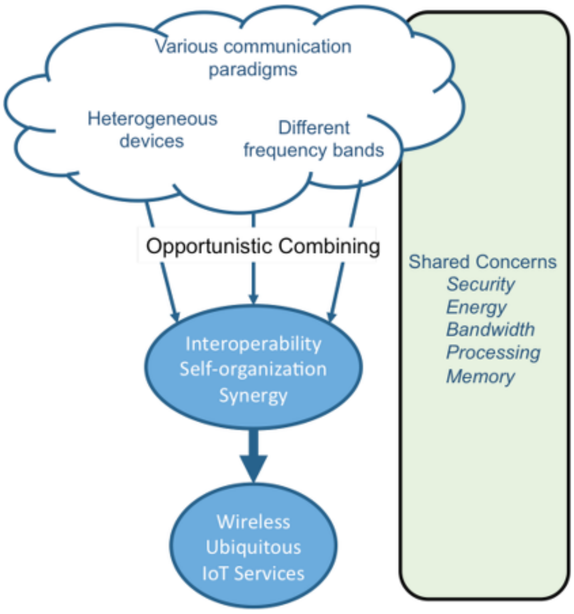

Figure 1. FUN's overal goal.

The goal of the FUN project team is to provide these next generation networks with a set of innovative and distributed self-organizing cooperative protocols to raise them to a new level of scalability, autonomy, adaptability, manageability and performance. We aim to break these silos to exploit the full synergy between devices, making them cooperate in a single holistic network. We will consider them as networks of heterogeneous devices rather than a collection of heterogeneous networks.

To realize the full potential of these ubiquitous networks, there is a need to provide them with a set of tools that allow them to *(i)* (self-)deploy, *(ii)* self-organize, *(iii)* discover and locate each other, resources and services and *(iv)* communicate. These tools will be the basics for enabling cooperation, co-existence and witnessing a global efficient behavior. The deployment of these mechanisms is challenging since it should be achieved in spite of several limitations. The main difficulties are to provide such protocols in a **secured** and **energy-efficient** fashion in spite of:

- dynamic topology changes due to various factors such as the unreliability of the wireless medium, the wireless interferences between devices, node mobility and energy saving mechanisms;

- hardware constraints in terms of CPU and memory capacities that limit the operations and data each node can perform/collect;
- lacks of interoperability between applicative, hardware and technological silos that may prevent from data exchange between different devices.

3.1.1. Objectives and methodology

To reach our overall goal, we will pursue the two following objectives. These two objectives are orthogonal and can be carried on jointly:

1. Providing realistic complete self-organizing tools *e.g. vertical perspective*.
2. Going to heterogeneous energy-efficient performing wireless networks *e.g. horizontal perspective*.

We give more details on these two objectives below. To achieve our main objectives, we will mainly apply the methodology depicted in Figure 2 combining both theoretical analysis and experimental validation. Mathematical tools will allow us to properly dimension a problem, formally define its limitations and needs to provide suitable protocols in response. Then, they will allow us to qualify the outcome solutions before we validate and stress them in real scenarios with regards to applications requirements. For this, we will realize proofs-of-concept with real scenarios and real devices. Differences between results and expectations will be analyzed in return in order to well understand them and integrate them by design for a better protocol self-adaptation capability.

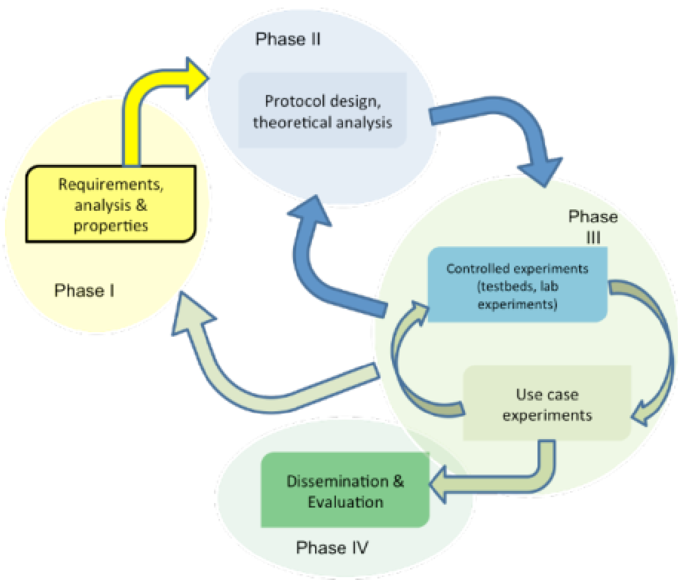

Figure 2. Methodology to be applied in FUN.

3.2. Vertical Perspective

As mentioned, future ubiquitous networks evolve in dynamic and unpredictable environments. Also, they can be used in a large scope of applications that have several expectations in terms of performance and different contextual limitations. In this heterogeneous context, IoT devices must support multiple applications and relay traffic with non-deterministic pattern.

To make our solutions practical and efficient in real conditions, we will adopt the dual approach both *top-down* and *bottom-up*. The *top-down* approach will ensure that we consider the application (such as throughput, delay, energy consumption, etc.) and environmental limitations (such as deployment constraints, etc.). The *bottom-up* approach will ensure that we take account of the physical and hardware characteristics such as memory, CPU, energy capacities but also physical interferences and obstacles. With this integrated perpective, we will be in capacity to design well adapted **cross-layer** integrated protocols [59]. We will design jointly routing and MAC layers by taking dynamics occurring at the physical layer into account with a constant concern for energy and security. We will investigate new adaptive frequency hopping techniques combined with routing protocols [59], [45].

This vision will also allow us to integrate external factors by design in our protocols, in an opportunistic way. Yet, we will leverage on the occurrence of any of these phenomena rather than perceiving them as obstacles or limitations. As an example, we will rely on node undergone mobility to enhance routing performance as we have started to investigate in [55], [37]. On the same idea, when specific features are available like controlled mobility, we will exploit it to improve connectivity or coverage quality like in [49], [57], [31], [25].

3.3. Horizontal perspective

We aim at designing efficient tools for a plethora of wireless devices supporting highly heterogeneous technologies. We will thus investigate these networks from a horizontal perspective, e.g. by considering heterogeneity in low level communications layers.

Given the spectrum scarcity, they will probably need to coexist in the same frequency bands and sometimes for different purposes (RFID tag reading may use the same frequency bands as the wireless sensors). One important aspect to consider in this setting is how these different access technologies will interact with each other, and what are the mechanisms needed to be put in place to guarantee that all services obtain the required share of resources when needed. This problem appears in different application domains, ranging from traffic offloading to unlicensed bands by cellular networks and the need to coexist with WiFi and radars, from a scenario in which multiple-purpose IoT clouds coexist in a city [56]. We will thus explore the dynamics of these interactions and devise ways to ensure smooth coexistence while considering the heterogeneity of the devices involved, the access mechanisms used as well as the requirements of the services provided.

To face the spectrum scarcity, we will also investigate new alternative communication paradigms such as phonon-based or light-based communications as we have initiated in [42] and we will work on the coexistence of these technologies with traditional communication techniques, specifically by investigating efficient switching techniques from one communication technology to the other (they were most focused on the security aspects, to prevent jamming attacks). Resilience and reliability of the whole system will be the key factors to be taken into account [43], [39], [18].

As a more prospective activity, we consider exploring software and communication security for IoT. This is challenging given that existing solutions do not address systems that are both constrained and networked [44]. Finally, in order to contribute to a better interoperability between all these technologies, we will continue to contribute to standardization bodies such as IETF and EPC Global.

4. Highlights of the Year

4.1. Highlights of the Year

4.1.1. Awards

Paper [17] has been awarded best paper of the 34th ACM/SIGAPP Symposium On Applied Computing (SAC 2019).
BEST PAPERS AWARDS :
[17]

A. BLANCHARD, N. KOSMATOV, F. LOULERGUE.*Logic against Ghosts: Comparison of Two Proof Approaches for a List Module*, in "SAC 2019 - The 34th ACM/SIGAPP Symposium On Applied Computing", Limassol, Cyprus, April 2019 [*DOI :* 10.1145/3297280.3297495], https://hal.inria.fr/hal-02100515

5. New Software and Platforms

5.1. AspireRFID ALE

FUNCTIONAL DESCRIPTION: AspireRFID middleware is a modular OW2 open source RFID middleware. It is compliant with EPC Global standards. This new module integrates the modifications of the new standard release, including new RP and LLRP definitions and fixing bugs. This module has been implemented in the framework of the MIAOU project.

- Participants: Ibrahim Amadou, Julien Vandaele, Nathalie Mitton and Rim Driss
- Contact: Nathalie Mitton

5.2. ETINODE-CONTIKI-PORT

KEYWORD: Iot

FUNCTIONAL DESCRIPTION: Contiki is an open source embedded OS for Internet of Things (IoT). It is light and portable to different hardware architectures. It embeds communication stacks for IoT. This driver allows the running of Contiki OS over Etinode-MSP430. The code dalso allows the use of radio chip and embedded sensors. This module has been implemented in the framework of the ETIPOPS project.

- Participants: Nathalie Mitton, Roudy Dagher and Salvatore Guzzo Bonifacio
- Contact: Salvatore Guzzo Bonifacio

5.3. ETINODE-DRIVERS

FUNCTIONAL DESCRIPTION: These drivers for Etinode-MSP430 control the different embedded sensors and hardware components available on an Etinode-MSP430 node such as gyroscope, accelerometer and barometric sensor. This module has been implemented in the framework of the ETIPOPS project.

- Participants: Nathalie Mitton, Roudy Dagher and Salvatore Guzzo Bonifacio
- Contact: Salvatore Guzzo Bonifacio

5.4. EVe-TCF

Embedded Verifier for Transitive Control Flow

KEYWORDS: Control Flow - JavaCard - Embedded systems - Embedded - Security - Code analysis

FUNCTIONAL DESCRIPTION: Verification of transitive control flow policies on JavaCard 2.x bytecode. Control flow policies expressed using a DSL language are embedded in JavaCard packages (CAP files) using EVe-TCF convert tool. Control flow policies are then statically verified on-device at loading-time thanks to an embedded verifier (designed for smart cards in EVe-TCF). EVe-TCF (Embedded Verifier for Transitive Control Flow) also contains an off-device (i.e. PC tool) to simulate on-device loading process of JavaCard 2.x platforms with GlobalPlatform 2.x installed.

- Participants: Arnaud Fontaine and Isablle Simplot Ryl
- Contact: Nathalie Mitton

5.5. GOLIATH

Generic Optimized LIghtweight communication stack for Ambient TecHnologies

KEYWORDS: WSN - WSN430

FUNCTIONAL DESCRIPTION: GOLIATH (Generic Optimized LIghtweight communication stack for Ambient TecHnologies) is a full protocol stack for wireless sensor networks. This module has been implemented in the framework of the ETIPOPS project.

- Participants: David Simplot Ryl, Fadila Khadar, Nathalie Mitton and Salvatore Guzzo Bonifacio
- Contact: Nathalie Mitton
- URL: https://gforge.inria.fr/projects/goliath/

5.6. 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/

5.7. T-SCAN

KEYWORDS: Rfid - RFID Middleware

FUNCTIONAL DESCRIPTION: T-Scan is an interface ensuring the translation from a SGTIN tag format to an ONS hostname format according to the EPCGlobal standards. It allows the sending of a DNS request to look up the EPC-IS aides to which the product belongs in order to access the data relative to that product. This module has been implemented in the framework of the TRACAVERRE project.

- Participants: Gabriele Sabatino and Nathalie Mitton
- Contact: Gabriele Sabatino

5.8. FIT IoT-Lab

Participants: Nathalie Mitton [contact person], Julien Vandaele, Matthieu Berthome.

FIT IoT-LAB (http://www.iot-lab.info) is a very large scale open testbed that features over 2700 wireless sensor nodes and more than 200 robots spread across six different sites in France. Nodes are either fixed or mobile and can be allocated in various topologies throughout all sites. A variety of wireless sensors are available, with different processor architectures (MSP430, STM32 and Cortex-A8) and different wireless chips (802.15.4 PHY at 800 MHz or 2.4 GHz). In addition, "open nodes" can receive custom wireless sensors for inclusion in IoT-LAB testbed. This platform is completely open and can be used by any one wishing to run experiment on wireless sensors and robots.

The Lille site displays 2 subsets of the platforms:

- Haute Borne: this site features 256 M3 sensor nodes operating in the 2.4GHz band and 64 mobile robots (32 turtlebots and 32 wifibots) completely remotely programmable.
- Opennodes: this site features 64 hardware open slots to allow any one to plug his own hardware and benefits from the platform debugging and monitoring tools.

6. New Results

6.1. Security and Verification

Participants: Rehan Malak, Allan Blanchard, Antoine Gallais, Valeria Loscri, Nathalie Mitton.

6.1.1. Security

Numerous medium access control (MAC) have been proposed for Low-power Lossy Networks (LLNs) over the recent years. They aim at ensuring both energy efficiency and robustness of the communication transmissions. Nowadays, we observe deployments of LLNs for potentially critical application scenarios (e.g., plant monitoring, building automation), which require both determinism and security guarantees. They involve battery-powered devices which communicate over lossy wireless links. Radio interfaces are turned off by a node as soon as no traffic is to be sent or relayed. Denial-of-sleep attacks consist in exhausting the devices by forcing them to keep their radio on. In [21], we focus on jamming attacks whose impact can be mitigated by approaches such as time-division and channel hopping techniques. We use the IEEE 802.15.4e standard to show that such approaches manage to be resistant to jamming but yet remain vulnerable to selective jamming. We discuss the potential impacts of such onslaughts, depending on the knowledge gained by the attacker, and to what extent envisioned protections may allow jamming attacks to be handled at upper layers.

6.1.2. Verification

Modern verification projects continue to offer new challenges for formal verification. One of them is the linked list module of Contiki, a popular open-source operating system for the Internet of Things. It has a rich API and uses a particular list representation that make it different from the classical linked list implementations. Being widely used in the OS, the list module is critical for reliability and security. A recent work verified the list module using ghost arrays. In [17], [35], we report on a new verification effort for this module. Realized in the Frama-C/Wp tool, the new approach relies on logic lists. A logic list provides a convenient high-level view of the linked list. The specifications of all functions are now proved faster and almost all automatically, only a small number of auxiliary lemmas and a couple of assertions being proved interactively in Coq. The proposed specifications are validated by proving a few client functions manipulating lists. During the verification, a more efficient implementation for one function was found and verified. We compare the new approach with the previous effort based on ghost arrays, and discuss the benefits and drawbacks of both techniques.

While deductive verification is increasingly used on real-life code, making it fully automatic remains difficult. The development of powerful SMT solvers has improved the situation, but some proofs still require interactive theorem provers in order to achieve full formal verification. Auto-active verification relies on additional guiding annotations (assertions, ghost code, lemma functions, etc.) and provides an important step towards a greater automation of the proof. However, the support of this methodology often remains partial and depends on the verification tool. [18] presents an experience report on a complete functional verification of several C programs from the literature and real-life code using auto-active verification with the C software analysis platform Frama-C and its deductive verification plugin . The goal is to use automatic solvers to verify properties that are classically verified with interactive provers. Based on our experience, we discuss the benefits of this methodology and the current limitations of the tool, as well as proposals of new features to overcome them.

6.2. Visible Light Communication

Participants: Antonio Costanzo, Valeria Loscri.

Visible Light Communication (VLC) exploits optical frequencies, diffused by usual LED lamps, for adding data communication features to illuminating systems. This paradigm has attracted a growing interest in both scientific and industrial community in the latter decade. Nevertheless, classical wireless communication mechanisms for physical and Medium Access Control (MAC) layers are hardly available for VLC, due to the massive external interference caused by sunlight. Moreover, effects related to the data frames features need to be taken into account in order to improve the effectiveness of the VLC paradigm. Such as an instance, the preamble length of a packet in order to synchronize the data transmission represents an important factor in VLC. A too long preamble allows a better synchronization while impacting negatively in terms of overhead. Nevertheless, a too short preamble may be not effective for synchronizing the transmission. The more suitable selection of the preamble length is strictly related to the noise environments. In order to make an adaptive selection able to choice the more suitable preamble length, we have designed and integrated in our VLC system a machine learning algorithm based on multi-arm bandit approach, in order to dynamically select the best configuration [28]. Another important approach to face the high interference impacting on the VLC performance is represented by the treatment of the noise through a signal processing approach in order to estimate it and proceed with a mitigation of the noise [10], [29], [36], [30]. This approach has been implemented and tested by the means of a real prototype. Results obtained show the effectiveness of a similar approach.

6.3. Alternative communications

Participant: Valeria Loscri.

In the last few years, there has been an increasing interest in the study of "alternative communication" paradigms, ranging from the exploitation of the visible light as carrier information, the exploitation of a different portion of the spectrum in the THz frequency, the leverage of artificial molecules for transmitting information (i.e. artificial molecular communication). Another interesting approach is consisting on a different perspective of the interaction between the signals and the environment. Right now, the environment has always been considered as something that cannot be "changed", a kind of obstacle for the wireless transmission. A new paradigm is arising, based on the metamaterial surface, where the interaction between the signals and the environment can be adapted in order to improve the performance of the communication.

6.3.1. TeraHez communications

In [32] and [23] we have investigated a metasurface and how the design of this metasurface has to be realized in order to adapt the behavior of the optical and THz signals based on the specific application considered.

6.3.2. Molecular communications

Concerning the artificial molecular communication paradigm, a fundamental aspect to be considered is that the most of times the target application of this type of communication is a biological system. A fundamental question arising is then: how maximize the effectiveness of the communication by keeping a low impact in terms of interference for the biological system? We have tried to answer to these fundamental questions by considering a signal processing approach in [11], [19].

6.4. Long range communications

Participants: Nathalie Mitton, Brandon Foubert, Ibrahim Amadou.

In the context of smart farming, communications still pose a key challenge. Ubiquitous access to the Internet is not available worldwide, and battery capacity is still a limitation. Inria and the Sencrop company are collaborating to develop an innovative solution for wireless weather stations, based on multi-technology communications, to enable smart weather stations deployment everywhere around the globe. We discuss this model in [12] and assess the quality of a LoRA signal in different conditions [16].

6.5. Vehicular networks

Participants: Nathalie Mitton, Valeria Loscri.

6.5.1. Positioning

Typical Global Navigation Satellite System (GNSS) receivers offer precision in the order of meters. This error margin is excessive for vehicular safety applications, such as forward collision warning, autonomous intersection management, or hard braking sensing. In [14] we develop CooPS, a GNSS positioning system that uses Vehicle to Vehicle (V2V) and Vehicle to Infrastructure (V2I) communications to cooperatively determine absolute and relative position of the ego-vehicle with enough precision. To that end, we use differential GNSS through position vector differencing to acquire track and across-track axes projections, employing elliptical and spherical geometries. We evaluate CooPS performance by carrying out real experiments using off-the-shelf IEEE 802.11p equipment at the campus of the Federal University of Rio de Janeiro. We obtain an accuracy level under 1.0 and 1.5 m for track (where-in-lane) and across-track (which-lane) axes, respectively. These accuracy levels were achieved using a 2.5 m accuracy circular error probable (CEP) of 50% and a 5 Hz navigation update rate GNSS receiver.

6.5.2. Vehicular social networks

In recent years, the concept of social networking combined with the Internet of Vehicles has brought to the definition of the Social IoV (SIoV) paradigm, i.e., a social network where every vehicle is capable of establishing social relationships in an autonomous way with other vehicles or road infrastructure equipment. In SIoV, social networking is applied to vehicular networks according to how social ties are built upon, i.e., either among vehicles or humans. An analysis of the SIoT-based social relations in a vehicular network scenario for establishing a Social Internet of Vehicles and providing insights on this growing research area [33]. By considering the specific features of the Online Social Networks (ONSs) and Vehicular Social Networks (VSNs), we realize that there are limitations and advantages on both these systems. In [15] we have proposed SOVER, a hybrid OSN-VSN framework, allowing the communication between both the communities, the OSNs and VSNs. In [24] we investigate the twofold nature of SIoV, both based on human factors and relationships and as an instance of the Social Internet of Things (SIoT. Based on this twofold nature, it is possible to distinguish different applications and use-cases.)

6.6. On the use of controlled mobility

Participant: Nathalie Mitton.

Relying on controlled mobility as enabled by drones or robots could be a great asset for task management, data collection or quality of network deployment.

6.6.1. Robots

Robots and controlled mobility can help in the dynamic coverage of an area. In [22], we address the problem of defining a wireless sensor network by deploying sensors with the aim of guaranteeing the coverage of the area and the connectivity among the sensors. The wireless sensor networks are widely studied since they provide several services, e.g., environmental monitoring and target tracking. We consider several typologies of sensors characterized by different sensing and connectivity ranges. A cost is associated with each typology of sensors. In particular, the higher the sensing and connectivity ranges, the higher the cost. We formulate the problem of deploying sensors at minimum cost such that each sensor is connected to a base station with either a one-or a multi-hop and the area is full covered. We present preliminary computational results by solving the proposed mathematical model, on several instances. We provide a simulation-based analysis of the performances of such a deployment from the routing perspective.

Robots could be helpful when called upon an alert sent by sensors. But to intervene quickly, they need to locate or follow back the alert source as fast as possible. Two new algorithms (GFGF1 and GFGF2) for event finding in wireless sensor and robot networks based on the Greedy-Face-Greedy (GFG) routing are proposed in [27]. The purpose of finding the event (reported by sensors) is to allocate the task to the closest robot to act upon the event. Using two scenarios (event in or out of the network) and two topologies (random and random with hole) it is shown that GFGF1 always find the closest robot to the event but with more than twice higher communication cost compared to GFG, especially for the outside of the network scenario. GFGF2 features more than 4 times communication cost reduction compared to GFG but with percentage of finding the closest robot up to 90%.

6.6.2. Drones

Disaster scenarios are particularly devastating in urban environments, which are generally very densely populated. Disasters not only endanger the life of people, but also affect the existing communication infrastructure. In fact, such an infrastructure could be completely destroyed or damaged; even when it continues working, it suffers from high access demand to its resources within a short period of time, thereby compromising the efficiency of rescue operations. [31], [25] leverage the ubiquitous presence of wireless devices (e.g., smartphones) in urban scenarios to assist search and rescue activities following a disaster. This work considers multi-interface wireless devices and drones to collect emergency messages in areas affected by natural disasters. Specifically, it proposes a collaborative data collection protocol that organizes wireless devices in multiple tiers by targeting a fair energy consumption in the whole network, thereby extending the network lifetime. Moreover, it introduces a scheme to control the path of drones so as to collect data in a short time. Simulation results in realistic settings show that the proposed solution balances the energy consumption in the network by means of efficient drone routes, thereby effectively assisting search and rescue operations.

6.7. Self-organization, routing and orchestration

Participants: Nathalie Mitton, Valeria Loscri, Brandon Foubert.

By offering low-latency and context-aware services, fog computing will have a peculiar role in the deployment of Internet of Things (IoT) applications for smart environments. Unlike the conventional remote cloud, for which consolidated architectures and deployment options exist, many design and implementation aspects remain open when considering the latest fog computing paradigm. In [9], we focus on the problems of dynamically discovering the processing and storage resources distributed among fog nodes and, accordingly, orchestrating them for the provisioning of IoT services for smart environments. In particular, we show how these functionalities can be effectively supported by the revolutionary Named Data Networking (NDN) paradigm. Originally conceived to support named content delivery, NDN can be extended to request and provide named computation services, with NDN nodes acting as both content routers and in-network service executors. To substantiate our analysis, we present an NDN fog computing framework with focus on a smart campus scenario, where the execution of IoT services is dynamically orchestrated and performed by NDN nodes in a distributed fashion. A simulation campaign in ndnSIM, the reference network simulator of the NDN research community, is also presented to assess the performance of our proposal against state-of-the-art solutions. Results confirm the superiority of the proposal in terms of service provisioning time, paid at the expenses of a slightly higher amount of traffic exchanged among fog nodes.

[26] proposes FLY-COPE, a complete self-organization architecture that relies on cooperative communications and drone-assisted data collection, allowing a fast location of victims and rescuing operation organization in disaster relief operation. FLY-COPE mainly combines two components: i) a ground component that spontaneously emerges from any communicating devices (piece of infrastructure, mobile phone, etc) that cooperate to alert rescuers and remain all alive as long as possible and ii) an aerial component comprising UAV to communicate efficiently with ground devices. We show by simulation and/or by experimentation that each component of FLY-COPE allows substantial energy saving for efficient and fast disaster response.

The IPv6 Routing Protocol for Low-Power and Lossy Networks (RPL) builds a Direction Oriented Directed Acyclic Graph (DODAG) rooted at one node. This node may act as a border router to provide Internet connectivity to the members of the DODAG but such a situation creates a single point of failure. Upon border router failure, all nodes connected to the DODAG are affected as all ongoing communications are instantly broken and no new communications can be initiated. Moreover, nodes close to the border router should forward traffic from farther nodes in addition to their own, which may cause congestion and energy depletion inequality. In [20], we specify a full solution to enable border router redundancy in RPL networks. To achieve this, we propose a mechanism leveraging cooperation between colocated RPL networks. It enables failover to maintain Internet connectivity and load balancing to improve the overall energy consumption and bandwidth. Our contribution has been implemented in Contiki OS and was evaluated through experiments performed on the FIT IoT-LAB testbed.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- Sencrop
  Participants: Brandon Foubert, Nathalie Mitton [contact person].

  This collaboration aims to develop a complete multi-technology bilateral wireless communication stack for agriculture sensor networks.
- Enedis and NooliTic
  Participants: Ibrahim Amadou, Nathalie Mitton [contact person].

  This collaboration aims to investigate a novel localization approach based on wireless propagations. It is a tri-partite contract between our Inria team, the SME NooliTic and Enedis.
- Expleo
  Participants: Ibrahim Amadou, Nathalie Mitton [contact person].

  This collaboration aims to transfer self-deployment protocols to Expleo.

8. Partnerships and Cooperations

8.1. Regional Initiatives

8.1.1. StoreConnect

Participants: Nathalie Mitton [contact person], Valeria Loscri, Antonio Costanzo, Ibrahim Amadou.

Title: StoreConnect

Type: FUI

Duration: September 2016 - October 2018

Coordinator: NEOSENSYS

Others partners: Inria FUN, SPIRALS and STARS, TeVolys, Ubudu, Smile, STIME, Leroy Merlin

The aim of StoreConnect is to provide French large retailers with efficient and powerful tools in the in-store customer interaction.

8.1.2. LumiCAR

Participants: Valeria Loscri [contact person], Antonio Costanzo, Meysam Mayahi.

Title: LumiCAR

Type: ISITE

Duration: October 2019 - October 2021

Vehicle-to-Vehicle and Vehicle-RSU (Roadside Units) communication (V2X) has become a very active topic of research in recent years as it appears to be a means of improving road safety and make effective and timely intervention of road safety actors. To date, most research activities are based on the use of conventional radio frequency (RF) technology. However, faced with multiple constraints, these vehicular communications are not always effective. In the LumiCar project we will base the V2X communication mainly on the Visible Light Communication (VLC) technology and we will focus on the coexistence of the VLC with other technologies. VLC has already started to work in other indoor applications such as connected stores for geolocation of customers. The properties offered by light (speed, directional, controlled containment ...) suggest that VLC technology is more suitable for vehicular communications and can effectively meet the needs of a reliable, robust and with increasing flow to consider new applications such as virtual reality in future cars. In addition, VLC technology can be recognized as a "green" technology because it is based on the exploitation of LEDs and lamps already used for lighting and visibility. It is therefore a question of optimizing the use (by the transmission of information) of an energy already consumed.

8.2. National Initiatives

8.2.1. Exploratory Action

8.2.1.1. Ethicam

Participants: Valeria Loscri [contact person], Carola Rizza.

Duration: October 2019 - October 2022

The evolution of the Internet of Things (IoT) towards the Internet of Everything (IoE) paradigm represents an important and emerging research direction, capable to connect and interconnect massive number of heterogeneous nodes, both inanimate and living entities, encompassing molecules, nanosensors, vehicles and people. This new paradigm demands new engineering communication solutions to overcome miniaturization and spectrum scarcity. Novel pervasive communication paradigms will be conceived by the means of a cutting edge multidisciplinary research approach integrating (quasi) particles (e.g. phonons) and specific features of the (meta)material (e.g. chirality) in the design of the communication mechanisms. In particular, by the means of the meta-materials, it would be possible to control the propagation environment. More specifically, through this paradigm it will be possible to manipulate not only the desired signals, but also the interfering signals.

8.2.2. ADT

8.2.2.1. Catimex

Participants: Matthieu Berthome, Nathalie Mitton [contact person], Julien Vandaele.

Duration: September 2017 -June 2019

Coordinator: Inria FUN

The purpose of this project is to foster research transfer in IoT from ADT members to their industrial partners by widening experimental features and PoC realization. It is lead in closed partnership with Inria Chile and Université of Strasbourg.

8.2.3. Equipements d'Excellence

8.2.3.1. FIT

Participants: Nathalie Mitton [contact person], Julien Vandaele, Matthieu Berthome.

Title: Future Internet of Things

Type: EquipEx

Duration: March 2010 - December 2019

Coordinator: UPMC

See also: http://fit-equipex.fr/

Abstract: FIT (Future Internet of Things) aims to develop an experimental facility, a federated and competitive infrastructure with international visibility and a broad panel of customers. It will provide this facility with a set of complementary components that enable experimentation on innovative services for academic and industrial users. The project will give French Internet stakeholders a means to experiment on mobile wireless communications at the network and application layers thereby accelerating the design of advanced networking technologies for the Future Internet. FIT is one of 52 winning projects from the first wave of the French Ministry of Higher Education and Research's "Equipements d'Excellence" (Equipex) research grant program. Coordinated by Professor Serge Fdida of UPMC Sorbonne Universités and running over a nine-year period, the project will benefit from a 5.8 million euro grant from the French government.

8.3. European Initiatives

8.3.1. H2020 Projects

8.3.1.1. VESSEDIA

Participants: Rehan Malak, Nathalie Mitton, Allan Blanchard [contact person].

Title: Verification Engineering of Safety and Security Critical Dynamic Industrial Applications

Program: H2020

Duration: January 2017 - Dec. 2019

Coordinator: Technikon Forschungs und Planungsgesellschaft MBH (TEC)

The VESSEDIA project will bring safety and security to many new software applications and devices. In the fast evolving world we live in, the Internet has brought many benefits to individuals, organizations and industries. With the capabilities offered now (such as IPv6) to connect billions of devices and therefore humans together, the Internet brings new threats to the software developers and VESSEDIA will allow connected applications to be safe and secure. VESSEDIA proposes to enhance and scale up modern software analysis tools, namely the mostly open-source Frama-C Analysis platform, to allow developers to benefit rapidly from them when developing connected applications. At the forefront of connected applications is the IoT, whose growth is exponential and whose security risks are real (for instance in hacked smart phones). VESSEDIA will take this domain as a target for demonstrating the benefits of using our tools on connected applications. VESSEDIA will tackle this challenge by 1) developing a methodology that allows to adopt and use source code analysis tools efficiently and produce similar benefits than already achieved for highly-critical applications (i.e. an exhaustive analysis and extraction of faults), 2) enhancing the Frama-C toolbox to enable efficient and fast implementation, 3) demonstrating the new toolbox capabilities on typical IoT (Internet of Things) applications including an IoT Operating System (Contiki), 4) developing a standardization plan for generalizing the use of the toolbox, 5) contributing to the Common Criteria certification process, and 6) defining a label "Verified in Europe" for validating software products with European technologies such as Frama-C. This project yields to set of publications in 2019: [17], [18], [35].

8.3.1.2. CyberSANE

Participants: Valeria Loscri, Nathalie Mitton [contact person], Edward Staddon.

Title: Cyber Security Incident Handling, Warning and Response System for the European Critical Infrastructures

Program: H2020

Duration: September 2019 - September 2022

CyberSANE aims to enhance the security and resilience of Critical Information Infrastructures (CIIs) by providing a dynamic collaborative, warning and response system supporting and guiding security officers and operators (e.g. Incident Response professionals) to recognize, identify, dynamically analyze, forecast, treat and respond to advanced persistent threats (APTs) and handle their daily cyber incidents utilizing and combining both structured data (e.g. logs and network traffic) and unstructured data (e.g. data coming from social networks and dark web).

In achieving that aim, CyberSANE will introduce a holistic and privacy-aware approach in handling security incidents, addressing the complexity of these nets consisting of cyber assets hosted in cross-border, heterogeneous Critical Information Infrastructures (CIs). Moreover, CyberSANE is fully in-line with relevant regulations (such as the GDPR and NIS directive), which requires organizations to increase their preparedness, improve their cooperation with each other, and adopt appropriate steps to manage security risks, report and handle security incidents.

8.4. International Initiatives

8.4.1. Inria International Labs

8.4.1.1. Agrinet

Participants: Christian Salim, Brandon Foubert, Nathalie Mitton [contact person].

Title: Agrinet

International Partner (Institution - Laboratory - Researcher): Stellenbosch University, South Africa, Riaan Wolhuter

Type: LIRIMA Associate team

Duration: 2017-2020

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

The current drought and limited water resources in many parts of Southern Africa and beyond, already have a significant impact on agriculture and hence, food production. Sustainable food security depends upon proper plant and crop management respectful of soils and natural re- sources, such as water. This includes very important South African farming areas, such as the Western Cape and Northern Cape. In France, agriculture is also hugely important. Not just nationally, but also in Europe. The system proposed can be applied to a variety of crops. The economic- and social consequences are profound and any contribution towards more efficient farming within increasingly onerous natural constraints, should be a priority. To address these constraints, we propose to develop a flexible, rapidly deployable, biological/agricultural data acquisition platform and associated machine learning algorithms to create advanced agricultural monitoring and management techniques, to improve crop management and use of natural resources. The project also addresses an industry with very high socioeconomic impact.

Publications issued from that project in 2019 are: [12], [16].

8.4.2. Inria International Partners

8.4.2.1. Declared Inria International Partners

Université Mediterranea di Reggio Calabria (UNIC) (Italy): The objective of this collaboration is the design of an innovative architecture that enables autonomic and decentralized fruition of the services offered by the network of smart objects in many heterogeneous and dynamic environments, such that is independent of the network topology, in a reliable and flexible way. The result is an 'ecosystem' of self-organized and self-sustained objects, capable of making data and services available to the users wherever and whenever required, thus supporting the fruition of an 'augmented' reality thanks to a new environmental and social awareness.

8.4.2.2. Informal International Partners

Anna-Maria Vegni from Roma Tre University, Italy: The purpose of this collaboration is to study alternative communication paradigms and investigate their limitations and different effects on performances. In this framework, joint publications have been obtained, among them in 2019 [23], [33], [15], [24], [32].

8.4.3. Participation in Other International Programs

8.4.3.1. International Initiatives

CroMo

Title: Crowd data in the mobile cloud

International Partner (Institution - Laboratory - Researcher):

Universidade Federal do Rio de Janeiro (Brazil) - GTA Laboratory - Luis Henrique Costa

Duration: 2015 - 2019

Start year: 2015

CroMo's main goal is to investigate alternatives to efficiently offload multiple data collected from mobile users to the cloud. To achieve this goal, CroMo will focus on three complementary objectives:

- Objective 1 (Data acquisition): In a wireless environment, data can be sourced at a multitude of wireless devices. Hence, the first objective of this project is to identify the most relevant information from all the data available by using local criteria. The notion of local can be concerned with a single wireless device or a set of nearby wireless devices. The goal is to only send relevant data to the cloud or to assign a higher priority to it.

- Objective 2 (Data transmission): The large-scale sensing system forces massive transmission to the cloud. Hence, transmitting the data in a reliable and timely fashion is the purpose of this second objective.
- Objective 3 (Data computation): Mobile clouds must be available for wireless users to receive and process data. Hence, the cloud infrastructure must be efficient enough to process data from users in a efficient fashion. The third objective of this project is to evaluate cloud availability and to propose performance improvements for data computation. Such improvements are concerned with cloud infrastructure adaptation according to users' demands.

In this context, our project is original and ambitious. Indeed, compared to other studies in wireless networking, our project is focused on a global approach from raw data acquisition to information creation at the mobile cloud infrastructure.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

Several researchers have visited our group in 2019, mainly from our partner universities but not only:

- Gewu Bu, LIP6, France, January 2019
- Noura Mares, University of Sfax, Tunisia, from June 2019 until July 2019
- Marco Di Renzo, Centrale Supelec, France, August 2019
- Riaan Wolhuter, Stellenbosch University, South Africa, September 2019

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. General Chair, Scientific Chair

- Valeria Loscri is General Chair of ACM NanoCom 2020
- Nathalie Mitton is a member of the Steering committee of CIoT
- Nathalie Mitton is demo/poster chair of infocom 2019
- Valeria Loscri is Short Papers, Posters & Demos chair and Local Organizer of WiMoB 2019

9.1.1.2. Chair of Conference Program Committees

- Antoine Gallais is TPC co-chair of CoRes 2019.
- Valeria Loscri is TPC co-chair of IEEE ACM Nanocom 2019, CoRES 2020, TPC chair for Persist-IoT workshop at mobihoc 2019, TPC chair for Persist-IoT at Infocom 2020
- Nathalie Mitton is TPC co-chair of ICC 2019.

9.1.1.3. Member of the Conference Program Committees

- Valeria Loscri is/was member in the Technical Program Committee (TPC) in IoTDI 2020, IEEE GLOBECOM 2020 & 2019, VTC 2020 & 2019, SECON 2020 & 2019, GIoTS 2019, CCNC 2019, ICC 2020 & 2019, WiMob 2020 & 2019 WF-IoT
- Nathalie Mitton is/was in the Technical Program Committee (TPC) of REFRESH 2020, Infocom 2020, Percom 2020 & 2019, WCNC 2020 & 2019, DCOSS 2020 & 2019, CORES 2020 VTC 2020 & 2019, IEEE GLOBECOM 2020 & 2019, GIIS 2020, ICC 2020, AdHoc-Now 2019.
- Antoine Gallais was/is in the Technical Program Committee (TPC) of IEEE GLOBECOM 2019, IEEE ICNC 2019

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- Nathalie Mitton is editorial board members of AHSWN since 2011
- Nathalie Mitton is editorial board member of Adhoc Networks since 2012
- Nathalie Mitton is editorial board member of IET-WSS since 2013
- Nathalie Mitton is editorial board member of ComSoc MMTC e-letter since 2014
- Nathalie Mitton is editorial board member of Wiley Transactions Emerging Telecommunications Technologies since 2016
- Nathalie Mitton is editorial board member of Wireless Communications and Mobile Computing since 2016
- Nathalie Mitton is editorial board member of MDPI Future Internet since 2018
- Valeria Loscri is editorial board member of IEEE Transactions on Nanobioscience journal since 2017
- Valeria Loscri is editorial board member of Elsevier Computer Networks journal since 2016
- Valeria Loscri is editorial board member of Robotics Software Design and Engineering of the International Journal of Advanced Robotic Systems since 2016
- Valeria Loscri is editorial board member of Elsevier Journal of Networks and Computer Applications (JNCA) journal since 2016
- Valeria Loscri is editorial board member of Wiley Transactions Emerging Telecommunications Technologies since 2019

9.1.2.2. Guest editorial activities

- Nathalie Mitton is (co-)editor of Sensor Special Issues of MDPI Sensors on Optimization and Communication in UAV Networks and Lead Guest Editor of the Green Data Collection and Processing in Smart Cities SI at Annals of Telecoms.
- Valeria Loscri is co-editor of Future Internet Special issues of MDPI The Internet of Things in smart environments

9.1.3. Invited Talks

- Nathalie Mitton gave an invited talks at the following events: Journées scientifiques Inria in June, Recontres Inria EuraSanté in June, Journées SPOC at UTC in November and Smart Logistic at IoT Week in December and IoT Days at Paris Sorbonne in December.
- Julien Vandaele and Matthieu Berthome gave an invited talk at TechTalk S Station F in April.

9.1.4. Leadership within the Scientific Community

- Nathalie Mitton is a member of the Steering Committee of the GDR RSD Rescom
- Valeria Loscri is a Member of Social Network Technical Committee
- Valeria Loscri is a Member of Emerging Technologies Initiatives for Molecular, Biological and Multi-Scale Communications (ETI-MBMC)
- Valeria Loscri is a member of the Quantum Communications & Information Technology Emerging Technical Subcommittee (QCIT)
- Valeria Loscri is a member of the 'Research Group on IoT Communications and Networking Infrastructure' at ComSoc Communities
- Valeria Loscri is Newsletter Chair for the Technical Committee for Social Networks (TCSN) IEEE Comsoc since April 2019.

9.1.5. Scientific Expertise

- Nathalie Mitton

  - She is an elected member of the evaluation community of Inria.
  - She has acted as a reviewer for ANRT and ANR project submissions.
  - She is also member of the scientific committees of the competitiveness cluster of MATIKEN and for CITC (International Contactless Technologies Center).
  - She is member of the CA of CITC.
  - She has been a member of the PRIX Roberval committee.
  - She has been a member of the Bourses L'Oreal FRANCE.
  - She is a member of the HCERES visiting committee for the LISIS laboratory.
  - She has been appointed as reviewer for the Polish National Science Centre.
  - She has been appointed as reviewer for the Heinz Maier-Leibnitz-Prize 2020.
- Valeria Loscri
  - She has been reviewer for TOP grant for senior researchers for Netherlands Organisation for Scientific Research (NWO) program.
  - She has been reviewer for proposal for grant competition at the Center for Excellence in Applied Computational Science and Engineering (CEACSE), UTC (USA).
  - She has been external evaluator of ERC Consolidator Grants.
  - She was Scientific European Responsible for Inria Lille - Nord Europe till September 2019.
  - She is Scientific International Responsible for Inria Lille from September 2019.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

E-learning

Mooc, Nathalie Mitton, "Villes intelligentes : défis technologiques et sociétaux", 5-week mooc by the IPL CityLab@Inria team, FUN, Inria, in November 2019

SPOC, Nathalie Mitton, EIT Digital "Technological challenges of participatory Smart cities", 5-week by the IPL CityLab@Inria team in November 2019

Remote curse, Nathalie Mitton, Internet of things, 5-week + face to face week in May 2019

Master: Valeria Loscri, Objets Communicants, 24h (Mineure Habitat Intelligent), Ecole des Mines de Douai, France

Master: Nathalie Mitton, Wireless networks, 16h eqTD (Master TiiR), Université Lille 1, France

Master: Nathalie Mitton, Wireless sensor networks, 16h eqTD (Master MINT), Université Lille 1 and Telecom Lille 1, France

Master: Nathalie Mitton, Smart objects, 10h CM + 12h TP, Ecole centrale de Lille, France

Master: Nathalie Mitton, Industrial Internet of Things, 10h CM Ecole centrale de Lille, France

Master: Ibrahim Amadou, Industrial to Internet of Things, 12h TP Ecole centrale de Lille, France

Master: Nathalie Mitton, Introduction to Internet of Things, 4h CM Ecole centrale de Lille, France

Master: Ibrahim Amadou, Introduction to Internet of Things, 8h TP Ecole centrale de Lille, France

Master: Nathalie Mitton, Wireless sensor networks, 16h eqTD (Master ROC), IMT, France

Licence: Brandon Foubert, Conception Orientée Objet, 42.06h eqTD (Licence 3), FST, France

9.2.2. Supervision

PhD in progress: Brandon Foubert, Communication sans fil Polymorphe pour l'Agriculture Connectée, Université Lille 1, 2018-2021, Nathalie Mitton

PhD in progress: Edward Staddon, Threat detection, identification and quarantine in wireless IoT based Critical Infrastructures, Université Lille 1, 2019-2022, Nathalie Mitton & Valeria Loscri

PhD in progress: Carola Rizza, Nouveaux paradigmes de communication basés sur les technologies émergentes, Université Lille 1, 2019-2022, Valeria Loscri

PhD in progress: Emilie Bout, Denial-of-sleep over IoT networks, Université Lille 1, 2019-2022, Valeria Loscri & Antoine Gallais

PhD in progress: Meysam Mayahi, Communication Protocols based on alternative paradigm for wireless mobile devices, Université Lille 1, 2019-2022, Valeria Loscri

PhD in progress: Mohammad Hussein Ghosn, Université Lille 1, 2019-2022, Valeria Loscri & Alia Ghadar

9.2.3. Juries

- PhD and HDR committees:
  - Valeria Loscri is/was member of the following PhD thesis committees:
    - Xiaojun Xi, CentraleSupelée, Dec. 2019.
    - Jian Song, CentraleSupelée, Dec. 2019.
    - Jorge Martinez, Universidade Politecnica Valencia, Spain, June 2019 (reviewer).
  - Nathalie Mitton is/was member of the following PhD thesis committees:
    - Domga Rodrigue Komguem INSA Lyon/University of Yaoundé (reviewer)
    - Maxime Mroue, Université de Nantes, December 2019 (chair)
    - Alwan Safwan, Université Paris Est, December 2019
    - Gewu Bu, Université Paris Sorbonne, December 2019 (reviewer)
    - Marija Stojanova, Université de Lyon, December 2019
    - Henry-Joseph Audeoud, Université de Grenoble, December 2019
    - Francisco Jose Fabra Collado, Universitat Politecnica de Valencia, December 2019 (reviewer)
    - Lilian Besson, Centrale Supelec Rennes, novembre 2019 (reviewer, chair)
    - Dorin Rautu, ENSEEITH, October 2019 (chair)
    - Mariem Harmassi, Université de La Rochelle, September 2019 (reviewer)
    - Nicolas De Araujo Moreira, Université de Lille, July 2019
    - Wei Hu, Centrale Lille, July 2019 (chair)
    - Philippe Pittoli, Université de Strasbourg (reviewer) may 2019
    - Fadhlallah Baklouti, Université Bretagne Sud, February 2019
    - Mohammed Islam Naas, Université de Bretagne Occidentale (reviewer), February 2019
  - Nathalie Mitton was a member of the HDR defense committees
    - Thomas Watteyne, Sorbonne Université, May 2019
- Researcher selection committees:
  - Nathalie Mitton was a member of the Professor competition selection committee at INSA Lyon
  - Nathalie Mitton was a member of the Assistant Professor (MdC) committee for Université de Lyon.

  - Nathalie Mitton was a member of a CPPI selection for Inria Saclay.
  - Nathalie Mitton was a member of the Inria chaire committee for Centrale Supelec Rennes.
  - Nathalie Mitton was a member of the Inria SRP/ARP selection.
  - Nathalie Mitton was a member of the national Inria junior researcher competition committee.
  - Valeria Loscri was a member of the Inria junior researcher competition committee for Lille center.
  - Valeria Loscri was a member of the MdC committee for Centrale Supelec.
- PhD follow-up committees:
  - Nathalie Mitton is/was reviewer of the following PhD follow-up committees:
    * Rodrigue Domgua Rodriguez, INSA Lyon
    * Alexis Bitaillou, Université de Nantes
    * Vaseileios Kotsiou, Université de Strasbourg
  - Valeria Loscri is/was reviewer of the following PhD follow-up committees:
    * Lin Chen, Centrale Supélec.

9.3. Popularization

9.3.1. Internal or external Inria responsibilities

- Nathalie Mitton is the referent researcher for the creation of the MATH Laboratory in Nord

9.3.2. Articles and contents

- Nathalie Mitton published a paper in Techniques de l'Ingénieur.
- Valeria Loscri released an interview "Innovation Valeria Loscri, la lumière des LED contre le wifi électromagnétique" for the newspaper La Voix du Nord

10. Bibliography

Major publications by the team in recent years

[1] A. Costanzo, V. Loscri. *Error Compensation in Indoor Positioning Systems based on Software Defined Visible Light Communication*, in "Physical Communication", 2019, https://hal.archives-ouvertes.fr/hal-02087506

[2] 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° 1, p. 51-60 [*DOI :* 10.1109/TNB.2018.2884999], https://hal.archives-ouvertes.fr/hal-01928205

[3] F. Guerriero, V. Loscri, P. Pace, R. Surace. *Neural Networks and SDR Modulation schemes for wireless mobile nodes: a synergic approach*, in "Ad Hoc Networks", 2017 [*DOI :* 10.1016/J.ADHOC.2016.09.016], https://hal.inria.fr/hal-01386749

[4] F. Mezghani, N. Mitton. *Alternative opportunistic alert diffusion to support infrastructure failure during disasters*, in "Sensors", October 2017 [*DOI :* 10.3390/s17102370], https://hal.inria.fr/hal-01614744

[5] R. Morabito, R. Petrolo, V. Loscri, N. Mitton. *LEGIoT: a Lightweight Edge Gateway for the Internet of Things*, in "Future Generation Computer Systems", 2017, https://hal.inria.fr/hal-01614714

[6] J. B. PINTO NETO, L. C. GOMES, F. C. ORTIZ, T. T. ALMEIDA, M. E. M. CAMPISTA, L. H. MACIEL KOSMALSKI COSTA, N. MITTON.*An Accurate Cooperative Positioning System for Vehicular Safety Applications*, in "Computers and Electrical Engineering (COMPELECENG)", 2020, https://hal.inria.fr/hal-02364355

[7] V. TOLDOV, L. CLAVIER, V. LOSCRI, N. MITTON.*A Thompson Sampling Approach to Channel Exploration-Exploitation Problem in Multihop Cognitive Radio Networks*, in "27th annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)", Valencia, Spain, September 2016, https://hal.inria.fr/hal-01355002

[8] V. TOLDOV, L. CLAVIER, N. MITTON.*Multi-channel Distributed MAC protocol for WSN-based wildlife monitoring*, in "WiMob 2018 - 14th International Conference on Wireless and Mobile Computing, Networking and Communications", Limassol, Cyprus, October 2018, https://hal.inria.fr/hal-01866809

Publications of the year

Articles in International Peer-Reviewed Journal

[9] M. AMADEO, G. RUGGERI, C. CAMPOLO, A. MOLINARO, V. LOSCRI, C. T. CALAFATE.*Fog Computing in IoT Smart Environments via Named Data Networking: A Study on Service Orchestration Mechanisms*, in "Future internet", November 2019, vol. 11, n^{o} 11, 222 [*DOI :* 10.3390/FI11110222], https://hal.inria.fr/hal-02354073

[10] A. COSTANZO, V. LOSCRI.*Error Compensation in Indoor Positioning Systems based on Software Defined Visible Light Communication*, in "Physical Communication", 2019, https://hal.archives-ouvertes.fr/hal-02087506

[11] 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

[12] B. FOUBERT, N. MITTON.*Autonomous Collaborative Wireless Weather Stations: A Helping Hand for Farmers*, in "ERCIM News", October 2019, n^{o} 119, p. 37-38, forthcoming, https://hal.inria.fr/hal-02308015

[13] F. MEZGHANI, N. MITTON.*Opportunistic multi-technology cooperative scheme and UAV relaying for network disaster recovery*, in "Information", January 2020, vol. 11, n^{o} 1, https://hal.inria.fr/hal-02431659

[14] J. B. PINTO NETO, L. C. GOMES, F. C. ORTIZ, T. T. ALMEIDA, M. E. M. CAMPISTA, L. H. MACIEL KOSMALSKI COSTA, N. MITTON.*An Accurate Cooperative Positioning System for Vehicular Safety Applications*, in "Computers and Electrical Engineering", 2019, https://hal.inria.fr/hal-02364355

[15] A. M. VEGNI, V. LOSCRI, A. BENSLIMANE.*SOLVER: A Framework for the Integration of Online Social Networks with Vehicular Social Networks*, in "IEEE Network", July 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-02178424

International Conferences with Proceedings

[16] I. AMADOU, B. FOUBERT, N. MITTON.*LoRa in a haystack: a study of the LoRa signal behavior*, in "WiMob 2019 - 15th International Conference on Wireless and Mobile Computing, Networking and Communications", Barcelona, Spain, October 2019, https://hal.inria.fr/hal-02284062

[17] *Best Paper*
A. BLANCHARD, N. KOSMATOV, F. LOULERGUE.*Logic against Ghosts: Comparison of Two Proof Approaches for a List Module*, in "SAC 2019 - The 34th ACM/SIGAPP Symposium On Applied Computing", Limassol, Cyprus, April 2019 [*DOI :* 10.1145/3297280.3297495], https://hal.inria.fr/hal-02100515.

[18] A. BLANCHARD, F. LOULERGUE, N. KOSMATOV.*Towards Full Proof Automation in Frama-C Using Auto-active Verification*, in "NASA Formal Methods", Houston, TX, United States, Springer, May 2019, p. 88-105 [*DOI :* 10.1007/978-3-030-20652-9_6], https://hal.inria.fr/hal-02317055

[19] 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

[20] B. FOUBERT, J. MONTAVONT.*Sharing is caring: a cooperation scheme for RPL network resilience and efficiency*, in "ISCC 2019 - 24th Symposium on Computers and Communications", Barcelona, Spain, June 2019, https://hal.inria.fr/hal-02095410

[21] A. GALLAIS, T.-H. HEDLI, V. LOSCRI, N. MITTON.*Denial-of-Sleep Attacks against IoT Networks*, in "CoDIT 2019 - 6th International Conference on Control, Decision and Information Technologies", Paris, France, April 2019, https://hal.inria.fr/hal-02060608

[22] F. GUERRIERO, L. DI PUGLIA PUGLIESE, N. MITTON.*Efficient wireless sensor deployment at minimum cost*, in "AdHoc-Now 2019 - 18th International Conference on Ad Hoc Networks and Wireless", Luxembourg, Luxembourg, October 2019, https://hal.inria.fr/hal-02197689

[23] L. LA SPADA, V. LOSCRI, A. M. VEGNI.*MetaSurface Structure Design and Channel Modelling for THz Band Communications*, in "UBTCN'19 - Workshop on Ultra-high Broadband Terahertz Communication for 5G and Beyond networks in conjunction with IEEE INFOCOM 2019", Paris, France, April 2019, https://hal.inria.fr/hal-02050843

[24] V. LOSCRI, P. MANZONI, M. NITTI, G. RUGGERI, A. M. VEGNI.*A social internet of vehicles sharing SIoT relationships*, in "PERSIST-IoT workshop in conjunction with MobiHoc", Catania, Italy, July 2019 [*DOI :* 10.1145/1122445.1122456], https://hal.inria.fr/hal-02136896

[25] F. MEZGHANI, P. KORTOÇI, N. MITTON, M. DI FRANCESCO.*A Multi-tier Communication Schemefor Drone-assisted Disaster Recovery Scenarios*, in "PIMRC 2019 - IEEE International Symposium on Personal, Indoor and Mobile Radio Communications", Istanbul, Turkey, September 2019, https://hal.inria.fr/hal-02152133

[26] F. MEZGHANI, N. MITTON.*The potential of cooperative communications to speed up disaster relief operations*, in "Information and Communication Technologies for Disaster Management (ICT-DM)", Paris, France, December 2019, https://hal.inria.fr/hal-02375707

[27] J. STANULOVIC, N. MITTON, I. MEZEI.*Robot Task Allocation based on Greedy-Face-Greedy Algorithm*, in "Telecommunications Forum (TELFOR)", Belgrade, Serbia, November 2019, https://hal.inria.fr/hal-02344837

Conferences without Proceedings

[28] A. COSTANZO, V. LOSCRI.*A Learning Approach for Robust Carrier Recovery in Heavily Noisy Visible Light Communication*, in "WCNC 2019 - IEEE Wireless Communications and Networking Conference", Marrakesh, Morocco, April 2019, https://hal.archives-ouvertes.fr/hal-02022625

[29] A. COSTANZO, V. LOSCRI.*Système de localisation à Lumière Visible avec compensation du bruit environnemental*, in "CORES 2019", Narbonne, France, June 2019, https://hal.archives-ouvertes.fr/hal-02119952

[30] A. COSTANZO, V. LOSCRI.*Visible Light Indoor Positioning in a Noise-aware Environment*, in "WCNC 2019 - IEEE Wireless Communications and Networking Conference", Marrakesh, Morocco, April 2019, https://hal.archives-ouvertes.fr/hal-02022610

[31] F. MEZGHANI, N. MITTON.*Drone assisted multi-technology rescue operations*, 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, https://hal.archives-ouvertes.fr/hal-02122806

[32] A. TAIBI, A. DURANT, V. LOSCRI, A. M. VEGNI, L. LA SPADA.*Controlling Light by Curvilinear MetaSurfaces*, in "ACM NanoCom 2019 - 6th ACM International Conference on Nanoscale Computing and Communication", Dublin, Ireland, September 2019 [*DOI :* 10.1145/1122445.1122456], https://hal.archives-ouvertes.fr/hal-02177907

[33] A. M. VEGNI, V. LOSCRI, P. MANZONI.*Analysis of Small-World Features in Vehicular Social Networks*, in "CCNC 2019 - IEEE Consumer Communications & Networking Conference", Las Vegas, United States, January 2019, https://hal.archives-ouvertes.fr/hal-01907422

Scientific Books (or Scientific Book chapters)

[34] V. LOSCRI, N. MITTON, R. PETROLO.*Cloud, network and sensing in a Smart City, toward a cloud of meshed cooperative heterogeneous things*, in "Smart Cities in the Post-algorithmic EraIntegrating Technologies, Platforms and Governance", Edward Elgar Publishing, December 2019, https://hal.inria.fr/hal-02431666

Other Publications

[35] A. BLANCHARD, N. KOSMATOV, F. LOULERGUE.*La logique contre les fantômes: comparaison de deux approches pour la preuve d'un module de listes chaînées* *, June 2019, 18e journées Approches Formelles dans l'Assistance au Développement de Logic (AFADL), https://hal.inria.fr/hal-02317143

[36] A. COSTANZO, V. LOSCRI, M. BIAGI.*A Noise Mitigation Approach for VLC Systems*, June 2019, Global Lifi Congress, Poster, https://hal.archives-ouvertes.fr/hal-02082236

References in notes

[37] N. ANCIAUX, L. BOUGANIM, T. DELOT, S. ILARRI, L. KLOUL, N. MITTON, P. PUCHERAL.*Opportunistic Data Services in Least Developed Countries: Benefits, Challenges and Feasibility Issues*, in "ACM SIGMOD Record", March 2014, https://hal.inria.fr/hal-00971805

[38] J. BACHRACH, C. TAYLOR.*Localization in sensor networks*, in "Artificial Intelligence", 2004, p. 227–310

[39] A. BLANCHARD, N. KOSMATOV, F. LOULERGUE.*Ghosts for Lists: A Critical Module of Contiki Verified in Frama-C*, in "Tenth NASA Formal Methods Symposium - NFM 2018", Newport News, United States, April 2018, https://hal.inria.fr/hal-01720401

[40] P. BOSE, P. MORIN, I. STOJMENOVIC, J. URRUTIA.*Routing with guaranteed delivery in ad hoc wireless networks*, in "Proc. of the 3rd Int. Workshop on Discrete Algorithms and Methods for Mobile Computing and Comm. (DIAL-M)", Seattle, WA, USA, August 1999, p. 48-55

[41] H. CHAOUCHI.*Internet of Things. Connecting Objects*, Wiley and Sons, January 2010

[42] A. COSTANZO, V. LOSCRI, S. COSTANZO.*Adaptive Dual Color Visible Light Communication (VLC) System*, in "6th World Conference on Information Systems and Technologies", Napoli, Italy, March 2018, https://hal.inria.fr/hal-01687543

[43] A. DAS, A. SEN, C. QIAO, N. GHANI, N. MITTON.*RAPTOR: a network tool for mitigating the impactof spatially correlated failures in infrastructure networks*, in "Annals of Telecommunications - annales des télécommunications", February 2018, vol. 73, n^{o} 1, p. 153-164, https://hal.inria.fr/hal-01584867

[44] S. DUQUENNOY, B. AL NAHAS, O. LANDSIEDEL, T. WATTEYNE.*Orchestra: Robust Mesh Networks Through Autonomously Scheduled TSCH*, in "Proceedings of the International Conference on Embedded Networked Sensor Systems (ACM SenSys)", 2015

[45] M. EL AMINE SEDDIK, V. TOLDOV, L. CLAVIER, N. MITTON.*From Outage Probability to ALOHA MAC Layer Performance Analysis in Distributed WSNs*, in "WCNC 2018 - IEEE Wireless Communications and Networking Conference", Barcelona, Spain, April 2018, https://hal.inria.fr/hal-01677687

[46] E. ELHAFSI, N. MITTON, B. PAVKOVIC, D. SIMPLOT-RYL.*Energy-aware Georouting with Guaranteed Delivery in Wireless Sensor Networks with Obstacles*, in "International Journal of Wireless Information Networks", 2009, vol. 16, n^{o} 3, p. 142–153

[47] B. FABIAN, O. GÜNTHER.*Distributed ONS and its Impact on Privacy*, in "IEEE International Conference on Communications", 2007, p. 1223–1228

[48] A. GALLAIS, J. CARLE, D. SIMPLOT-RYL, I. STOJMENOVIC.*Localized Sensor Area Coverage with Low Communication Overhead*, in "IEEE Transactions on Mobile Computing (TMC)", May 2008, vol. 5, n^{o} 7, p. 661-672

[49] N. GOUVY, N. MITTON, J. ZHENG.*Greedy Routing Recovery Using Controlled Mobility in Wireless Sensor Networks*, in "ADHOC-NOW - The 12th International Conference on Ad Hoc Networks and Wireless - 2013", Wroclaw, Poland, July 2013, http://hal.inria.fr/hal-00830374

[50] F. INGELREST, N. MITTON, D. SIMPLOT-RYL.*A Turnover based Adaptive HELLO Protocol for Mobile Ad Hoc and Sensor Networks*, in "Proc. of the 15th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS'07)", Bogazici University, Istanbul, Turkey, October 2007

[51] X. LI, N. MITTON, D. SIMPLOT-RYL.*Mobility Prediction Based Neighborhood Discovery in Mobile Ad Hoc Networks*, in "Proc. of the IFIP/TC6 NETWORKING 2011", Valencia, Spain, 2011

[52] A. A. MBACKÉ, N. MITTON, H. RIVANO.*RFID Anticollision in Dense Mobile Environments*, in "WCNC 2017 - IEEE Wireless Communications and Networking Conference", San Francisco, United States, March 2017, https://hal.inria.fr/hal-01418170

[53] N. MITTON, D. SIMPLOT-RYL.*Wireless Sensor and Robot Networks From Topology Control to Communication Aspects*, WorldScientific, February 2014, 284, https://hal.inria.fr/hal-00927526

[54] A. PAL.*Localization Algorithms in Wireless Sensor Networks: Current Approaches and Future Challenges*, in "Network Protocols and Algorithms", 2010, vol. 2, n^{o} 1

[55] R. PETROLO, T. DELOT, N. MITTON, A. MOLINARO, C. CAMPOLO.*O-SPIN: an Opportunistic Data Dissemination Protocol for Folk-enabled Information System in Least Developed Countries*, in "ADHOC-NOW - 13th International Conference on Ad-Hoc Networks and Wireless", Benidorm, Spain, June 2014, https://hal.inria.fr/hal-00960023

[56] R. PETROLO, V. LOSCRI, N. MITTON, E. HERRMANN.*Adaptive Filtering as a Service for Smart City Applications*, in "14th IEEE International Conference on Networking, Sensing and Control (ICNSC)", Cosenza, Italy, May 2017, https://hal.inria.fr/hal-01482715

[57] C. RAZAFIMANDIMBY, V. LOSCRI, A. M. VEGNI.*A neural network and IoT-based scheme for performance assessment in Internet of Robotic Things*, in "I4T - 1st International Workshop on Interoperability, Integration, and Interconnection of Internet of Things Systems", Berlin, Germany, April 2016, https://hal.inria.fr/hal-01261842

[58] L. SCHMIDT, N. MITTON, D. SIMPLOT-RYL, R. DAGHER, R. QUILEZ.*DHT-based distributed ALE engine in RFID Middleware*, in "Proc. 7th IEEE International Conference on RFID-Technologies and Applications (RFID-TA)", Barcelona, Spain, 2011

[59] V. TOLDOV.*Adaptive MAC layer for interference limited WSN*, Université Lille 1 Sciences et technologies, January 2017, https://hal.inria.fr/tel-01493094

Project-Team INOCS

INtegrated Optimization with Complex Structure

IN COLLABORATION WITH: Centre de Recherche en Informatique, Signal et Automatique de Lille

IN PARTNERSHIP WITH:
Ecole Centrale de Lille

Université Libre de Bruxelles

RESEARCH CENTER
Lille - Nord Europe

THEME
Optimization, machine learning and statistical methods

Table of contents

Project-Team INOCS

Creation of the Team: 2015 May 01, updated into Project-Team: 2019 May 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
A6.2.3. - Probabilistic methods
A6.2.6. - Optimization
A9.6. - Decision support

Other Research Topics and Application Domains:

B4. - Energy
B4.3. - Renewable energy production
B4.4. - Energy delivery
B4.5. - Energy consumption
B6. - IT and telecom
B6.3.2. - Network protocols
B7. - Transport and logistics
B7.1. - Traffic management
B7.1.2. - Road traffic
B8.1. - Smart building/home
B8.1.1. - Energy for smart buildings
B8.2. - Connected city
B8.4. - Security and personal assistance

1. Team, Visitors, External Collaborators

Research Scientists

Luce Brotcorne [Team leader, Inria, Researcher, HDR]
Miguel Anjos [Inria, International Chair & Professor at Polytechnique Montréal, Canada]

Faculty Members

Diego Cattaruzza [Centrale Lille, Associate Professor]
Bernard Fortz [Université libre de Bruxelles, Belgium, Professor]
Martine Labbé [Université libre de Bruxelles, Belgium, Professor]
Maxime Ogier [Centrale Lille, Associate Professor]
Frédéric Semet [Centrale Lille, Professor]

Technical Staff

Tifaout Almeftah [Inria, Engineer, from Dec 2019]
Mathieu Besançon [Inria, Engineer]
Martin Bue [Inria, Engineer, from Apr 2019 until Sep 2019]

PhD Students

Yaheng Cui [Inria]
Concepción Domínguez Sánchez [Fund for research training in industry and agriculture (FRIA), Belgium]

Fränk Plein [National Fund for Scientific Research (FNRS), Belgium]
Moises Rodríguez Madrena [University of Sevilla, Spain]
Wenjuan Gu [Centrale Lille]
Matteo Petris [Inria, from Oct 2019]
Luis Alberto Salazar Zendeja [National Council of Science and Technology (CONACYT), Mexico]
Léonard Von Niederhausern [Inria, until Jun 2019]
Yuan Yuan [Centrale Lille]

Post-Doctoral Fellows
Pablo Escalona Rodriguez [Inria, until Sep 2019]
Arnaud Laurent [Inria]
Guillerme Duvillié [Université libre de Bruxelles, Belgium]
Victor Bucarey [Université libre de Bruxelles, Belgium]

Visiting Scientists
Yasemin Arda Da Silveira [University of Liège, Belgium, Mar 2019]
Maria Del Carmen Gale Pola [University of Zaragoza, Spain, until Feb 2019]
Anton Kleywegt [Georgia Institute of Technology, USA, from Apr 2019 until May 2019]
Daniel Pereda Herrera [University of Chile, from Nov 2019]
Sebastián Dávila Gálvez [University of Chile, from June 2019 until Dec 2019]
Natividad Gonzalez Blanco [University of Seville, Spain, from May 2019 until July 2019]
Federica Laureana [University of Salerno, Italy, from Feb 2019 until May 2019]

Administrative Assistant
Nathalie Bonte [Inria]

2. Overall Objectives

2.1. Introduction

INOCS is a cross-border "France-Belgium" project team in the Applied Mathematics Computation and Simulation Inria domain. The main goal of this team is the study of optimization problems involving complex structures. The scientific objectives of INOCS are related to modeling and methodological concerns. The INOCS team will focus on:

1. integrated models for problems with complex structure (CS) taking into account the whole structure of the problem;
2. on the development of solution methods taking explicitly into account *the nature and the structure of the decisions as well as the properties of the problem*.

Even if CS problems are in general NP-hard due to their complex nature, exact solution methods or matheuristics (heuristics based on exact optimization methods) will be developed by INOCS. The scientific contribution of INOCS will result in a toolbox of models and methods to solve challenging real life problems.

2.2. Schedule of tasks

The research program development of INOCS is to move alternatively:

- *from problems towards new approaches in optimization*: Models and solution algorithms will be developed to fit the structure and properties of the problem. From them, new generic approaches will be used to optimize problems with similar properties.
- *from innovative approaches towards problems*: The relevance of the proposed approaches will be assessed by designing new models and/or solution methods for various classes of problems. These models and methods will be based on the extension and integration of specific, well studied, models and methods.

Even if these two axes are developed sequentially in a first phase, their interactions will lead us to explore them jointly in the mid-term.

3. Research Program

3.1. Introduction

An optimization problem consists in finding a best solution from a set of feasible solutions. Such a problem can be typically modeled as a mathematical program in which decision variables must:

1. satisfy a set of constraints that translate the feasibility of the solution and
2. optimize some (or several) objective function(s). Optimization problems are usually classified according to types of decision to be taken into strategic, tactical and operational problems.

We consider that an optimization problem presents a complex structure when it involves decisions of different types/nature (i.e. strategic, tactical or operational), and/or presenting some hierarchical leader-follower structure. The set of constraints may usually be partitioned into global constraints linking variables associated with the different types/nature of decision and constraints involving each type of variables separately. Optimization problems with a complex structure lead to extremely challenging problems since a global optimum with respect to the whole sets of decision variables and of constraints must be determined.

Significant progresses have been made in optimization to solve academic problems. Nowadays large-scale instances of some NP-Hard problems are routinely solved to optimality. *Our vision within INOCS is to make the same advances while addressing CS optimization problems*. To achieve this goal we aim to develop global solution approaches at the opposite of the current trend. INOCS team members have already proposed some successful methods following this research lines to model and solve CS problems (e.g. ANR project RESPET, Brotcorne *et al.* 2011, 2012, Gendron *et al.* 2009, Strack *et al.* 2009). However, these are preliminary attempts and a number of challenges regarding modeling and methodological issues have still to be met.

3.2. Modeling problems with complex structures

A classical optimization problem can be formulated as follows:

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & x \in X. \end{aligned} \tag{1}$$

In this problem, X is the set of feasible solutions. Typically, in mathematical programming, X is defined by a set of constraints. x may be also limited to non-negative integer values.

INOCS team plan to address optimization problem where two types of decision are addressed jointly and are interrelated. More precisely, let us assume that variables x and y are associated with these decisions. A generic model for CS problems is the following:

$$\begin{aligned} \min \quad & g(x,y) \\ s.t. \quad & x \in X, \\ & (x,y) \in XY, \\ & y \in Y(x). \end{aligned} \tag{2}$$

In this model, X is the set of feasible values for x. XY is the set of feasible values for x and y jointly. This set is typically modeled through linking constraints. Last, $Y(x)$ is the set of feasible values for y for a given x. In INOCS, we do not assume that $Y(x)$ has any properties.

The INOCS team plans to model optimization CS problems according to three types of optimization paradigms: large scale complex structures optimization, bilevel optimization and robust/stochastic optimization. These paradigms instantiate specific variants of the generic model.

Large scale complex structures optimization problems can be formulated through the simplest variant of the generic model given above. In this case, it is assumed that $Y(x)$ does not depend on x. In such models, X and Y are associated with constraints on x and on y, XY are the linking constraints. x and y can take continuous or integer values. Note that all the problem data are deterministically known.

Bilevel programs allow the modeling of situations in which a decision-maker, hereafter the leader, optimizes his objective by taking explicitly into account the response of another decision maker or set of decision makers (the follower) to his/her decisions. Bilevel programs are closely related to Stackelberg (leader-follower) games as well as to the principal-agent paradigm in economics. In other words, bilevel programs can be considered as demand-offer equilibrium models where the demand is the result of another mathematical problem. Bilevel problems can be formulated through the generic CS model when $Y(x)$ corresponds to the optimal solutions of a mathematical program defined for a given x, i.e. $Y(x) = \arg\min\{h(x,y)|y \in Y_2, (x,y) \in XY_2\}$ where Y_2 is defined by a set of constraints on y, and XY_2 is associated with the linking constraints.

In robust/stochastic optimization, it is assumed that the data related to a problem are subject to uncertainty. In stochastic optimization, probability distributions governing the data are known, and the objective function involves mathematical expectation(s). In robust optimization, uncertain data take value within specified sets, and the function to optimize is formulated in terms of a min-max objective typically (the solution must be optimal for the worst-case scenario). A standard modeling of uncertainty on data is obtained by defining a set of possible scenarios that can be described explicitly or implicitly. In stochastic optimization, in addition, a probability of occurrence is associated with each scenario and the expected objective value is optimized.

3.3. Solving problems with complex structures

Standard solution methods developed for CS problems solve independent sub-problems associated with each type of variables without explicitly integrating their interactions or integrating them iteratively in a heuristic way. However these subproblems are intrinsically linked and should be addressed jointly. In *mathematicaloptimization* a classical approach is to approximate the convex hull of the integer solutions of the model by its linear relaxation. The main solution methods are (1) polyhedral solution methods which strengthen this linear relaxation by adding valid inequalities, (2) decomposition solution methods (Dantzig Wolfe, Lagrangian Relaxation, Benders decomposition) which aim to obtain a better approximation and solve it by generating extreme points/rays. Main challenges are (1) the analysis of the strength of the cuts and their separations for polyhedral solution methods, (2) the decomposition schemes and (3) the extreme points/rays generations for the decomposition solution methods.

The main difficulty in solving *bilevel problems* is due to their non convexity and non differentiability. Even linear bilevel programs, where all functions involved are affine, are computationally challenging despite their apparent simplicity. Up to now, much research has been devoted to bilevel problems with linear or convex follower problems. In this case, the problem can be reformulated as a single-level program involving complementarity constraints, exemplifying the dual nature, continuous and combinatorial, of bilevel programs.

4. Application Domains

4.1. Energy

In energy, the team mainly focuses on pricing models for demand side management. Demand side management methods are traditionally used to control electricity demand which became quite irregular recently and resulted in inefficiency in supply. We have explored the relationship between energy suppliers and customers who are connected to a smart grid. The smart grid technology allows customers to keep track of hourly prices and shift their demand accordingly, and allows the provider to observe the actual demand response to its pricing strategy. We tackle pricing problems in energy according to the bilevel optimization approaches. Some research works in this domain are supported by bilateral grants with EDF.

4.2. Transportation and Logistics

In transportation and logistics, the team addresses mainly integrated problems, which require taking into account simultaneously different types of decision. Examples are location and routing, inventory management and routing or staff scheduling and warehouse operations management. Such problems occur from the supply chain design level to the logistic facility level.

4.3. Telecommunications

In telecommunications, the team mainly focuses on network design problems and on routing problems. Such problems are optimization problems with complex structure, since the optimization of capacity installation and traffic flow routing have to be addressed simultaneously.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

- Martine Labbé received the EURO Gold Medal in June 2019. This is the highest distinction in Operations Research in Europe.

6. New Software and Platforms

6.1. HappyChic-ApproPick

KEYWORDS: Operational research - Optimization - Java

FUNCTIONAL DESCRIPTION: This software is a prototype developed for the bilateral contract with the company HappyChic. This software is a solver for an integrated warehouse order picking problem with manual picking operations. More precisely, the following problems are solved: (1) the assignment of references to storage positions, based on the iterative solving of minimum cost flow problems, (2) the division of clients orders into several parcels, respecting weight and size constraints, using a dynamic programming algorithm based on the split algorithm, (3) the batching of parcels into trolleys to perform picking tours, using a dynamic programming algorithm based on the split algorithm. The objective function is to minimize the total walking distance. This software is designed to deal with the large-sized industrial instances of HappyChic (considering hundreds of clients, thousands of positions and product references) in a short computation time (few minutes).

- Contact: Maxime Ogier

6.2. KEOLIS-MEDIATOUR

KEYWORDS: Operational research - Mathematical Optimization - Staff scheduling

FUNCTIONAL DESCRIPTION: This software is a prototype developed under a bilateral contract with the company Keolis. This software is a solver which aims to optimize the schedule of mediation staff. More precisely, for each member of the mediation staff working in a public transportation network, MEDIATOUR determines his/her schedule along the day, i.e. when and where he/she is present. Various operational constraints must be taken into account such as the coverage of the network. This software is designed to solve large-scale industrial instances (the subway network of Lille) in short computation times (less than 1 minute).

- Contact: Frederic Semet

6.3. PARROT

Planning Adapter Performing ReRouting and Optimization of Timing

KEYWORDS: Decision aid - Railway - Scheduling

FUNCTIONAL DESCRIPTION: This is a decision support system addressing the problem of the rescheduling railway schedules on the Belgian network when maintenance operations are planned in the short term (2-3 weeks in advance). The deliverable is a software tool that will take as input: (1) the schedules initially planned for the different trains, (2) the initial routes of the trains, (3) maintenance operations / changes of elements in the form of constraints (unavailable routes etc.). It then provides in output: (1) the new train schedule, (2) the new routing of the fleet. The modifications must respect the constraints corresponding to the operations of maintenance. For example, in some cases it is common to leave at least a few minutes interval between two trains using the same track in the station. This constraint must then be propagated if a maintenance operation delays the arrival of a train. New schedules and routings have to be created following a specific goal. Changes made to schedules and routings must minimize: (1) variations on the time spent at the station, (2) the number of partially canceled trains (additional correspondence (s) or stations that are no longer served), (2) the number of fully canceled trains (no stations served).

- Contact: Martine Labbé

7. New Results

7.1. Large scale complex structure optimization

Joint order batching and picker routing problem: Order picking is the process of retrieving products from inventory. It is mostly done manually by dedicated employees called pickers and is considered the most expensive of warehouse operations. To reduce the picking cost, customer orders can be grouped into batches that are then collected by traveling the shortest possible distance. We proposed an industrial case study for the HappyChic company where the warehouse has an acyclic layout: pickers are not allowed to backtrack. We developed a two-phase heuristic approach to solve this industrial case [64]. Moreover, we propose an exponential linear programming formulation to tackle the joint order batching and picker routing problem. Variables, or columns, are related to the picking routes in the warehouse. Computing such routes is generally an intractable routing problem and relates to the well known traveling salesman problem (TSP). Nonetheless, the rectangular warehouse's layouts can be used to efficiently solve the corresponding TSP and take into account in the development of an efficient subroutine, called oracle. We therefore investigate whether such an oracle allows for an effective exponential formulation. Experimented on a publicly available benchmark, the algorithm proves to be very effective. It improves many of the best known solutions and provides very strong lower bounds. Finally, this approach is also applied to the HappyChic industrial case to demonstrate its interest for this field of application [67], [45].

Logistics network design problem: Planning transportation operations within a supply chain is a difficult task that is often outsourced to logistics providers, in practice. At the tactical level, the problem of distributing products through a multi-echelon network is defined in the literature as the Logistics Service Network Design Problem (LSNDP). We study a LSNDP variant inspired by the management of restaurant supply chains. In this problem, a third party carrier seeks to cost-effectively source and fulfill customer demands of products through a tri-echelon supply chain composed of suppliers, warehouses, and customers. We propose an exact solution method based on partial Benders decompositions, where the master problem is strengthened by the addition of aggregated information derived from the subproblem. More specifically, we introduce a high-level dynamic Benders approach where the aggregated information used to strengthen the master is refined iteratively. In an extensive computational study, we demonstrate that our dynamic Benders strategy produces provably high-quality solutions and we validate the interest of refining the master problem in the course of a partial Benders decomposition-based scheme [72].

Multi commodity vehicle routing problem: We study vehicle routing problems considering multiple commodities, with applications in the local fresh food supply chains. The studied supply chain contains two echelons with three sets of actors: suppliers, distribution centers and customers. Suppliers are farmers that produce some fresh foods. Distribution centers are in charge of consolidation and delivery of the products to customers. Distribution centers collect products from the suppliers that perform direct trips. Products are delivered to the customers with a fleet of vehicles performing routes. Each customer requires several commodities, and the farmers produce a limited quantity of these commodities. For the minimization of the transportation cost, it is beneficial that a single customer is delivered by several vehicles. However, for the convenience of the customer, it is imposed that a single commodity is delivered at once by a single vehicle. Hence, different commodities have been explicitly considered. The complete problem is named Multi-Commodity two-echelon Distribution Problem (MC2DP). The restricted problem that addresses only the delivery from a single distribution center is named Commodity constrained Split Delivery Vehicle Routing Problem (C-SDVRP). We first propose a heuristic based on the Adaptive Large Neighborhood Search (ALNS) for the C-SDVRP [22]. Then, we address the whole problem (MC2DP) with collection and delivery operations and multiple distribution centers. In order to tackle this complex problem, we propose to decompose the problem: collection and delivery are sequentially solved [53], [54].

Generalized routing problems: We study routing problems that arise in the context of last mile delivery when multiple delivery options are proposed to the customers. The most common option to deliver packages is home/workplace delivery. Besides, the delivery can be made to pick-up points such as dedicated lockers or stores. In recent years, a new concept called trunk/in-car delivery has been proposed. Here, customers' packages can be delivered to the trunks of cars. Our goal is to model and develop efficient solution approaches for routing problems in this context, in which each customer can have multiple shipping locations. First, we study the single-vehicle case in the considered context, which is modeled as a Generalized Traveling Salesman Problem with Time Windows (GTSPTW). Four mixed integer linear programming formulations and an efficient branch-and-cut algorithm are proposed. Then, we study the multi-vehicle case which is denoted Generalized Vehicle Routing Problem with Time Windows (GVRPTW). An efficient column generation based heuristic is proposed to solve it [60], [61], [59], [37].

Joint management of demand and offer for last-mile delivery systems: E-commerce is a thriving market around the world and suits very well the busy lifestyle of today's customers. This growing in e-commerce poses a huge challenge for transportation companies, especially in the last mile delivery. We addressed first a fleet composition problem for last-mile delivery service. This problem occurs at a tactical level when the composition of the fleet has to be decided in advance. It is the case for companies that offer last-mile delivery service. Most of them subcontract the transportation part to local carriers and have to decide the day before which vehicles will be needed to cover a partially known demand. We assumed that the distribution area is divided into a limited number of delivery zones and the time horizon into time-slots. The demand is characterized by packages to be transported from pick-up zones to delivery zones given a delivery time slot. We have proposed an optimization problem aiming jointly to manage the offer and the demand. More precisely, discrete choice models representing the choices made by couriers and customers are integrated into the same optimization model. The originality of the contribution is based on the integration of variables of different natures in the same model and the development of integrated resolution methods. On the basis of a closed form of discrete choice models, we have reformulated the problem as a non-linear optimization problem. The resolution of this model by classical solvers requires coupling exact methods with heuristics in order to define a first initial solution [47], [58].

Delay Management in Public Transportation: The Delay Management Problem arises in Public Transportation networks, and is characterized by the necessity of connections between different vehicles. The attractiveness of Public Transportation networks is strongly related to the reliability of connections, which can be missed when delays or other unpredictable events occur. Given a single initial delay at one node of the network, the Delay Management Problem is to determine which vehicles have to wait for the delayed ones, with the aim of minimizing the dissatisfaction of the passengers. We derived strengthened mixed integer linear programming formulations and new families of valid inequalities for that problem. The implementation of branch-and-cut methods and tests on a benchmark of instances taken from real networks show the potential of the proposed formulations and cuts [20].

Discrete Ordered Median Problem: The discrete ordered median problem consists in locating p facilities in order to minimize an ordered weighted sum of distances between clients and closest open facility. We formulate this problem as a set partitioning problem using an exponential number of variables. Each variable corresponds to a set of demand points allocated to the same facility with the information of the sorting position of their corresponding costs. We develop a column generation approach to solve the continuous relaxation of this model. Then, we apply a branch-price-and-cut algorithm to solve small to large sized instances of DOMP in competitive computational time [21].

Genome wide association studies: We studied the Polymorphic Alu Insertion Recognition Problem (PAIRP). Alu (Arthrobacter luteus) forms a major component of repetitive DNA and are frequently encountered during the genotyping of individuals. The basic approach to find Alus consists of (1) aligning sequence reads from a set of individual(s) with respect to a reference genome and (2) comparing the possible Alu insertion induced by the alignment with the Alu insertions positions already known for the reference genome. The sequence genome of the reference individual is known and will be highly similar, but not identical, to the genome of the individual(s) being sequenced. Hence, at some locations they will diverge. Some of this divergence is due to the insertion of Alu polymorphisms. Detecting Alus has a central role in the field of Genetic Wide Association Studies because basic elements are a common source of mutation in humans. We investigated the PAIRP relationship with the the Clique Partitioning of Interval Graphs (CPIG). Our results [29] provide insights of the complexity of the problem, a characterization of its combinatorial structure and an exact approach based on Integer Linear Programming to exactly solve the correspond instances.

A branch-and-cut algorithm for the maximum k-balanced subgraph of a signed graph: A signed graph is k-balanced if its vertex set can be partitioned into at most k sets in such a way that positive edges are found only within the sets and negative edges go between sets. The maximum k-balanced subgraph problem is the problem of finding a subgraph of a graph G that is k-balanced and maximum according to the number of vertices. This problem has applications in clustering problems appearing in collaborative vs conflicting environments. We provide a representatives formulation for the problem and present a partial description of the associated polytope, including the introduction of strengthening families of valid inequalities. A branch-and-cut algorithm is described for finding an optimal solution to the problem. An ILS metaheuristic is implemented for providing primal bounds for this exact method and a branching rule strategy is proposed for the representatives formulation. Computational experiments, carried out over a set of random instances and on a set of instances from an application, show the effectiveness of the valid inequalities and strategies adopted in this work [75].

Feature Selection in Support Vector Machine: This work focuses on support vector machine (SVM) with feature selection. A MILP formulation is proposed for the problem. The choice of suitable features to construct the separating hyperplanes has been modelled in this formulation by including a budget constraint that sets in advance a limit on the number of features to be used in the classification process. We propose both an exact and a heuristic procedure to solve this formulation in an efficient way. Finally, the validation of the model is done by checking it with some well-known data sets and comparing it with classical classification methods [24].

7.2. Bilevel Programming

Pricing for Energy Management: Power systems face higher flexibility requirements from generation to consumption due to the increasing penetration of non-controllable distributed renewable energy. In this context, demand side management aims at reducing excessive load fluctuation and match the price of energy to their real cost for the grid. Pricing models for demand side management methods are traditionally used to control electricity demand. First, we proposed bilevel pricing models to explore the relationship between energy suppliers and customers who are connected to a smart grid. The smart grid technology allows customers to keep track of hourly prices and shift their demand accordingly, and allows the provider to observe the actual demand response to its pricing strategy. Moreover, we assumed that the smart grid optimizes the usage of a renewable energy generation source and a storage capacity. Results over a rolling horizon were obtained (Léonard Von Niederhausern PhD thesis [76]). Next, we considered four types of actors: furnishers sell electricity, local agents trade and consume energy, aggregators trade energy and provide energy to end-users, who consume it. This gives rise to three levels of optimization. The interaction between aggregators and their end-users is modeled with a bilevel program, and so is the interaction between furnishers, and local agents and aggregators. Since solving bilevel programs is difficult in itself, solving trilevel programs requires particular care. We proposed three possible approaches, two of them relying on a characterization of the intermediary optimization level [13], [76].

Finally, Time and-Level-of-Use is a recently proposed energy pricing scheme, designed for the residential sector and providing suppliers with robust guarantee on the consumption. We formulate the supplier decision as a bilevel, bi-objective problem optimizing for both financial loss and guarantee. A decomposition method is proposed, related to the optimal value transformation. It allows for the computation of an exact solution by finding possible Pareto optimal candidate solutions and then eliminating dominated ones. Numerical results on experimental residential power consumption data show the method effectively finds the optimal candidate solutions while optimizing costs only or incorporating risk aversion at the lower-level [38], [46].

Linear bilevel optimization: One of the most frequently used approaches to solve linear bilevel optimization problems consists in replacing the lower-level problem with its Karush–Kuhn–Tucker (KKT) conditions and by reformulating the KKT complementarity conditions using techniques from mixed-integer linear optimization. The latter step requires to determine some big-M constant in order to bound the lower level's dual feasible set such that no bilevel-optimal solution is cut off. In practice, heuristics are often used to find a big-M although it is known that these approaches may fail. In [69], we consider the hardness of two proxies for the above mentioned concept of a bilevel-correct big-M. First, we prove that verifying that a given big-M does not cut off any feasible vertex of the lower level's dual polyhedron cannot be done in polynomial time unless P = NP. Second, we show that verifying that a given big-M does not cut off any optimal point of the lower level's dual problem (for any point in the projection of the high-point relaxation onto the leader's decision space) is as hard as solving the original bilevel problem.

Market regulation: We proposed a bilevel programming model to study a problem of market regulation through government intervention. One of the main characteristics of the problem is that the government monopolizes the raw material in one industry, and competes in another industry with private firms for the production of commodities. Under this scheme, the government controls a state-owned firm to balance the market; that is, to minimize the difference between the produced and demanded commodities. On the other hand, a regulatory organism that coordinates private firms aims to maximize the total profit by deciding the amount of raw material bought from the state-owned firm. Two equivalent single-level reformulations are proposed to solve the problem. Additionally, three heuristic algorithms are designed to obtain good-quality solutions with low computational effort. Extensive computational experimentation is carried out to measure the efficiency of the proposed solution methodologies. A case study based on the Mexican petrochemical industry is presented. Additional instances generated from the case study are considered to validate the robustness of the proposed heuristic algorithms [28].

Product pricing: One of the main concerns in management and economic planning is to sell the right product to the right customer for the right price. Companies in retail and manufacturing employ pricing strategies to maximize their revenues. The Rank Pricing Problem considers a unit-demand model with unlimited supply and uniform budgets in which customers have a rank-buying behavior. Under these assumptions, the problem is first analyzed from the perspective of bilevel pricing models and formulated as a non linear bilevel program with multiple independent followers. We also present a direct non linear single level formulation. Two different linearizations of the models are carried out and two families of valid inequalities are obtained which, embedded in the formulations by implementing a branch-and-cut algorithm, allow us to tighten the upper bound given by the linear relaxation of the models. We show the efficiency of the formulations, the branch-and-cut algorithms and some preprocessing through extensive computational experiments [73].

Next in [68], we analyze a product pricing problem with single-minded customers, each interested in buying a bundle of products. The objective is to maximize the total revenue and we assume that supply is unlimited for all products. We contribute to a missing piece of literature by giving some mathematical formulations for this single-minded bundle pricing problem. We first present a mixed-integer nonlinear program with bilinear terms in the objective function and the constraints. By applying classical linearization techniques, we obtain two different mixed-integer linear programs. We then study the polyhedral structure of the linear formulations and obtain valid inequalities based on an RLT-like framework. We develop a Benders decomposition to project strong cuts from the tightest model onto the lighter models. We conclude this work with extensive numerical experiments to assess the quality of the mixed-integer linear formulations, as well as the performance of the cutting plane algorithms and the impact of the preprocessing on computation times.

Bilevel Minimum Spanning Tree Problem: Consider a graph G whose edge set is partitioned into a set of red edges and a set of blue edges, and assume that red edges are weighted and contain a spanning tree of G. Then, the Bilevel Minimum Spanning Tree Problem (BMSTP) consists in pricing (i.e., weighting) the blue edges in such a way that the total weight of the blue edges selected in a minimum spanning tree of the resulting graph is maximized. We propose different mathematical formulations for the BMSTP based on the properties of the Minimum Spanning Tree Problem and the bilevel optimization. We establish a theoretical and empirical comparison between these new formulations and we also provide reinforcements that together with a proper formulation are able to solve medium to big size instances [26].

Bilevel programming models for location problems: First, we addressed a multi-product location problem in which a retail firm has several malls with a known location. A particular product comes in p types. Each mall has a limited capacity for products to be sold at that location, so the firm has to choose what products to sold at what mall. Furthermore, the firm can apply discrete levels of discount on the products. The objective of the firm is to find what products to sell at which mall, with what level of discount, so that its profit is maximized. Consumers are located in points of the region. Each consumer has a different set of acceptable products, and will purchase one of these, or none if it is not convenient for her. Consumers maximize their utility. The agents (firm and consumers) play a Stackelberg game, in which the firm is the leader and the customers the follower. Once the firm decides the products to sell at each mall and the possible discounts, consumers purchase (or not) one of their acceptable products wherever their utility is maximized. We model the problem using bilevel formulations, which are compared on known instances from the literature [74]. Second we studied a location problem of controversial facilities. On the one hand, a leader chooses among a number of fixed potential locations which ones to establish. On the second hand, one or several followers who, once the leader location facilities have been set, choose their location points in a continuous framework. The leader's goal is to maximize some proxy to the weighted distance to the follower's location points, while the follower(s) aim is to locate his location points as close as possible to the leader ones. We develop the bilevel location model for one follower and for any polyhedral distance, and we extend it for several followers and any so-called p-norm. We prove the NP-hardness of the problem and propose different mixed integer linear programming formulations. Moreover, we develop alternative Benders decomposition algorithms for the problem. Finally, we report some computational results comparing the formulations and the Benders decompositions on a set of instances [23].

Stackelberg games: First we analyzed general Stackelberg games (SGs) and Stackelberg security games (SSGs). SGs are hierarchical adversarial games where players select actions or strategies to optimize their payoffs in a sequential manner. SSGs are a type of SGs that arise in security applications, where the strategies of the player that acts first consist in protecting subsets of targets and the strategies of the followers consist in attacking one of the targets. We review existing mixed integer optimization formulations in both the general and the security setting and present new formulations for the the second one. We compare the SG formulations and the SSG formulations both from a theoretical and a computational point of view. We identify which formulations provide tighter linear relaxations and show that the strongest formulation for the security version is ideal in the case of one single attacker. Our computational experiments show that the new formulations can be solved in shorter times [18].

Second, we formulate a Stackelberg Security game that coordinates resources in a border patrol problem. In this security domain, resources from different precincts have to be paired to conduct patrols in the border due to logistic constraints. Given this structure the set of pure defender strategies is of exponential size. We describe the set of mixed strategies using a polynomial number of variables but exponentially many constraints that come from the matching polytope. We then include this description in a mixed integer formulation to compute the Strong Stackelberg Equilibrium efficiently with a branch and cut scheme. Since the optimal patrol solution is a probability distribution over the set of exponential size, we also introduce an efficient sampling method that can be used to deploy the security resources every shift. Our computational results evaluate the efficiency of the branch and cut scheme developed and the accuracy of the sampling method. We show the applicability of the methodology by solving a real world border patrol problem [15].

Third, in [39], we discuss the impact of fairness constraints in Stackelberg Security Games. Fairness constraints can be used to avoid discrimination at the moment of implementing police patrolling. We present two ways of modelling fairness constraints, one with a detailed description of the population and the other with labels. We discuss the implementability of these constraints. In the case that the constraints are not implementable we present models to retrieve pure strategies in a way that they are the closest in average to the set of fairness constraints.

Finally, in [65], we focus on Stackelberg equilibria for discounted stochastic games. We begin by formalizing the concept of Stationary Strong Stackelberg Equlibrium (SSSE) policies for such games. We provide classes of games where the SSSE exists, and we prove via counterexamples that SSSE does not exist in the general case. We define suitable dynamic programming operators whose fixed points are referred to as Fixed Point Equilibrium (FPE). We show that the FPE and SSSE coincide for a class of games with Myopic Follower Strategy. We provide numerical examples that shed light on the relationship between SSSE and FPE and the behavior of Value Iteration, Policy Iteration and Mathematical programming formulations for this problem. Finally, we present a security application to illustrate the solution concepts and the efficiency of the algorithms studied.

7.3. Robust/Stochastic programming

Locating stations in a one-way electric car sharing system under demand uncertainty: In [16], we focused on a problem of locating recharging stations in one-way station based electric car sharing systems which operate under demand uncertainty. We modeled this problem as a mixed integer stochastic program and develop a Benders decomposition algorithm based on this formulation. We integrated a stabilization procedure to our algorithm and conduct a large-scale experimental study on our methods. To conduct the computational experiments, we developed a demand forecasting method allowing to generate many demand scenarios. The method was applied to real data from Manhattan taxi trips.

Bookings in the European Gas Market: Characterisation of Feasibility and Computational Complexity Results: As a consequence of the liberalisation of the European gas market in the last decades, gas trading and transport have been decoupled. At the core of this decoupling are so-called bookings and nominations. Bookings are special long-term capacity right contracts that guarantee that a specified amount of gas can be supplied or withdrawn at certain entry or exit nodes of the network. These supplies and withdrawals are nominated at the day-ahead. These bookings then need to be feasible, i.e., every nomination that complies with the given bookings can be transported. While checking the feasibility of a nomination can typically be done by solving a mixed-integer nonlinear feasibility problem, the verification of feasibility of a set of bookings is much harder. We consider the question of how to verify the feasibility of given bookings for a number of special cases. For our physics model we impose a steady-state potential-based flow model and disregard controllable network elements. We derive a characterisation of feasible bookings, which is then used to show that the problem is in coNP for the general case but can be solved in polynomial time for linear potential-based flow models. Moreover, we present a dynamic programming approach for deciding the feasibility of a booking in tree-shaped networks even for nonlinear flow models [25]. Further, in [71], we show that the feasibility of a booking also can be decided in polynomial time on single-cycle networks.

Robust bilevel programs: Bilevel optimization problems embed the optimality conditions of a sub-problem into the constraints of a decision-making process. A general question of bilevel optimization occurs where the lower-level is solved (only) to near-optimality. Solving bilevel problems under limited deviations of the lower-level variables was introduced under the term "ϵ-approximation" of the pessimistic bilevel problem. In [77] the authors define special properties and a solution method for this variant in the so-called independent case, i.e., where the lower-level feasible set is independent of the upper-level decision. In [66], we generalized the approach of *Wiesemann et al. 2013*, to problems with constraints involving upper- and lower-level variables in the constraints at both levels. The purpose of this generalization is to protect the upper-level feasibility against uncertainty of near-optimal solutions of the lower-level. We call this near-optimal robustness and the generalization is a near-optimal robust bilevel problem (NORBiP). NORBiP is a bilinear bilevel problem, and this makes it very hard in general. We have defined and implemented a solution algorithm for the linear linear NORBiP [66].

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Utocat (2018-2020): Study optimization problems arising in the blockchain

8.2. Bilateral Grants with Industry

- Program PGMO funded by the Fondation Mathématiques Jacques Hadamard. EDF is the industrial partner (2017-2019)
- Program PGMO funded by the Fondation Mathématiques Jacques Hadamard. A generic framework for routing and scheduling problems (2019-2021)
- Program PGMO funded by the Fondation Mathématiques Jacques Hadamard. Integrated models for the dimensioning and location of charging electric vehicles stations in the presence of renewable energy sources: Models and Algorithms (2019-2020)

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. ANR

ANR project AGIRE "Aide à la Gestion Intelligente des Ressources dans les Entrepôts - Decision system for smart management of resources in warehouses" in collaboration with Ecole des Mines de Saint-Etienne (Gardanne), IFSTTAR (Champs-sur-Marne), HappyChic (Tourcoing). This project addresses human resources management in warehouses which supply either sale points (B2B) or final consumers (B2C). Nowadays, such warehouses are under pressure. This is mainly due to the no inventory policy at the sale points and to the constant growth of e-commerce sales in France and Europe. In terms of logistics, this translates into an increasing number of parcels to prepare and to ship to satisfy an order, which is known typically a few hours before. Moreover, the total number of products to be packed varies very significantly from day-to-day by a factor of at least 3 (https://fr.wikipedia.org/wiki/Happychic).

The novelty of the project is twofold: (1) The human factor is explicitly be taken into account. It is integrated in the mathematical models and algorithms that are developed for the project. The aim is to improve the quality of employees' work ensuring the efficiency of the logistic system; (2) Problems at different decision levels are integrated and tackled jointly. At the tactical level, the main problematics are workload smoothing and the management of the storage zone. At operational level, the major issues concern the rearrangement of the picking zone, the picking tours, and the dynamic reorganization of activities to manage uncertainties.

ANR project PI-Commodality "Co-modal freight transportation chains: an approach based on physical internet" in collaboration with CGS-ARMINES (Paris), LAAS (Toulouse), DHL (2016 - 2019). The PI-co-modality project aims to design new sustainable logistic services between preset origins and destinations. It is based on innovative approaches both in terms of: (1) Logistics and transportation services: by considering the PI-internet approach, specifically: mesh logistics and transportation networks based on available capacities, by designing consistent integrated co-modal chains; (2) Methodology: by addressing the underlying problems according to two approaches: centralized and decentralized, by proposing news realistic models relevant for practitioner taking into account the consistency, by developing state-of-the-art decision making algorithms.

9.1.2. F.R.S.-FNRS (Belgium)

Bilevel optimization is a branch of mathematical optimization that deals with problems whose constraints embed an auxiliary optimization problem. The F.R.S.-FNRS research project "bilevel optimization" (2018-2019) will study such bilevel problems with bilinear objectives and simple second level problems. Each follower chooses one strategy in a given fixed set of limited size. Two classes of such problems will be studied: Pricing Problems and Stackelberg Security Games.

In pricing problems, prices for products must be determined to maximize the revenue of a leader given specific behaviors of customers (followers). More precisely, we will consider the single minded pricing problem and the rank pricing problem.

In Stackelberg games, mixed strategies to cover targets,must be determined in order to maximize the defender expected payoff given that attackers (followers) attack targets that maximize their own payoffs.

9.2. Regional Initiatives

9.2.1. Lille

The ELSAT research program addresses the issues involved in sustainable transportation and mobility. Within ELSAT, INOCS is involved on two projects devoted to hybrid optimization methods in logistics and to city logistics in collaboration with LAMIH (Université de Valenciennes), LGI2A (Université d'Artois) and LEOST (IFSTTAR). ELSAT is supported by the CPER 2015-2020 (State-Region Contract).

9.2.2. Brussels

ValueBugs is a citizen participatory research project, funded by INNOVIRIS (2018-2020). The objective of ValueBugs is to collectively develop a method for decentralized insect production in cities while enhancing the value of food waste on a small scale. In practical terms, peelings are consumed by insect larvae that have reached the end of their development and offer many promising outlets: feed for hens, farmed fish, pets... and much more! This new, totally innovative sector will be a new tool to be put in the hands of every citizen: we must therefore imagine it collectively.

9.3. International Initiatives

9.3.1. Inria International Labs

Inria Chile
Associate Team involved in the International Lab:

9.3.1.1. BIPLOS

Title: BIlevel Problems in LOgistics and Security

International Partner (Institution - Laboratory - Researcher):

University of Chile - Complex Engineering Systems Institute (ISCI) - Ordonez Fernando

Start year: 2017

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

This projet is devoted to bilevel optimisation problems with application in the security and logistics domains. Stackelberg games, including one defender and several followers, and competitive location problems will be considered. Mixed integer linear optimisation models and efficient algorithms to solve them will be developed.

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

9.3.2.1. LOBI

Title: Learning within Bilevel Optimization

International Partner (Institution - Laboratory - Researcher):

Polytechnique Montréal (Canada) - Research Group in Decision Analysis (GERAD) - Gilles Savard

Start year: 2018

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

The interplay between optimization and machine learning is one of the most important developments in modern computational science. Simultaneously there is a tremendous increase in the availability of large quantities of data in a multitude of applications, and a growing interest in exploiting the information that this data can provide to improve decision-making. Given the importance of big data in business analytics, its explicit integration into an optimization process is a challenge with high potential impact. The innovative project is concerned with the interconnection between machine learning approaches and a particular branch of optimization called bilevel optimization in this "big data" context. More precisely, we will focus on the development of new approaches integrating machine learning within bilevel optimization (LOBI: "learning au sein de l'Optimisation BIniveau") for two important practical applications, the pricing problem in revenue management and the energy ressource aggregation problem in smart grids. The applications arise from current industry collaborations of the teams involved, and will serve as testbeds to demonstrate the potential impact of the proposed approach.

9.3.2.2. North-European associated team

Title: Physical-internet services for city logistics

International Partner (Institution - Laboratory - Researcher):

Norwegian School of Economics - Stein Wallace

Start year: 2017

In this project, we consider an urban logistic terminal and new logistics services which could be developed according to a Physical Internet approach. The main objective is to evaluate the services using optimization models created within the project. We are developing optimization models to identify win-win cooperation between carriers based on supply and demand. We aim to explore how to include stochasticity in the description of the supplies and demands, as well as travel times, and to what extent the plans within a day can improve by such knowledge. The second task is to develop solution algorithms for these models. These are real scientific challenges as we are facing stochastic mixed integer problems.

9.3.3. Inria International Partners

9.3.3.1. Informal International Partners

Department of Statistics and Operations Research, University of Vienna, Austria

Centre for Quantitative Methods and Operations Management, HEC-Liège, Belgium

Interuniversity Centre on Entreprise Networks, Transportation and Logistics (CIRRELT), Montreal, Canada

Department of Industrial Engineering, University of Talca, Curicó, Chile

Complex Engineering Systems Institute (ISCI), University of Chile, Santiago, Chile

Department of Mathematics, Trier University, Germany

The Centre for Business Analytics, University College Dublin, Ireland

Department of Electrical, Electronic, and Information Engineering, University of Bologna, Italy

Department of Mathematics, University of Padova, Italy

Department of Electrical and Information Engineering, University of Padova, Italy

Department of Mathematics, University of Salerno, Italy

Department of Control and Computer Engineering, Politecnico di Torino, Italy

Department of Mathematics, University of Aveiro, Portugal

Department of Statistics and Operations Research, Universidade de Lisboa, Portugal

Department of Statistics and Operational Research, University of Murcia, Spain

Institute of Mathematics, University of Seville, Spain

Stewart School of Industrial and Systems Engineering, Georgia Tech Institute of Technology, USA

9.3.4. Participation in Other International Programs

9.3.4.1. Inria International Chairs

IIC ANJOS Miguel

Title: Power Peak Minimization for the Smart Gird

International Partner (Institution - Laboratory - Researcher):

Polytechnique Montréal (Canada) - Miguel Anjos

Duration: 2016 - 2020

Start year: 2016

9.4. International Research Visitors

9.4.1. Visits of International Scientists

- Yasemin Arda Da Silveira, HEC-Liège, University of Liège, Belgium, Mar 2019
- Maria Del Carmen Gale Pola, University of Zaragoza, Spain, Feb 2019
- Anton Kleywegt, Georgia Institute of Technology, USA, from Apr 2019 until May 2019
- Daniel Pereda Herrera, University of Chile, Chile, from Nov 2019
- Sebastián Dávila, University of Chile, Chile, from June 2019 until Dec 2019
- Natividad Gonzalez Blanco, University of Sevilla, Spain, from May 2019 until July 2019
- Federica Laureanam, University of Salerno, Italy, from Feb 2019 until May 2019

9.4.1.1. Internships

Sebastián Dávila, Ph.D. student at University of Chile, June to December 2019

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- 23rd Belgian Mathematical Optimization Workshop, La Roche, Belgium, April 25-26, 2019: Bernard Fortz (Organizer)

10.1.1.2. Member of the Organizing Committees

- 8th Winter School on Network Optimization, Estoril, Portugal, January 14-18, 2019: Bernard Fortz

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

LI Brazilian Symposium on Operational Research (LI SBPO), Limeira/SP - Brazil, September 2019: Martine Labbé

ORBEL 2019, Hasselt, Belgium, January 2019: Bernard Fortz, Martine Labbé

ROADEF2019 - 19ème Conférence de la Société Française de Recherche Opérationnelle et d'Aide à la Décision, Le Havre, France, February 2019: Luce Brotcorne, Diego Cattaruzza, Bernard Fortz, Frédéric Semet

INOC 2019, International Network Optimization Conference, Avignon, France, June 2019: Bernard Fortz

EURO 2019, European Conference of Operational Research Societies, Dublin, Ireland, June 2019: Bernard Fortz

5th JuliaCon, Baltimore, USA, July 22-26, 2019: Mathieu Besançon

TRISTAN X-Triennial Symposium on Transportation Analysis, Jun 2019, Hamilton Island, Australia, June 2019: Frédéric Semet

Workshop of the Transportation and Logistics Society, INFORMS, Jun 2019, Vienna, Austria, June 2019: Frédéric Semet

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- EURO Journal on Computational Optimization: Martine Labbé - Editor in chief, Bernard Fortz - Editor
- Computers and Operations Research: Luce Brotcorne - Member of the Advisory Board
- INFORMS Journal on Computing: Bernard Fortz - Associate editor
- International Transactions in Operations Research: Bernard Fortz, Martine Labbé - Associate editors
- Transportation Science: Martine Labbé - Member of the Advisory Board

10.1.3.2. Reviewer - Reviewing Activities

Annals of Operations Research, Applied Computing and Informatics, Central European Journal of Operations Research, Computers & Operations Research, Computational Optimization and Applications, Discrete Applied Mathematics, EURO Journal on Transportation and Logistics, European Journal of Operational Research, IISE Transactions, INFORMS Journal on Computing, International Journal of Management Science and Engineering Management, Mathematical Programming Computation, Networks, Omega, Operations Research, Optimization and Engineering, RAIRO - Operations Research, Transportation Science, IEEE Transactions on Power Systems, IEEE Transactions on Smart Grids, IEEE Power Engineering Letters: Mathieu Besançon, Luce Brotcorne, Diego Cattaruzza, Bernard Fortz, Martine Labbé, Maxime Ogier, Frédéric Semet.

10.1.4. Invited Talks

- Mathieu Besançon was invited speaker at the JuliaNantes workshop in Nantes, France, June 2019
- Martine Labbé was plenary speaker at CPAIOR in Tessalonik, Greece, June 2019
- Martine Labbé was plenary speaker at AMSI Optimise in Perth, Australia, June 2019
- Martine Labbé was plenary speaker at the 7th International Conference on Variable Neighborhood Search, Rabat, Marocco, October 2019
- Frédéric Semet was invited speaker at the Workshop on the Logistics of Autonomous Vessels, Bergen, Norway, May 2019

10.1.5. Leadership within the Scientific Community

EURO Working Group “Pricing and Revenue Management”: Luce Brotcorne - coordinator

EURO Working Group “European Network Optimization Group (ENOG)”: Bernard Fortz - coordinator

EURO Working Group “Vehicle routing and logistics optimization (VEROLOG)”: Frédéric Semet - Member of the board

INFORMS Women in OR/MS: Luce Brotcorne - International liaison

ORBEL (Belgian Operations Research Society): Bernard Fortz - Member of the board of administration

ORBEL representative for EURO and IFORS: Bernard Fortz

CNRS GdR 3002: Operations Research: Frédéric Semet - Member of the steering committee

10.1.6. Scientific Expertise

Scientific orientation committee of the Interuniversity Centre on Entreprise Networks, Transportation and Logistics (CIRRELT), Canada: Bernard Fortz, Frédéric Semet - Members

Center for Mathematics, Fundamental Applications and Operations Research, University of Lisbon: Martine Labbé - Member

DFG Review Panel “Mathematics” for Clusters of Excellence, 2018: Martine Labbé - Member

Scientific committee of France-Netherlands Exchange Program: Luce Brotcorne - Member

Evaluation committee for Inria/MITACS Exchange Program: Luce Brotcorne - Member

Evaluation committee COST GTRI: Luce Brotcorne - Member

Fund for research training in industry and agriculture (FRIA) PE1 - jury 1: Bernard Fortz - Member

Scientific board of PICOM competitiveness cluster: Frédéric Semet - Member

Agence Nationale de la Recherche (ANR): Luce Brotcorne, Frédéric Semet - Reviewer

Fond de Recherche Nature et Technologie du Québec: Frédéric Semet - Reviewer

Research Council of Norway: Frédéric Semet - Reviewer

Evaluation committee NSERC - EG 1509: Bernard Fortz - Member

10.1.7. Research Administration

Committee for the Technological Development (CDT): Luce Brotcorne - Member

CRIStAL: Frédéric Semet - Deputy-director

Scientific council of Centrale Lille: Frédéric Semet - Elected member

Scientific council of OPTIMA thematic group of CRIStAL: Diego Cattaruzza Member

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Bernard Fortz, Recherche Opérationnelle et Applications, 30hrs, M1, University of Mons (Charleroi campus), Belgium

Master: Bernard Fortz, Continuous Optimization, 24hrs, M1 & M2, Université libre de Bruxelles, Belgium

Master: Frédéric Semet, Non-linear Optimization, 30hrs, M2, Centrale Lille

Master: Frédéric Semet, Operations Research, 28hrs, M2, Centrale Lille

Master: Luce Brotcorne, Optimisation, 14hrs, M1, Polytech Lille

Master: Luce Brotcorne, Recherche opérationnelle, 16hrs, M1 apprentissage, Polytech Lille

Master: Diego Cattaruzza, Maxime Ogier, Frédéric Semet, Prescriptive analytics and optimization, 64hrs, M1, Centrale Lille

Master: Diego Cattaruzza, Maxime Ogier, Object-Oriented Programming, 48hrs, M1, Centrale Lille

Master: Diego Cattaruzza, Maxime Ogier, Operations Research, 16hrs, M1, Centrale Lille

Licence: Diego Cattaruzza, Maxime Ogier, Object-Oriented Programming, 36hrs, L3, Centrale Lille

Licence: Frédéric Semet, Advanced programming and Complexity, 24hrs, L3, Centrale Lille

Licence: Diego Cattaruzza, Maxime Ogier, Object-Oriented Programming, 40hrs, L2, Centrale Lille

Licence: Bernard Fortz, Algorithmique 1, 12hrs, L1, Université libre de Bruxelles, Belgium

Licence: Bernard Fortz, Algorithmique et Recherche Opérationnelle, 24hrs, L3, Université libre de Bruxelles, Belgium

10.2.2. Supervision

PhD: Simon Bélières, Mathematical programming for tactical transportation planning in a multi-product supply chain, November 2019, Nicolas Jozefowiez, Frédéric Semet

PhD: Wenjuan Gu, Location routing for short and local fresh food supply chain, November 2019, Maxime Ogier, Frédéric Semet

PhD: Yuan Yuan, Vehicle Routing Problems with Synchronization for City Logistics, October 2019, Diego Cattaruzza, Frédéric Semet

PhD: Léonard Von Niederhausern, Design and pricing of new services in energy in a competitive environment, March 2019, Luce Brotcorne, Didier Aussel

PhD in progress: Luis Alberto Salazar Zendeja, Formulations and resolution methods for network interdiction problems, from November 2018, Diego Cattaruzza, Martine Labbé, Frédéric Semet

PhD in progress: Matteo Petris, Column generation approaches for integrated operationnal problems, from October 2019, Diego Cattaruzza, Maxime Ogier, Frédéric Semet

PhD in progress: Concepción Domínguez Sánchez, Mixed Integer Linear Models and Algorithms for Pricing Problems, from October 2017, Martine Labbé

PhD in progress: Moises Rodriguez Madrena, Problems in data analysis and location Theory, from January 2019, Martine Labbé, Justo Puerto

PhD in progress: Fränk Plein, Models and methods for the robust verification of booked capacities in gas networks in a decentralized setting, from October 2017, Martine Labbé

PhD in progress: Jérôme De Boeck, Optimization problems in energy, from October 2015, Bernard Fortz

PhD in progress: Mathieu Besançon, Approche bi-niveau de réponse à la demande dans les réseaux électriques intelligents, from September 2018, Miguel Anjos, Luce Brotcorne, Frédéric Semet

PhD in progress: Yaheng Cui, Models and methods for decentralized decision in logistics networks, from Oct 2016, Luce Brotcorne, Eric Ballot

10.2.3. Juries

- Ikram Bouras, PhD, Université de Montpellier, 2019, “Fixed Charge Network Design Problem with User Optimal flows”: Luce Brotcorne (reviewer)
- Umar Hashmi, PhD, Université PSL, Paris, 2019, “Optimization and Control of Storage in Smart Grids”: Luce Brotcorne (reviewer)
- Léonard von Niederhäusern, PhD, Centrale Lille, 2019, “Design and pricing of new energy services in a competitive environment”: Bernard Fortz (president)
- Meihui Gao, PhD, Université de Lorraine, 2019, “Models and methods for Network Function Virtualization (NFV) architectures”: Bernard Fortz (reviewer)
- Yuan Yuan, PhD, Centrale Lille, 2019, “Models and Algorithms for Last Mile Delivery Problems with Multiple Shipping Options”: Martine Labbé (member)
- José Neto, HDR, Université Evry, 2019, “Some contributions to mathematical programming and combinatorial optimization”: Bernard Fortz (reviewer)
- Marc-Antoine Coindreau, PhD, University of Lausanne, Switzerland, 2019, “Managing Advanced Synchronization Aspects in Logistics Systems”: Frédéric Semet (reviewer)
- Ruslan Sadykov, HDR, Université de Bordeaux, 2019, “Modern Branch-Cut-and-Price”: Frédéric Semet (reviewer)

10.3. Popularization

- Luce Brotcorne, Journées scientifiques de l’Inria, Lyon, June 2019
- Frédéric Semet, Cité de l’IA - Intelligence Artificielle, MEDEF Lille Métropole, April 2019
- Maxime Ogier, Frédéric Semet, 4ème séminaire humAIn : IA, optimisation et retail - logistique, Polytech Lille, October 2019

10.3.1. Internal or external Inria responsibilities

- Frédéric Semet, 30 minutes de sciences, Inria Lille Nord Europe, April 2019

11. Bibliography

Major publications by the team in recent years

[1] N. Absi, D. Cattaruzza, D. Feillet, M. Ogier, F. Semet.*A heuristic branch-cut-and-price algorithm for the ROADEF/EURO challenge on Inventory Routing*, in "Transportation Science", 2019, forthcoming, https://hal-emse.ccsd.cnrs.fr/emse-02163171

[2] S. Afsar, L. Brotcorne, P. Marcotte, G. Savard.*Achieving an optimal trade-off between revenue and energy peak within a smart grid environment*, in "Renewable Energy", March 2016, https://hal.inria.fr/hal-01230915

[3] Q. Botton, B. Fortz, L. Gouveia, M. Poss.*Benders Decomposition for the Hop-Constrained Survivable Network Design Problem*, in "INFORMS Journal on Computing", 2013, vol. 25, n^{o} 1, p. 13-26 [*DOI :* 10.1287/ijoc.1110.0472], http://joc.journal.informs.org/content/25/1/13.abstract

[4] L. Brotcorne, D. Aussel, S. Lepaul, L. Von Niederha¨usen.*A Trilevel Model for Best Response in Energy Demand-Side Management*, in "European Journal of Operational Research", 2020, https://hal.inria.fr/hal-02414600

[5] L. Brotcorne, F. Cirinei, P. Marcotte, G. Savard.*An exact algorithm for the network pricing problem*, in "Discrete Optimization", 2011, vol. 8, n° 2, p. 246–258, https://dx.doi.org/10.1016/j.disopt.2010.09.003

[6] L. Brotcorne, M. Labbé, P. Marcotte, G. Savard.*Joint design and pricing on a network*, in "Operation Research", 2008, vol. 56, p. 1104-1115, https://dx.doi.org/10.1287/opre.1080.0617

[7] V. Dal Sasso, L. De Giovanni, M. Labbé.*Strengthened Formulations and Valid Inequalities for Single Delay Management in Public Transportation*, in "Transportation Science", 2019, vol. 53, n° 5, p. 1213-1499, https://hal.inria.fr/hal-01925451

[8] B. Fortz, E. Gorgone, D. Papadimitriou.*A Lagrangian heuristic algorithm for the time-dependent combined network design and routing problem*, in "Networks", 2017, vol. 69, n° 1, p. 110–123, http://dx.doi.org/10.1002/net.21721

[9] B. Fortz, M. Labbé, F. Louveaux, M. Poss.*Stochastic binary problems with simple penalties for capacity constraints violations*, in "Mathematical Programming", 2013, vol. 138, p. 199-221, http://dx.doi.org/10.1007/s10107-012-0520-4

[10] M. Restrepo, F. Semet, T. Pocreau.*Integrated Shift Scheduling and Load Assignment Optimization for Attended Home Delivery*, in "Transportation Science", 2019, vol. 53, p. 917-1212, https://hal.inria.fr/hal-01963916

Publications of the year

Articles in International Peer-Reviewed Journal

[11] N. Absi, D. Cattaruzza, D. Feillet, M. Ogier, F. Semet.*A heuristic branch-cut-and-price algorithm for the ROADEF/EURO challenge on Inventory Routing*, in "Transportation Science", 2019, forthcoming, https://hal-emse.ccsd.cnrs.fr/emse-02163171

[12] E. Alekseeva, L. Brotcorne.*Bilevel approach to optimize electricity prices*, in "Yugoslav Journal of Operations Research", 2019, https://hal.inria.fr/hal-02414645

[13] L. Brotcorne, D. Aussel, S. Lepaul, L. Von Niederha¨usen.*A Trilevel Model for Best Response in Energy Demand-Side Management*, in "European Journal of Operational Research", 2020, https://hal.inria.fr/hal-02414600

[14] L. Brotcorne, G. Perboli, M. A. Rosano, Q. Wei.*A Managerial Analysis of Urban Parcel Delivery: A Lean Business Approach*, in "Sustainability", 2019, https://hal.inria.fr/hal-02414629

[15] V. Bucarey, C. Casorrán, M. Labbé, F. Ordóñez, O. Figueroa.*Coordinated defender strategies for border patrols*, in "European Journal of Operational Research", 2019, forthcoming, https://hal.inria.fr/hal-01917782

[16] H. Calik, B. Fortz.*A Benders decomposition method for locating stations in a one-way electric car sharing system under demand uncertainty*, in "Transportation Research Part B: Methodological", July 2019, vol. 125, p. 121-150 [*DOI :* 10.1016/j.trb.2019.05.004], https://hal.inria.fr/hal-02409510

[17] H. CALVETE, C. DOMÍNGUEZ, C. GALÉ, M. LABBÉ, A. MARIN.*The Rank Pricing Problem: models and branch-and-cut algorithms*, in "Computers and Operations Research", May 2019, vol. 105, p. 12-31, https://hal.inria.fr/hal-01782958

[18] C. CASORRÁN, B. FORTZ, M. LABBÉ, F. ORDÓÑEZ.*A study of general and security Stackelberg game formulations*, in "European Journal of Operational Research", 2019, vol. 278, n^{o} 3, p. 855 - 868 [*DOI :* 10.1016/J.EJOR.2019.05.012], https://hal.inria.fr/hal-01917798

[19] M. A. D. R. S. CRUZ, H. R. D. O. ROCHA, M. H. M. PAIVA, M. E. V. SEGATTO, E. CAMBY, G. CAPOROSSI.*An algorithm for cost optimization of PMU and communication infrastructure in WAMS*, in "International Journal of Electrical Power and Energy Systems", March 2019, vol. 106, p. 96-104 [*DOI :* 10.1016/J.IJEPES.2018.09.020], https://hal.inria.fr/hal-01944258

[20] V. DAL SASSO, L. DE GIOVANNI, M. LABBÉ.*Strengthened Formulations and Valid Inequalities for Single Delay Management in Public Transportation*, in "Transportation Science", 2019, vol. 53, n^{o} 5, p. 1213-1499 [*DOI :* 10.1287/TRSC.2018.0874], https://hal.inria.fr/hal-01925451

[21] S. DELEPLANQUE, M. LABBÉ, D. PONCE LOPEZ, J. PUERTO.*A Branch-Price-and-Cut Procedure for the Discrete Ordered Median Problem*, in "INFORMS Journal on Computing", 2019, forthcoming, https://hal.inria.fr/hal-01954865

[22] W. GU, D. CATTARUZZA, M. OGIER, F. SEMET.*Adaptive large neighborhood search for the commodity constrained split delivery VRP*, in "Computers and Operations Research", 2019, vol. 112, 104761 [*DOI :* 10.1016/J.COR.2019.07.019], https://hal.archives-ouvertes.fr/hal-02317246

[23] M. LABBÉ, M. LEAL, J. PUERTO.*New models for the location of controversial facilities: A bilevel programming approach*, in "Computers and Operations Research", 2019, vol. 107, p. 95 -106 [*DOI :* 10.1016/J.COR.2019.03.003], https://hal.inria.fr/hal-01933601

[24] M. LABBÉ, L. I. MARTÍNEZ-MERINO, A. M. RODRÍGUEZ-CHÍA.*Mixed Integer Linear Programming for Feature Selection in Support Vector Machine*, in "Discrete Applied Mathematics", 2019, vol. 261, p. 276-304 [*DOI :* 10.1016/J.DAM.2018.10.025], https://hal.inria.fr/hal-01924379

[25] M. LABBÉ, F. PLEIN, M. SCHMIDT.*Bookings in the European Gas Market: Characterisation of Feasibility and Computational Complexity Results*, in "Optimization and Engineering", 2019, forthcoming [*DOI :* 10.1007/S11081-019-09447-0], https://hal.inria.fr/hal-01954262

[26] M. LABBÉ, M. POZO, J. PUERTO.*Computational comparisons of different formulations for the Bilevel Minimum Spanning Tree Problem*, in "International Transactions in Operational Research", 2019, forthcoming [*DOI :* 10.1111/ITOR.12680], https://hal.inria.fr/hal-01937013

[27] P. H. MACÊDO DE ARAÚJO, M. CAMPÊLO, R. CORREA, M. LABBÉ.*The Geodesic Classification Problem on Graphs*, in "Electronic Notes in Theoretical Computer Science", 2019, vol. 346, p. 65 - 76 [*DOI :* 10.1016/J.ENTCS.2019.08.007], https://hal.inria.fr/hal-02393316

[28] H. MARAVILLO, J.-F. CAMACHO-VALLEJO, J. PUERTO, M. LABBÉ.*A market regulation bilevel problem: a case study of the Mexican petrochemical industry*, in "Omega", 2019, forthcoming [*DOI :* 10.1016/J.OMEGA.2019.08.012], https://hal.inria.fr/hal-01922518

[29] L. PORRETTA, D. CATANZARO, B. HALLDÓRSSON, B. FORTZ.*A Branch & Price algorithm for the minimum cost clique cover problem in max-point tolerance graphs*, in "4OR: A Quarterly Journal of Operations Research", March 2019, vol. 17, n^{o} 1, p. 75-96 [*DOI :* 10.1007/S10288-018-0377-3], https://hal.inria.fr/hal-01943982

[30] M. RESTREPO, F. SEMET, T. POCREAU.*Integrated Shift Scheduling and Load Assignment Optimization for Attended Home Delivery*, in "Transportation Science", 2019, vol. 53, n^{o} 4, p. 917-1212 [*DOI :* 10.1287/TRSC.2018.0857], https://hal.inria.fr/hal-01963916

Invited Conferences

[31] L. BROTCORNE, A. JOFRÉ, D. PEREDA.*Bilevel optimization applied to strategic pricing in electricity markets and extension to markets with massive entry of renewable energies and distributed generation*, in "PGMO", Paris, France, November 2019, https://hal.inria.fr/hal-02414774

[32] M. LABBÉ.*Bilevel optimisation and pricing problems*, in "The 7th International Conference on Variable Neighborhood Search", Rabat, Morocco, October 2019, https://hal.inria.fr/hal-02393387

[33] M. LABBÉ.*Bilevel optimisation, Stackelberg games and pricing problems*, in "CPAIOR", Thessalonik, Greece, June 2019, https://hal.inria.fr/hal-02393354

[34] M. LABBÉ.*Bilevel programming and price optimization problems*, in "Workshop PGMO "Jeux de pilotage de flexibilités de consommation électrique : dynamique et aspect composite"", Paris, France, April 2019, https://hal.inria.fr/hal-02393340

[35] M. LABBÉ.*Bilevel programming, Stackelberg games and pricing problems*, in "AMSI Optimise", Perth, Australia, June 2019, https://hal.inria.fr/hal-02393373

[36] M. LABBÉ.*EURO Gold Medal Laureate Lecture*, in "EURO Conference", Dublin, Ireland, June 2019, https://hal.inria.fr/hal-02393379

[37] F. SEMET, Y. YUAN, D. CATTARUZZA, M. OGIER, D. VIGO.*Combining different types of services in the last-mile delivery*, in "Workshop on the Logistics of Autonomous Vessels", Bergen, Norway, May 2019, https://hal.inria.fr/hal-02419302

International Conferences with Proceedings

[38] L. BROTCORNE, M. BESANÇON, M. F. ANJOS, J. A. GOMEZ-HERRERA.*A Bilevel Approach to Optimal Price-Setting of Time-and-Level-of-Use Tariffs*, in "Eurogen19", Guimaraes, Portugal, September 2019, https://hal.inria.fr/hal-02414702

[39] V. BUCAREY, M. LABBÉ.*Discussion of Fairness and Implementability in Stackelberg Security Games*, in "GameSec 2019 - 10th Conference on Decision and Game Theory for Security", Stockholm, Sweden, October 2019, https://hal.inria.fr/hal-02267147

Conferences without Proceedings

[40] Y. ARDA, D. CATTARUZZA, V. FRANCOIS, M. OGIER.*Home Chemotherapy Planning: An Integrated Production Scheduling and Multi-Trip Vehicle Routing Problem*, in "EURO 2019 - 30th European Conference On Operational Research", Dublin, Ireland, June 2019, https://hal.archives-ouvertes.fr/hal-02317435

[41] Y. ARDA, D. CATTARUZZA, V. FRANCOIS, M. OGIER.*Home Chemotherapy Planning: An Integrated Production Scheduling and Multi-Trip Vehicle Routing Problem*, in "VeRoLog 2019 - Workshop of the EURO Working Group on Vehicle Routing and Logistics optimization", Sevilla, Spain, June 2019, https://hal.archives-ouvertes.fr/hal-02317439

[42] Y. ARDA, D. CATTARUZZA, V. FRANCOIS, M. OGIER.*Synchronize production and administration tasks for home chemotherapy services*, in "SynchroTrans 2019 - Second International Workshop on Synchronization in Transport", Nantes, France, September 2019, https://hal.archives-ouvertes.fr/hal-02317426

[43] S. BELIERES, M. HEWITT, N. JOZEFOWIEZ, F. SEMET, T. VAN WOENSEL.*Partial Benders decomposition for logistics network design*, in "ROADEF 2019 – 20ème Congrès Annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision", Le Havre, France, February 2019, https://hal.inria.fr/hal-02421624

[44] M. BESANÇON, L. BROTCORNE, M. F. ANJOS.*A Bilevel Model for Large-scale Time-and-Level-of-Use Pricing*, in "ROADEF 2019 - 20ème Congrès Annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision", Le Havre, France, February 2019, https://hal.inria.fr/hal-02055738

[45] O. BRIANT, H. CAMBAZARD, D. CATTARUZZA, N. CATUSSE, A.-L. LADIER, M. OGIER.*A column generation approach for the joint order batching and picker routing problem*, in "Workshop of the EURO Working Group on Vehicle Routing and Logistics optimization (VeRoLog)", Seville, Spain, June 2019, https://hal.archives-ouvertes.fr/hal-02317032

[46] L. BROTCORNE, M. BESANÇON, J. A. GOMEZ-HERRERA, M. F. ANJOS.*Time of use pricing for demand response*, in "Informs 19", Seattle, United States, October 2019, https://hal.inria.fr/hal-02414755

[47] L. BROTCORNE, A. KLEYWEGT, Y. MAGNOUCHE.*Pricing for Drivers and Customers for Goods Deliveries, Triennal Symposium on Transportation*, in "Tristan 19", Hamilton Island, Australia, June 2019, https://hal.inria.fr/hal-02414733

[48] B. FORTZ.*New models and preprocessing techniques for segment routing optimization*, in "Optimization Days 2019", Montréal, Canada, May 2019, https://hal.inria.fr/hal-02409764

[49] B. FORTZ.*New models and preprocessing techniques for segment routing optimization*, in "International Network Optimization Conference (INOC) 2019", Avignon, France, June 2019, https://hal.inria.fr/hal-02409773

[50] B. FORTZ.*New models and preprocessing techniques for segment routing optimization*, in "Operations Research 2019", Dresden, Germany, September 2019, https://hal.inria.fr/hal-02409784

[51] B. FORTZ.*New models and preprocessing techniques for segment routing optimization*, in "INFORMS Annual Meeting 2019", Seattle, United States, October 2019, https://hal.inria.fr/hal-02409791

[52] B. FORTZ, M. LEITNER.*Using variables aggregation and Benders decomposition for solving large-scale extended formulations*, in "ROADEF 2019 - 20ème Congrès Annuel de le Société Française de Recherche Opérationnelle et d'Aide à la Décision", Le Havre, France, February 2019, https://hal.inria.fr/hal-02409714

[53] W. GU, C. ARCHETTI, D. CATTARUZZA, M. OGIER, F. SEMET, M. G. SPERANZA.*A multi-commodity transportation planning problem in supply chain management*, in "ROADEF 2019 – 20ème Congrès Annuel

de la Société Française de Recherche Opérationnelle et d'Aide à la Décision", Le Havre, France, February 2019, https://hal.archives-ouvertes.fr/hal-02131675

[54] W. GU, C. ARCHETTI, D. CATTARUZZA, M. OGIER, F. SEMET, M. G. SPERANZA.*Sequential approaches to solve a multi-commodity transportation planning problem*, in "VeRoLog 2019 - Workshop of the EURO Working Group on Vehicle Routing and Logistics optimization", Seville, Spain, June 2019, https://hal.archives-ouvertes.fr/hal-02318115

[55] M. LABBÉ, S. DAVILA, F. ORDÓÑEZ, V. MARIANOV, F. SEMET.*Bilevel programming models for multi-product location problems*, in "IX International Workshop on Locational Analysis and Related Problems-Bilevel programming models for multi-product location problems", Cadiz, Spain, January 2019, https://hal.inria.fr/hal-02393329

[56] A. LAURENT, L. BROTCORNE, B. FORTZ.*Modélisation du problème de pricing dans la blockchain Ethereum*, in "ROADEF 2019 - 20ème Congrès Annuel de le Société Française de Recherche Opérationnelle et d'Aide à la Décision", Le Havre, France, February 2019, https://hal.inria.fr/hal-02409724

[57] P. NOLZ, N. ABSI, D. CATTARUZZA, D. FEILLET.*Two-echelon distribution with city hub capacity management*, in "SynchroTrans 2019 - Second International Workshop on Synchronization in Transport", Nantes, France, September 2019, https://hal.archives-ouvertes.fr/hal-02317431

[58] F. SEMET, E. ALEKSEEVA, L. BROTCORNE, Y. MAGNOUCHE, E. SOUFFLET.*Estimating Vehicle Fleet Composition for Last-Mile Delivery Service*, in "TRISTAN X-Triennial Symposium on Transportation Analysis", Hamilton Island, Australia, June 2019, https://hal.inria.fr/hal-02374921

[59] F. SEMET, Y. YUAN, D. CATTARUZZA, M. OGIER, D. VIGO.*A column generation based heuristic for the generalized vehicle routing problem with time windows*, in "Journées de l'Optimisation 2019", Montréal, Canada, May 2019, https://hal.inria.fr/hal-02419289

[60] Y. YUAN, D. CATTARUZZA, M. OGIER, F. SEMET, D. VIGO.*The generalized vehicle routing problem with time windows*, in "ROADEF2019 - 20ème Congrès Annuel de la Société Française de Recherche Opérationnelle et d'Aide à la Décision", Le Havre, France, February 2019, https://hal.archives-ouvertes.fr/hal-02383126

[61] Y. YUAN, D. CATTARUZZA, M. OGIER, F. SEMET, D. VIGO.*The generalized vehicle routing problem with time windows*, in "Workshop of the EURO Working Group on Vehicle Routing and Logistics optimization (VeRoLog 2019)", Seville, Spain, June 2019, https://hal.archives-ouvertes.fr/hal-02383149

[62] Y. YUAN, Y. LYU, X. SHEN, I. TSANG, D.-Y. YEUNG.*Marginalized Average Attentional Network for Weakly-Supervised Learning*, in "ICLR 2019 - Seventh International Conference on Learning Representations", New-Orleans, United States, May 2019, https://hal-enpc.archives-ouvertes.fr/hal-02057597

Scientific Books (or Scientific Book chapters)

[63] *Operations Research Proceedings 2018*, 2019 [*DOI :* 10.1007/978-3-030-18500-8], https://hal.inria.fr/hal-02409572

[64] M. BUÉ, D. CATTARUZZA, M. OGIER, F. SEMET.*A Two-Phase Approach for an Integrated Order Batching and Picker Routing Problem*, in "A View of Operations Research Applications in Italy, 2018", Springer, September 2019, https://hal.archives-ouvertes.fr/hal-01849980

Research Reports

[65] V. BUCAREY, E. DELLA VECCHIA, A. JEAN-MARIE, F. ORDÓÑEZ.*Stationary Strong Stackelberg Equilibrium in Discounted Stochastic Games*, Inria, May 2019, n° RR-9271, 62, https://hal.inria.fr/hal-02144095

Other Publications

[66] M. BESANÇON, M. F. ANJOS, L. BROTCORNE.*Near-optimal robust bilevel optimization*, November 2019, https://arxiv.org/abs/1908.04040 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02414848

[67] O. BRIANT, H. CAMBAZARD, D. CATTARUZZA, N. CATUSSE, A.-L. LADIER, M. OGIER.*An efficient and general approach for the joint order batching and picker routing problem*, March 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02078547

[68] V. BUCAREY, S. ELLOUMI, M. LABBÉ, F. PLEIN.*Models and Algorithms for the Product Pricing with Single-Minded Customers Requesting Bundles*, March 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02056763

[69] T. KLEINERT, M. LABBÉ, F. PLEIN, M. SCHMIDT.*There's No Free Lunch: On the Hardness of Choosing a Correct Big-M in Bilevel Optimization*, June 2019, working paper or preprint, https://hal.inria.fr/hal-02106642

[70] M. LABBÉ, P. MARCOTTE.*Bilevel Network Design*, 2019, working paper or preprint, https://hal.inria.fr/hal-01937014

[71] M. LABBÉ, F. PLEIN, M. SCHMIDT, J. THÜRAUF.*Deciding Feasibility of a Booking in the European Gas Market on a Cycle is in P*, November 2019, working paper or preprint, https://hal.inria.fr/hal-02360972

References in notes

[72] S. BELIERES, M. HEWITT, N. JOZEFOWIEZ, F. SEMET, T. VAN WOENSEL.*Partial Benders decomposition for logistics network design*, in "ROADEF 2019 – 20ème Congrès de la Société Française de Recherche Opérationnelle et d'Aide à la Décision", Le Havre, France, February 2019, https://hal.inria.fr/hal-02421624

[73] H. CALVETE, C. DOMÍNGUEZ, C. GALÉ, M. LABBÉ, A. MARIN.*The Rank Pricing Problem: models and branch-and-cut algorithms*, in "Computers and Operations Research", 2018, forthcoming, https://hal.inria.fr/hal-01782958

[74] S. DAVILA, M. LABBÉ, V. MARIANOV, F. ORDÓÑEZ, F. SEMET.*Bilevel programming models for multi-product location problems*, in "IWOBIP'18 - 2nd International Workshop on Bilevel Programming", Lille, France, June 2018, https://hal.inria.fr/hal-01964681

[75] R. FIGUEIREDO, Y. FROTA, M. LABBÉ.*A branch-and-cut algorithm for the maximum k-balanced subgraph of a signed graph*, in "Discrete Applied Mathematics", 2018, forthcoming, https://hal.inria.fr/hal-01937015

[76] L. VON NIEDERHA¨USEN.*Design and pricing of new energy services in a competitive environment*, Ecole Centrale de Lille, April 2019, https://tel.archives-ouvertes.fr/tel-02397416

[77] W. WIESEMANN, A. TSOUKALAS, P.-M. KLENIATI, B. RUSTEM.*Pessimistic bilevel optimization*, in "Journal on Optimization", 2013, vol. 23, p. 353–380

Project-Team LINKS

Linking Dynamic Data

IN COLLABORATION WITH: Centre de Recherche en Informatique, Signal et Automatique de Lille

IN PARTNERSHIP WITH:
CNRS

Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team LINKS

Creation of the Team: 2013 January 01, updated into Project-Team: 2016 June 01

Keywords:

Computer Science and Digital Science:

A2.1. - Programming Languages
A2.1.1. - Semantics of programming languages
A2.1.4. - Functional programming
A2.1.6. - Concurrent programming
A2.4. - Formal method for verification, reliability, certification
A2.4.1. - Analysis
A2.4.2. - Model-checking
A2.4.3. - Proofs
A3.1. - Data
A3.1.1. - Modeling, representation
A3.1.2. - Data management, quering and storage
A3.1.3. - Distributed data
A3.1.4. - Uncertain data
A3.1.5. - Control access, privacy
A3.1.6. - Query optimization
A3.1.7. - Open data
A3.1.8. - Big data (production, storage, transfer)
A3.1.9. - Database
A3.2.1. - Knowledge bases
A3.2.2. - Knowledge extraction, cleaning
A3.2.3. - Inference
A3.2.4. - Semantic Web
A4.7. - Access control
A4.8. - Privacy-enhancing technologies
A7. - Theory of computation
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
B6.3.4. - Social Networks
B6.5. - Information systems
B9.5.1. - Computer science
B9.5.6. - Data science
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientist
Joachim Niehren [Team leader, Inria, Senior Researcher, HDR]

Faculty Members
Iovka Boneva [Université de Lille, Associate Professor]
Florent Capelli [Université de Lille, Associate Professor]
Aurélien Lemay [Université de Lille, Associate Professor, HDR]
Charles Paperman [Université de Lille, Associate Professor]
Sylvain Salvati [Université de Lille, Professor, HDR]
Slawomir Staworko [Université de Lille, Associate Professor, HDR]
Sophie Tison [Université de Lille, Professor, HDR]

Technical Staff
Antonio Al Serhali [Inria, Engineer, from Nov 2019]
Jeremie Dusart [Inria, Engineer]

PhD Students
Nicolas Crosetti [Inria]
Lily Gallois [Université de Lille]
Paul Gallot [Inria]
Jose Martin Lozano [Université de Lille]
Momar Sakho [Inria]

Post-Doctoral Fellow
Bruno Guillon [Inria, until Aug 2019]

Visiting Scientists
Rustam Azimov [JetBrains, Saint-Petersburg State University (Russia), Apr 2019]
Semyon Grigorev [JetBrains, Saint-Petersburg State University (Russia), Apr 2019]

Administrative Assistant
Nathalie Bonte [Inria]

2. Overall Objectives

2.1. Overall Objectives

We will develop algorithms for answering logical querying on heterogeneous linked data collections in hybrid formats, distributed programming languages for managing dynamic linked data collections and workflows based on queries and mappings, and symbolic machine learning algorithms that can link datasets by inferring appropriate queries and mappings.

2.2. Presentation

The following three paragraphs summarize our main research objectives.

Querying Heterogeneous Linked Data We develop new kinds of schema mappings for semi-structured datasets in hybrid formats including graph databases, RDF collections, and relational databases. These induce recursive queries on linked data collections for which we will investigate evaluation algorithms, containment problems, and concrete applications.

Managing Dynamic Linked Data In order to manage dynamic linked data collections and workflows, we will develop distributed data-centric programming languages with streams and parallelism, based on novel algorithms for incremental query answering, study the propagation of updates of dynamic data through schema mappings, and investigate static analysis methods for linked data workflows.

Linking Data Graphs Finally, we will develop symbolic machine learning algorithms, for inferring queries and mappings between linked data collections in various graphs formats from annotated examples.

3. Research Program

3.1. Background

The main objective of LINKS is to develop methods for querying and managing linked data collections. Even though open linked data is the most prominent example, we will focus on hybrid linked data collections, which are collections of semi-structured datasets in hybrid formats: graph-based, RDF, relational, and NOSQL. The elements of these datasets may be linked, either by pointers or by additional relations between the elements of the different datasets, for instance the "same-as" or "member-of" relations as in RDF.

The advantage of traditional data models is that there exist powerful querying methods and technologies that one might want to preserve. In particular, they come with powerful schemas that constraint the possible manners in which knowledge is represented to a finite number of patterns. The exhaustiveness of these patterns is essential for writing of queries that cover all possible cases. Pattern violations are excluded by schema validation. In contrast, RDF schema languages such as RDFS can only enrich the relations of a dataset by new relations, which also helps for query writing, but which cannot constraint the number of possible patterns, so that they do not come with any reasonable notion of schema validation.

The main weakness of traditional formats, however, is that they do not scale to large data collections as stored on the Web, while the RDF data models scales well to very big collections such as linked open data. Therefore, our objective is to study mixed data collections, some of which may be in RDF format, in which we can lift the advantages of smaller datasets in traditional formats to much larger linked data collections. Such data collections are typically distributed over the internet, where data sources may have rigid query facilities that cannot be easily adapted or extended.

The main assumption that we impose in order to enable the logical approach, is that the given linked data collection must be correct in most dimensions. This means that all datasets are well-formed with respect to their available constraints and schemas, and clean with respect to the data values in most of the components of the relations in the datasets. One of the challenges is to integrate good quality RDF datasets into this setting, another is to clean the incorrect data in those dimensions that are less proper. It remains to be investigated in how far these assumptions can be maintained in realistic applications, and how much they can be weakened otherwise.

For querying linked data collections, the main problems are to resolve the heterogeneity of data formats and schemas, to understand the efficiency and expressiveness of recursive queries, that can follow links repeatedly, to answer queries under constraints, and to optimize query answering algorithms based on static analysis. When linked data is dynamically created, exchanged, or updated, the problems are how to process linked data incrementally, and how to manage linked data collections that change dynamically. In any case (static and dynamic) one needs to find appropriate schema mappings for linking semi-structured datasets. We will study how to automatize parts of this search process by developing symbolic machine learning techniques for linked data collections.

3.2. Querying Heterogeneous Linked Data

Our main objective is to query collections of linked datasets. In the static setting, we consider two kinds of links: explicit links between elements of the datasets, such as equalities or pointers, and logical links between relations of different datasets such as schema mappings. In the dynamic setting, we permit a third kind of links that point to "intentional" relations computable from a description, such as the application of a Web service or the application of a schema mapping.

We believe that collections of linked datasets are usually too big to ensure a global knowledge of all datasets. Therefore, schema mappings and constraints should remain between pairs of datasets. Our main goal is to be able to pose a query on a collection of datasets, while accounting for the possible recursive effects of schema mappings. For illustration, consider a ring of datasets D_1, D_2, D_3 linked by schema mappings M_1, M_2, M_3 that tell us how to complete a database D_i by new elements from the next database in the cycle.

The mappings M_i induce three intentional datasets I_1, I_2, and I_3, such that I_i contains all elements from D_i and all elements implied by M_i from the next intentional dataset in the ring:

$$I_1 = D_1 \cup M_1(I_2), \quad I_2 = D_2 \cup M_2(I_3), \quad I_3 = D_3 \cup M_3(I_1)$$

Clearly, the global information collected by the intentional datasets depends recursively on all three original datasets D_i. Queries to the global information can now be specified as standard queries to the intentional databases I_i. However, we will never materialize the intentional databases I_i. Instead, we can rewrite queries on one of the intentional datasets I_i to recursive queries on the union of the original datasets D_1, D_2, and D_3 with their links and relations. Therefore, a query answering algorithm is needed for recursive queries, that chases the "links" between the D_i in order to compute the part of I_i needed for the purpose of query answering.

This illustrates that we must account for the graph data models when dealing with linked data collections whose elements are linked, and that query languages for such graphs must provide recursion in order to chase links. Therefore, we will have to study graph databases with recursive queries, such as RDF graphs with SPARQL queries, but also other classes of graph databases and queries.

We study schemas and mappings between datasets with different kinds of data models and the complexity of evaluating recursive queries over graphs. In order to use schema mapping for efficiently querying the different datasets, we need to optimize the queries by taking into account the mappings. Therefore, we will study static analysis of schema mappings and recursive queries. Finally, we develop concrete applications in which our fundamental techniques can be applied.

3.3. Managing Dynamic Linked Data

With the quick growth of the information technology on the Web, more and more Web data gets created dynamically every day, for instance by smartphones, industrial machines, users of social networks, and all kinds of sensors. Therefore, large amounts of dynamic data need to be exchanged and managed by various data-centric web services, such as online shops, online newspapers, and social networks.

Dynamic data is often created by the application of some kind of service on the Web. This kind of data is intentional in the same spirit as the intentional data specified by the application of a schema mapping, or the application of some query to the hidden Web. Therefore, we will consider a third kind of links in the dynamic setting, that map to intentional data specified by whatever kind of function application. Such a function can be defined in data-centric programming languages, in the style of Active XML, XSLT, and NOSQL languages.

The dynamicity of data adds a further dimension to the challenges for linked data collections that we described before, while all the difficulties remain valid. One of the new aspects is that intentional data may be produced incrementally, as for instance when exchanged over data streams. Therefore, one needs incremental algorithms able to evaluate queries on incomplete linked data collections, that are extended or updated incrementally. Note that incremental data may be produced without end, such as a Twitter stream, so that one cannot wait for its completion. Instead, one needs to query and manage dynamic data with as low latency as possible. Furthermore, all static analysis problems are to be re-investigated in the presence of dynamic data.

Another aspect of dynamic data is distribution over the Web, and thus parallel processing as in the cloud. This raises the typical problems coming with data distribution: huge data sources cannot be moved without very high costs, while data must be replicated for providing efficient parallel access. This makes it difficult, if not impossible, to update replicated data consistently. Therefore, the consistency assumption has been removed by NOSQL databases for instance, while parallel algorithmic is limited to naive parallelization (i.e. map/reduce) where only few data needs to be exchanged.

We will investigate incremental query evaluation for distributed data-centered programming languages for linked data collections, dynamic updates as needed for linked data management, and static analysis for linked data workflows.

3.4. Linking Graphs

When datasets from independent sources are not linked with existing schema mappings, we would like to investigate symbolic machine learning solutions for inferring such mappings in order to define meaningful links between data from separate sources. This problem can be studied for various kinds of linked data collections. Before presenting the precise objectives, we will illustrate our approach on the example of linking data in two independent graphs: an address book of a research institute containing detailed personnel information and a (global) bibliographic database containing information on papers and their authors.

We remind that a schema allows to identify a collection of types each grouping objects from the same semantic class e.g., the collection of all persons in the address book and the collection of all authors in the bibliography database. As a schema is often lacking or underspecified in graph data models, we intend to investigate inference methods based on structural similarity of graph fragments used to describe objects from the same class in a given document e.g., in the bibliographic database every author has a name and a number of affiliations, while a paper has a title and a number of authors. Furthermore, our inference methods will attempt to identify, for every type, a set of possible keys, where by key we understand a collection of attributes of an object that uniquely identifies such an object in its semantic class. For instance, for a person in the address book two examples of a key are the name of the person and the office phone number of that person.

In the next step, we plan to investigate employing existing entity linkage solutions to identify pairs of types from different databases whose instances should be linked using compatible keys. For instance, persons in the address book should be linked with authors in the bibliographical database using the name as the compatible key. Linking the same objects (represented in different ways) in two databases can be viewed as an instance of a mapping between the two databases. Such mapping is, however, discriminatory because it typically maps objects from a specific subset of objects of given types. For instance, the mapping implied by linking persons in the address book with authors in the bibliographic database involves in fact researchers, a subgroup of personnel of the research institute, and authors affiliated with the research institute. Naturally, a subset of objects of a given type, or a subtype, can be viewed as a result of a query on the set of all objects, which on very basic level illustrates how learning data mappings can be reduced to learning queries.

While basic mappings link objects of the same type, more general mappings define how the same type of information is represented in two different databases. For instance, the email address and the postal address of an individual may be represented in one way in the address book and in another way in the bibliographic databases, and naturally, the query asking for the email address and the postal address of a person identified by a given name will differ from one database to the other. While queries used in the context of linking objects of compatible types are essentially unary, queries used in the context of linking information are n-ary and we plan to approach inference of general database mappings by investigating and employing algorithms for inference of n-ary queries.

An important goal in this research is elaborating a formal definition of *learnability* (feasibility of inference) of a given class of concepts (schemas of queries). We plan to following the example of Gold (1967), which requires not only the existence of an efficient algorithm that infers concepts consistent with the given input but the ability to infer every concept from the given class with a sufficiently informative input. Naturally, learnability depends on two parameters. The first parameter is the class of concepts i.e., a class of schema and a class of queries, from which the goal concept is to be inferred. The second parameter is the type of input that an inference algorithm is given. This can be a set of examples of a concept e.g., instances of RDF databases for which we wish to construct a schema or a selection of nodes that a goal query is to select. Alternatively, a more general interactive scenario can be used where the learning algorithm inquires the user about the goal concept e.g., by asking to indicate whether a given node is to be selected or not (as membership queries of Angluin (1987)). In general, the richer the input is, the richer class of concepts can be handled, however, the richer class of queries is to be handled, the higher computational cost is to be expected. The primary task is to find a good compromise and identify classes of concepts that are of high practical value, allow efficient inference with possibly simple type of input.

The main open problem for graph-shaped data studied by Links are how to infer queries, schemas, and schema-mappings for graph-structured data.

4. Application Domains

4.1. Linked Data Integration

There are many contexts in which integrating linked data is interesting. We advocate here one possible scenario, namely that of integrating business linked data to feed what is called Business Intelligence. The latter consists of a set of theories and methodologies that transform raw data into meaningful and useful information for business purposes (from Wikipedia). In the past decade, most of the enterprise data was proprietary, thus residing within the enterprise repository, along with the knowledge derived from that data. Today's enterprises and businessmen need to face the problem of information explosion, due to the Internet's ability to rapidly convey large amounts of information throughout the world via end-user applications and tools. Although linked data collections exist by bridging the gap between enterprise data and external resources, they are not sufficient to support the various tasks of Business Intelligence. To make a concrete example, concepts in an enterprise repository need to be matched with concepts in Wikipedia and this can be done via pointers or equalities. However, more complex logical statements (i.e. mappings) need to be conceived to map a portion of a local database to a portion of an RDF graph, such as a subgraph in Wikipedia or in a social network, e.g. LinkedIn. Such mappings would then enrich the amount of knowledge shared within the enterprise and let more complex queries be evaluated. As an example, businessmen with the aid of business intelligence tools need to make complex sentimental analysis on the potential clients and for such a reason, such tools must be able to pose complex queries, that exploit the previous logical mappings to guide their analysis. Moreover, the external resources may be rapidly evolving thus leading to revisit the current state of business intelligence within the enterprise.

4.2. Data Cleaning

The second example of application of our proposal concerns scientists who want to quickly inspect relevant literature and datasets. In such a case, local knowledge that comes from a local repository of publications belonging to a research institute (e.g. HAL) need to be integrated with other Web-based repositories, such as DBLP, Google Scholar, ResearchGate and even Wikipedia. Indeed, the local repository may be incomplete or contain semantic ambiguities, such as mistaken or missing conference venues, mistaken long names for the publication venues and journals, missing explanation of research keywords, and opaque keywords. We envision a publication management system that exploits both links between database elements, namely pointers to external resources and logical links. The latter can be complex relationships between local portions of data and remote resources, encoded as schema mappings. There are different tasks that such a scenario could entail such as (i) cleaning the errors with links to correct data e.g. via mappings from HAL to DBLP for the publications errors, and via mappings from HAL to Wikipedia for opaque keywords, (ii) thoroughly enrich the list of publications of a given research institute, and (iii) support complex queries on the corrected data combined with logical mappings.

4.3. Real Time Complex Event Processing

Complex event processing serves for monitoring nested word streams in real time. Complex event streams are gaining popularity with social networks such as with Facebook and Twitter, and thus should be supported by distributed databases on the Web. Since this is not yet the case, there remains much space for future industrial transfer related to Links' second axis on dynamic linked data.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Data Integration and Schema Validation

The ShEx language for defining RDF schemas was proposed and developed earlier by the Links team in cooperation with the W3C. S. Staworko et al. now studied the containment problem for ShEx schemas for RDF documents. They showed at *PODS* [10] – the best database theory conference – that the problem is decidable, but co-NEXP-hard. This is a joint work with P. Wieczorek from the University of Wroclaw, Poland.

5.1.2. Aggregates

Florent Capelli et al. showed at *STACS* [7] – a top conferences in theoretical computer science – a new knowledge compilation procedure for quantified Boolean formulas allowing to decide the satisfiability quantified Boolean formulas with bounded tree width in polynomial time. This can be applied in particular to first-order database queries with quantifiers. This is joined work with S. Mengel from the CNRS in Lens.

6. New Software and Platforms

6.1. ShEx validator

Validation of Shape Expression schemas

KEYWORDS: Data management - RDF

FUNCTIONAL DESCRIPTION: Shape Expression schemas is a formalism for defining constraints on RDF graphs. This software allows to check whether a graph satisfies a Shape Expressions schema.

RELEASE FUNCTIONAL DESCRIPTION: ShExJava now uses the Commons RDF API and so support RDF4J, Jena, JSON-LD-Java, OWL API and Apache Clerezza. It can parse ShEx schema in the ShEcC, ShEJ, ShExR formats and can serialize a schema in ShExJ.

To validate data against a ShExSchema using ShExJava, you have two different algorithms: - the refine algorithm: compute once and for all the typing for the whole graph - the recursive algorithm: compute only the typing required to answer a validate(node,ShapeLabel) call and forget the results.

- Contact: Iovka Boneva
- URL: http://shexjava.lille.inria.fr/

6.2. gMark

gMark: schema-driven graph and query generation

KEYWORDS: Semantic Web - Data base

FUNCTIONAL DESCRIPTION: gMark allow the generation of graph databases and an associated set of query from a schema of the graph.gMark is based on the following principles: - great flexibility in the schema definition - ability to generate big size graphs - ability to generate recursive queries - ability to generate queries with a desired selectivity

- Contact: Aurélien Lemay
- URL: https://github.com/graphMark/gmark

6.3. SmartHal

KEYWORD: Bibliography

FUNCTIONAL DESCRIPTION: SmartHal is a better tool for querying the HAL bibliography database, while is based on Haltool queries. The idea is that a Haltool query returns an XML document that can be queried further. In order to do so, SmartHal provides a new query language. Its queries are conjunctions of Haltool queries (for a list of laboratories or authors) with expressive Boolean queries by which answers of Haltool queries can be refined. These Boolean refinement queries are automatically translated to XQuery and executed by Saxon. A java application for extraction from the command line is available. On top of this, we have build a tool for producing the citation lists for the evaluation report of the LIFL, which can be easily adapter to other Labs.

- Contact: Joachim Niehren
- URL: http://smarthal.lille.inria.fr/

6.4. QuiXPath

KEYWORDS: XML - NoSQL - Data stream

SCIENTIFIC DESCRIPTION: The QuiXPath tools supports a very large fragment of XPath 3.0. The QuiXPath library provides a compiler from QuiXPath to FXP, which is a library for querying XML streams with a fragment of temporal logic.

FUNCTIONAL DESCRIPTION: QuiXPath is a streaming implementation of XPath 3.0. It can query large XML files without loading the entire file in main memory, while selecting nodes as early as possible.

- Contact: Joachim Niehren
- URL: https://project.inria.fr/quix-tool-suite/

6.5. X-FUN

KEYWORDS: Programming language - Compilers - Functional programming - Transformation - XML

FUNCTIONAL DESCRIPTION: X-FUN is a core language for implementing various XML, standards in a uniform manner. X-Fun is a higher-order functional programming language for transforming data trees based on node selection queries.

- Participants: Joachim Niehren and Pavel Labath
- Contact: Joachim Niehren

6.6. ShapeDesigner

KEYWORDS: Validation - Data Exploration - Verification

FUNCTIONAL DESCRIPTION: ShapeDesigner allows construct a ShEx or SHACL schema for an existing dataset. It combines algorithms to analyse the data and automatically extract shape constraints, and to edit and validate shape schemas.

- Contact: Jeremie Dusart
- URL: https://gitlab.inria.fr/jdusart/shexjapp

7. New Results

7.1. Querying Heterogeneous Linked Data

7.1.1. Data Integration and Schema Validation

Data integration requires knowledge about the structure of the various data. Such a structure is usually described by schemas. While for relational databases, schemas are hard-coded, this is not the case for many other formats. In XML for instance, several schema formalisms exists, such as DTD, XML Schema or Schematron. The Links Project-Team investigate the problem of defining schemas and use them to data, in particular for RDF and JSON Formats.

With P. Wieczorek of the University of Wroclaw, Poland, S. Staworko et al. have studied the containment problem of ShEx schemas for RDF documents in *PODS* [10].

Also, J. Dusart develops under the supervision of I. Boneva and S. Staworko the software *ShEx Validator* so as to foster the practical usage of ShEx. It is also worth noting that ShEx is now being adopted by several institutions such as *WikiData*.

7.1.2. Aggregates

Aggregation refers to computations that are alien to mere logical data manipulation (e.g. such as in relational algebra). Typically, aggregation means counting the number of answers, or performing other kinds of statistics. We have a slightly larger understanding as we may also include enumerating all answers with a *small delay*. Aggregation algorithms are generally subtle as they in most cases avoid the explicit generation of the whole set of answers. We study aggregation problems within the ANR project *Aggreg* coordinated by Niehren.

In the same spirit, Capelli et al. (in a joint work with Mengel from the CNRS in Lens) showed at *STACS* [7] a new knowledge compilation procedure which allows a polynomial algorithm to test the satisfiability quantified Boolean formulas with bounded tree width. In *Theory of Computing Systems*, [25], Capelli also gave a taxonomy of results according to various restrictions of tree-width of graphs.

Also, in *Theory of Computing Systems*, [25], Capelli gave a taxonomy of results according to various restrictions of tree-width of graphs.

Finally, in an article in *JCSS* [14], F. Capelli (with Bergougnoux and Kanté from Bordeaux and Clermont-Ferrand) propose an algorithm for counting the number of transversals (i.e. subset of nodes intersecting all hyperedges) in some hypergraphs.

7.1.3. Certain Query Answering

When data is incomplete, logical constraints and knowledge about its intended structure help to infer the answers of queries. This inference problem is known as *certain query answering*.

L. Gallois and S. Tison [6] presented in *IJCAI* - one of the main conferences of Artificial Intelligence. L. Gallois and S. Tison study boundedness of the chase procedure in the context of positive existential rules, providing decidability results for several classes and outlining the complexity of the problem. This work is done in collaboration with P. Bourhis and Graphik team-project. These results also belong to the PhD thesis of L. Gallois [11] supervised by S. Tison and P. Bourhis.

7.2. Managing Dynamic Linked Data

7.2.1. Complex Event Processing

Complex event processing requires to answer queries on streams of complex events, i.e., nested words or equivalently linearizations of data trees, but also to produce dynamically evolving data structures as output.

In an article published in *LATA* [17], I. Boneva, J. Niehren and M. Sakho studied certain query answering for hyperstreams - which are collections of connected streams - with *complex events* (i.e. that correspond to tree patterns). They showed that the problem is `EXP`-complete in general, and obtained PTIME algorithms when restricted to *linear* tree patterns (possibly with compression) and to deterministic tree automata.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

Strapdata C. Paperman is actively collaborating with the Strapdata company on efficient distributed graph database using an Apache novel technology to query distributed graph *Gremlin* that could benefit of the main product of Strapdata: Elassandra as a *database backend*.

9. Partnerships and Cooperations

9.1. Regional Initiatives

- Links is member of the CPER Data (2016-19)
- Lozano's PhD project (2016-19) is co-funded by the Region Nord-Pas de Calais
- Sakho's PhD project is co-funded by the Region Nord-Pas de Calais
- Gallot's PhD project (2017-20) is co-funded by the Region Nord-Pas de Calais
- Crosetti's PhD project (2018-21) is co-funded by the Region Haut de France. This is joined work with J. Ramon from the Inria project Magnet

9.2. National Initiatives

ANR Aggreg (2014-19): Aggregated Queries.
Participants: Joachim Niehren [correspondent], Aurélien Lemay, Adrien Boiret [University of Mons, Belgium], Florent Capelli.

- The coordinator is J. Niehren and the partners are the Université Paris 7 (A. Durand) including members of the Inria project DAHU (L. Ségoufin), the Université de Marseille (N. Creignou) and Université de Caen (E. Grandjean).
- Objective: the main goal of the Aggreg project is to develop efficient algorithms and to study the complexity of answering aggregate queries for databases and data streams of various kinds.

ANR Colis (2015-20): Correctness of Linux Scripts.
Participants: Joachim Niehren [correspondent], Aurélien Lemay, Sophie Tison, Adrien Boiret [University of Mons, Belgium], Vincent Hugot [INSA Centre-Val de Loire], Nicolas Bacquey [Twig], Paul Gallot, Sylvain Salvati.

- The coordinator is R. Treinen from the Université Paris 7 and the other partner is the Tocata project of Inria Saclay (C. Marché).
- Objective: This project aims at verifying the correctness of transformations on data trees defined by shell scripts for Linux software installation. The data trees here are the instance of the file system which are changed by installation scripts.

ANR DataCert (2015-20):
Participants: Iovka Boneva [correspondent], Sophie Tison, Jose Martin Lozano.

- Partners: The coordinator is E. Contejean from the Université Paris-Sud and the other partner is the Université de Lyon.
- Objective: the main goals of the Datacert project are to provide deep specification in Coq of algorithms for data integration and exchange and of algorithms for enforcing security policies, as well as to design data integration methods for data models beyond the relational data model.

ANR Headwork (2016-21):
Participants: Joachim Niehren [correspondent], Momar Sakho, Nicolas Crosetti, Florent Capelli.

- Scientific partners: The coordinateur is D. Gross-Amblard from the Druid Team (Rennes 1). Other partners include the Dahu team (Inria Saclay) and Sumo (Inria Bretagne).
- Industrial partners: Spipoll, and Foulefactory.
- Objective: The main object is to develop data-centric workflows for programming crowd sourcing systems in flexible declarative manner. The problem of crowd sourcing systems is to fill a database with knowledge gathered by thousands or more human participants. A particular focus is to be put on the aspects of data uncertainty and for the representation of user expertise.

ANR Delta (2016-21):
Participants: Joachim Niehren [correspondent], Sylvain Salvati, Aurélien Lemay, Nicolas Bacquey [Twig], Lily Gallois.

- Partners: The coordinator is M. Zeitoun from LaBRI, other partners are LIF (Marseille) and IRIF (Paris-Diderot).
- Objective: Delta is focused on the study of logic, transducers and automata. In particular, it aims at extending classical framework to handle input/output, quantities and data.

ANR Bravas (2017-22):
Participant: Sylvain Salvati [correspondent].

- Scientific Partners: The coordinator is Jérôme Leroux from LaBRI, Université de Bordeaux. The other partner is LSV, ENS Cachan.
- Objective: The goal of the BraVAS project is to develop a new and powerful approach to decide the reachability problems for Vector Addition Systems (VAS) extensions and to analyze their complexity. The ambition here is to crack with a single hammer (ideals over well-orders) several long-lasting open problems that have all been identified as a barrier in different areas, but that are in fact closely related when seen as reachability.

9.3. European Initiatives

Oxford, UK: An exchange project with the computer science lab of the University of Oxford is funded by the Université de Lille via the CRIStAL Lab. Links' members produced many common publications over the years with Oxford. Links' contact is C. Paperman.

Wroclaw, Poland: S. Staworko has regular exchange with the University of Wroclaw. This has led to a publication at *PODS* [10] together with P. Wieczorek.

Saint-Petersburg, Russia: S. Salvati and J. Niehren started a cooperation with the Saint-Petersburg State University, via a month-long visit by R. Azimov and S. Grigorev.

Oviedo, Spain: I. Boneva has an active cooperation with the University of Oviedo.

9.4. International Initiatives

9.4.1. Informal International Partners

Santiago de Chile, Chile: S. Staworko and I. Boneva have a collaboration with C. Riveros from the Pontifical Catholic University of Chile since 2018.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Selection

10.1.1.1. Chair of Conference Program Committees

J. Niehren was co-chair of the programm committee of WPTE 2019.

10.1.1.2. Member of the Conference Program Committees

- J. Niehren: member of the Program Committee of LATA 2019
- S. Tison: member of the program committee of FSCD (International Conference on Formal Structure for Computation and Deduction) 2019.
- S. Staworko: member of the program committee of EDBT 2020: International Conference on Extending Database Technology, Copenhagen, Denmark, March 30–April 2, 2020; Demonstration Track
- F. Capelli: member of the program committee of IJCAI 2019

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

J. Niehren is editor of Fundamenta Informaticae

S. Salvati is managing editor of JoLLI (Journal for Logic, Language and Information)

S. Tison is in the editorial committee of RAIRO-ITA (Theoretical Informatics and Applications)

10.1.3. Invited Talks

- S. Staworko has been invited in GT Automata, Logic, Games and Algebra (ALGA) to give the talk *Shape Expressions Schemas for RDF: Semantics, Complexity, and Inference* (Oct 2019, Paris)
- I. Boneva and J. Dusart have been invited to Ghent University, Belgium Wikidata and Wikibase Workshop: developing a Wikibase instance (3 - 5 July 2019) as expert on ShapeDesigner
- F. Capelli has been invited in the seminar of LIMD (Université de Savoie) (13/06/2019)
- F. Capelli has been invited in the Dagstuhl Seminar “Deduction Beyond Satisfiability” (10/09/2019)
- F. Capelli has been invited in the seminar of LACL (Université Paris-Est Creteil) (18/11/2019)
- F. Capelli has been invited in “Journée du GT ALGA” (Paris, 11/10/201)9

10.1.4. Scientific Expertise

- S. Tison: member of the coordinating committee of I-Site Université Lille Nord Europe, about innovation and relationship with social economical world
- S. Tison: head of CITC-Eurarfid until June 2019
- J. Niehren: member of the board of the comittee of project-teams of Inria Lille

10.1.5. Research Administration

- S. Salvati: elected alternate member of Inria Evaluation Commission (2019-2023)
- F. Capelli: co-organizer of *Groupe de Travail* of CNRS IMIA (Informatique Mathématique Intelligence Artificielle)
- F. Capelli: organiser of a AI DAY for the GDR IM
- F. Capelli: organiser of an international workshop on knowledge compilation (project SACRe Kocoon - http://kocoon.gforge.inria.fr/)

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

I. Boneva teaches computer science in DUT Informatique of Université de Lille

F. Capelli teaches computer science in UFR LEA of Université de Lille for around 200h per year (Licence and Master). He is also responsible of remediation of Licence 1 in its UFR.

A. Lemay teaches computer science in UFR LEA of Université de Lille for around 200h per year (Licence and Master). He is also responsible for computer science and numeric correspondent for its UFR.

J. Niehren gives two lessons for the 2nd year students of the Master MOCAD (Université de Lille): one on databases (20.5h) and one on information extraction (21h).

C. Paperman teaches computer science for a total of around 200h per year. He gives lessons in UFR MISASH (Université de Lille), in Licence and Master. He also gives a database lesson of 25h in Master MOCAD (Université de Lille).

S. Salvati teaches computer science for a total of around 230h per year in computer science departement of Université de Lille. That includes Introduction to Computer Science (L1, 50h), Logic (L3, 50h), Algorithmic and operational research (L3, 36h), Functional Programming (L3, 35h), Research Option (L3, 10h), Semantic Web (M2, 30h), Advanced Databases (M1, 20h). He is *directeur d'étude* of Master MIAGE FA. He is a member of conseil de departement in Computer Science department of Université de Lille and of the ad-hoc commission of Doctoral School that studies PhD applications in computer science.

S. Staworko teaches computer science for a total of around 200h in UFR MIME (Université de Lille). He is co-head of the master web-analyst at Université de Lille.

S. Tison teaches computer science for a total of around 120h at Université de Lille. That includes a course on Advanced algorithms and complexity (54h, M1) and Business Intelligence (36h, M1).

S. Tison is member of the selection Board for "Agrégation" in Mathematics, more specifically in charge of the option "Computer Science".

10.2.2. Supervision

PhD: L. Gallois, Recursive Queries, defended on December 19, 2019, supervised by P. Bourhis (Team SPIRALS) and S. Tison

PhD in progress: N. Crosetti, Privacy Risks of Aggregates in Data Centric-Workflows, supervised by F. Capelli, J. Niehren, J. Ramon (Team MAGNET) and S. Tison

PhD in progress: P. Gallot, On safety of data transformations, since October 2017, supervised by A. Lemay and S. Salvati

PhD in progress: J.M. Lozano, On data integration for mixed database formats, supervised by I. Boneva and S. Staworko

PhD in progress: M. Sakho, Hyperstreaming Query answering on graphs, since 2016, supervised by J. Niehren and I. Boneva

10.2.3. Juries

10.2.3.1. PhDs committees

- J. Niehren was a reviewer of the PhD thesis of *Antony Lick* at ENS Cachan 2019
- S. Tison was member of the jury for *Xinzhe Wu*, Université de Lille, March 2019
- S. Tison was member of the jury for *Stathis Delivorias*, Université de Montpellier, September 2019 (reviewer)
- S. Tison was member of the jury for *Nicolas Bloyet*, Université Bretagne Sud, December 2019 (reviewer)

10.2.3.2. HDR committees

- S. Tison was member of the jury for *Arnaud Carayol*, December 2019, Université Paris Est-Marne la Vallée (reviewer)
- S. Salvati was member of the jury for *Colin Riba*, ENS Lyon, December 2019

10.3. Popularization

10.3.1. Education

RIC Days S. Salvati organizes the *Recherche Innovation et Créativité days*. It presents research professions to student (primarily Master students)

10.3.2. Interventions

Introduction to programming I. Boneva has supervised second year students from DUT while they conducted activities on introduction to programming to 9-10 years old students in Villeneuve d'Ascq

TFJM^2 In 2019, L. Gallois participated in the organization of this event of the *Tournoi Français des jeunes mathématiciennes et mathématiciens*, a national contest of mathematics for high-school students

10.3.3. Internal action

- Inria by Lille: J. Niehren contributed an article in the december issue of "Inria by Lille" titled "Interroger les bases de données d'une manière plus intelligente"

11. Bibliography

Major publications by the team in recent years

[1] A. AMARILLI, C. PAPERMAN.*Topological Sorting with Regular Constraints*, in "45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)", Prague, Czech Republic, July 2018, https://hal.archives-ouvertes.fr/hal-01950909

[2] M. BENEDIKT, P. BOURHIS, M. V. BOOM.*Characterizing Definability in Decidable Fixpoint Logics* , in "ICALP 2017 - 44th International Colloquium on Automata, Languages, and Programming", Varsovie, Poland, I. CHATZIGIANNAKIS, P. INDYK, F. KUHN, A. MUSCHOLL (editors), July 2017, vol. 107, 14, ICALP-2017 Best paper award of Track B [*DOI :* 10.4230/LIPICS.ICALP.2017.107], https://hal.inria.fr/hal-01639015

[3] A. BOIRET, V. HUGOT, J. NIEHREN, R. TREINEN.*Logics for Unordered Trees with Data Constraints*, in "Journal of Computer and System Sciences", December 2018, 40, https://hal.inria.fr/hal-01176763

[4] I. BONEVA, J. G. LABRA GAYO, E. G. PRUD 'HOMMEAUX.*Semantics and Validation of Shapes Schemas for RDF*, in "ISWC2017 - 16th International semantic web conference", Vienna, Austria, October 2017, https://hal.archives-ouvertes.fr/hal-01590350

[5] A. BONIFATI, R. CIUCANU, S. STAWORKO.*Learning Join Queries from User Examples*, in "ACM Transactions on Database Systems", February 2016, vol. 40, n^{o} 4, p. 1-38, https://hal.inria.fr/hal-01187986

[6] P. BOURHIS, M. LECLÈRE, M.-L. MUGNIER, S. TISON, F. ULLIANA, L. GALLOIS.*Oblivious and Semi-Oblivious Boundedness for Existential Rules*, in "IJCAI 2019 - International Joint Conference on Artificial Intelligence", Macao, China, August 2019, https://hal-lirmm.ccsd.cnrs.fr/lirmm-02148142

[7] F. CAPELLI, S. MENGEL.*Tractable QBF by Knowledge Compilation*, in "36th International Symposium on Theoretical Aspects of Computer Science (STACS 2019)", Berlin, Germany, March 2019, https://arxiv.org/abs/1807.04263 , https://hal.archives-ouvertes.fr/hal-01836402

[8] D. DEBARBIEUX, O. GAUWIN, J. NIEHREN, T. SEBASTIAN, M. ZERGAOUI.*Early Nested Word Automata for XPath Query Answering on XML Streams*, in "Theoretical Computer Science", March 2015, n° 578, p. 100-127, https://hal.inria.fr/hal-00966625

[9] V. HUGOT, A. BOIRET, J. NIEHREN.*Equivalence of Symbolic Tree Transducers*, in "DLT 2017 - Developments in Language Theory", Liege, Belgium, August 2017, vol. 105, 12 [*DOI :* 10.1007/978-3-642-29709-0_32], https://hal.inria.fr/hal-01517919

[10] S. STAWORKO, P. WIECZOREK.*Containment of Shape Expression Schemas for RDF*, in "SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems (PODS)", Amsterdam, Netherlands, June 2019, https://hal.inria.fr/hal-01959143

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] L. GALLOIS.*Dialog between chase approach and string rewriting system approach*, Université de Lille, December 2019, https://tel.archives-ouvertes.fr/tel-02445754

Articles in International Peer-Reviewed Journal

[12] A. AMARILLI, F. CAPELLI, M. MONET, P. SENELLART.*Connecting Knowledge Compilation Classes and Width Parameters*, in "Theory of Computing Systems", June 2019, https://arxiv.org/abs/1811.02944 [*DOI :* 10.1007/S00224-019-09930-2], https://hal.inria.fr/hal-02163749

[13] M. BENEDIKT, P. BOURHIS, M. VANDEN BOOM.*Definability and Interpolation within Decidable Fixpoint Logics*, in "Logical Methods in Computer Science", 2019, https://hal.archives-ouvertes.fr/hal-02392433

[14] B. BERGOUGNOUX, F. CAPELLI, M. M. KANTÉ.*Counting Minimal Transversals of β-Acyclic Hypergraphs*, in "Journal of Computer and System Sciences", May 2019, https://arxiv.org/abs/1808.05017 [*DOI :* 10.1016/J.JCSS.2018.10.002], https://hal.inria.fr/hal-01923090

[15] A. BOIRET, V. HUGOT, J. NIEHREN, R. TREINEN.*Logics for Unordered Trees with Data Constraints*, in "Journal of Computer and System Sciences", January 2019, 40 [*DOI :* 10.1016/J.JCSS.2018.11.004], https://hal.inria.fr/hal-01176763

International Conferences with Proceedings

[16] E. ALLART, C. VERSARI, J. NIEHREN.*Computing Difference Abstractions of Metabolic Networks Under Kinetic Constraints*, in "CMSB 2019 - 17th International Conference on Computational Methods in Systems Biology", Trieste, Italy, Lecture Notes in Computer Science, Springer, September 2019, vol. 11773, p. 266-285 [*DOI :* 10.1007/978-3-030-31304-3_14], https://hal.archives-ouvertes.fr/hal-02302463

[17] I. BONEVA, J. NIEHREN, M. SAKHO.*Regular Matching and Inclusion on Compressed Tree Patterns with Context Variables*, in "LATA 2019 - 13th International Conference on Language and Automata Theory and Applications", Saint Petersburg, Russia, January 2019, https://hal.inria.fr/hal-01811835

[18] P. BOURHIS, M. LECLÈRE, M.-L. MUGNIER, S. TISON, F. ULLIANA, L. GALLOIS.*Oblivious and Semi-Oblivious Boundedness for Existential Rules*, in "IJCAI 2019 - International Joint Conference on Artificial Intelligence", Macao, China, August 2019, https://hal-lirmm.ccsd.cnrs.fr/lirmm-02148142

[19] F. CAPELLI.*Knowledge Compilation Languages as Proof Systems*, in "Theory and Applications of Satisfiability Testing – SAT 2019", Lisbon, Portugal, July 2019, p. 90-99, https://arxiv.org/abs/1903.04039 [*DOI :* 10.1007/978-3-030-24258-9_6], https://hal.inria.fr/hal-02445523

[20] F. CAPELLI, S. MENGEL.*Tractable QBF by Knowledge Compilation*, in "36th International Symposium on Theoretical Aspects of Computer Science (STACS 2019)", Berlin, Germany, March 2019, https://arxiv.org/abs/1807.04263 , https://hal.archives-ouvertes.fr/hal-01836402

[21] S. STAWORKO, P. WIECZOREK.*Containment of Shape Expression Schemas for RDF*, in "PODS 2019 - 38th ACM SIGMOD-SIGACT-SIGAI Symposium on PRINCIPLES OF DATABASE SYSTEMS", Amsterdam, Netherlands, ACM Press, June 2019, p. 303-319 [*DOI :* 10.1145/3294052.3319687], https://hal.inria.fr/hal-01959143

Other Publications

[22] I. BONEVA, J. DUSART, D. FERNÁNDEZ ALVAREZ, J. E. L. GAYO.*Shape Designer for ShEx and SHACL Constraints*, October 2019, ISWC 2019 - 18th International Semantic Web Conference, Poster, https://hal.archives-ouvertes.fr/hal-02268667

[23] I. BONEVA, J. DUSART, D. FERNÁNDEZ ALVAREZ, J. E. LABRA GAYO.*Semi Automatic Construction of ShEx and SHACL Schemas*, July 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02193275

[24] I. BONEVA, J. NIEHREN, M. SAKHO.*Approximating Certain Query Answers on Nested Hyperstreams*, April 2019, working paper or preprint, https://hal.inria.fr/hal-02092276

[25] F. CAPELLI.*Knowledge compilation languages as proof systems*, June 2019, https://arxiv.org/abs/1903.04039 - working paper or preprint, https://hal.inria.fr/hal-02163761

[26] F. CAPELLI, N. CROSETTI, J. NIEHREN, J. RAMON.*Dependency Weighted Aggregation on Factorized Databases*, January 2019, https://arxiv.org/abs/1901.03633 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01981553

Project-Team LOKI

technology & knowledge for interaction

IN COLLABORATION WITH: Centre de Recherche en Informatique, Signal et Automatique de Lille

IN PARTNERSHIP WITH:
Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Interaction and visualization

Table of contents

Project-Team LOKI

Creation of the Project-Team: 2019 July 01

Keywords:

Computer Science and Digital Science:

A5.1. - Human-Computer Interaction
A5.1.1. - Engineering of interactive systems
A5.1.2. - Evaluation of interactive systems
A5.1.3. - Haptic interfaces
A5.1.4. - Brain-computer interfaces, physiological computing
A5.1.5. - Body-based interfaces
A5.1.8. - 3D User Interfaces
A5.1.9. - User and perceptual studies
A5.2. - Data visualization
A5.6.1. - Virtual reality
A5.6.3. - Avatar simulation and embodiment
A5.6.4. - Multisensory feedback and interfaces
A5.7.2. - Music

Other Research Topics and Application Domains:

B2.2.6. - Neurodegenerative diseases
B2.8. - Sports, performance, motor skills
B6.1.1. - Software engineering
B9.2.1. - Music, sound
B9.5.1. - Computer science
B9.5.6. - Data science
B9.6.10. - Digital humanities
B9.8. - Reproducibility

1. Team, Visitors, External Collaborators

Research Scientists

Stéphane Huot [Team leader, Inria, Senior Researcher, HDR]
Sylvain Malacria [Inria, Researcher]
Mathieu Nancel [Inria, Researcher]
Marcelo Wanderley [Inria, International Chair & Professor at McGill University]
Edward Lank [Inria, International Chair from Nov 2019 & Professor at University of Waterloo]

Faculty Members

Géry Casiez [Université de Lille, Professor, HDR]
Thomas Pietrzak [Université de Lille, Associate Professor]

PhD Students

Axel Antoine [Université de Lille]
Marc Baloup [Inria]
Nicole Pong [Inria]

Thibault Raffaillac [Université de Lille, until Sep 2019]
Grégoire Richard [Inria, from Oct 2019]
Philippe Schmid [Inria, from Oct 2019]

Post-Doctoral Fellow
Raiza Sarkis Hanada [Inria]

Visiting Scientist
Edward Lank [Professor at University of Waterloo, from Aug 2019 until Oct 2019]

Administrative Assistant
Julie Jonas [Inria]

2. Overall Objectives

2.1. Introduction

Human-Computer Interaction (HCI) is a constantly moving field [38]. Changes in computing technologies extend their possible uses and modify the conditions of existing ones. People also adapt to new technologies and adapt them to their own needs [42]. Different problems and opportunities thus regularly appear that require to be addressed from both the user and the machine perspective, to understand and account for the tight coupling between human factors and interactive technologies. Our vision is to link together these two essential elements: *Knowledge & Technology for Interaction*.

2.2. Knowledge for Interaction

In the early 1960s, when computers were scarce, expensive, bulky, and formal-scheduled machines used for automatic computations, ENGELBART saw their potential as personal interactive resources. He saw them as *tools* we would purposefully use to carry out particular tasks and that would empower people by supporting intelligent use [35]. Others at the same time were seeing computers differently, as *partners*, intelligent entities to whom we would delegate tasks. These two visions still constitute the roots of today's predominant HCI paradigms, *use* and *delegation*. In the delegation approach, a lot of effort has been made to support oral, written and non-verbal forms of human-computer communication, and to analyze and predict human behavior. But the inconsistency and ambiguity of human beings, and the variety and complexity of contexts, make these tasks very difficult [46] and the machine is thus the center of interest.

2.2.1. Computers as tools

Our focus is not in what machines can understand or do by themselves, but in what people can do with them. We do not reject the delegation paradigm but clearly favor the one of tool use, aiming for systems that support intelligent use rather than intelligent systems. And as the frontier is getting thinner, ***one of our goals is to better understand what it takes for an interactive system to be perceived as a tool or a partner, and how the two paradigms can be combined for the best benefit of the user.***

2.2.2. Empowering tools

The ability provided by interactive tools to create and control complex transformations in real-time can support intellectual and creative processes in unusual but powerful ways. But mastering powerful tools is not simple and immediate, it requires learning and practice. ***Our research in HCI should not just focus on novice or highly proficient users, it should also care about intermediate ones willing to devote time and effort to develop new skills, be it for work or leisure.***

2.2.3. Transparent tools

Technology is most empowering when it is transparent: invisible in effect, it does not get in your way but lets you focus on the task. HEIDEGGER characterized this unobtruded relation to things with the term *zuhanden* (*ready-to-hand*). Transparency of interaction is not best achieved with tools mimicking human capabilities, but with tools taking full advantage of them given the context and task. For instance, the transparency of driving a car "*is not achieved by having a car communicate like a person, but by providing the right coupling between the driver and action in the relevant domain (motion down the road)*" [49]. Our actions towards the digital world need to be digitized and we must receive proper feedback in return. But input and output technologies pose somewhat inevitable constraints while the number, diversity, and dynamicity of digital objects call for more and more sophisticated perception-action couplings for increasingly complex tasks. ***We want to study the means currently available for perception and action in the digital world: Do they leverage our perceptual and control skills? Do they support the right level of coupling for transparent use? Can we improve them or design more suitable ones?***

2.3. Technology for Interaction

Studying the *interactive phenomena* described above is one of the pillars of HCI research, in order to understand, model and ultimately improve them. Yet, we have to make those phenomena happen, to make them possible and reproducible, be it for further research or for their diffusion [37]. However, because of the high viscosity and the lack of openness of actual systems, this requires considerable efforts in designing, engineering, implementing and hacking hardware and software interactive artifacts. This is what we call "*The Iceberg of HCI Research*", of which the hidden part supports the design and study of new artifacts, but also informs their creation process.

2.3.1. "Designeering Interaction"

Both parts of this iceberg are strongly influencing each other: The design of interaction techniques (the visible top) informs on the capabilities and limitations of the platform and the software being used (the hidden bottom), giving insights into what could be done to improve them. On the other hand, new architectures and software tools open the way to new designs, by giving the necessary bricks to build with [39]. These bricks define the adjacent possible of interactive technology, the set of what could be designed by assembling the parts in new ways. Exploring ideas that lie outside of the adjacent possible require the necessary technological evolutions to be addressed first. This is a slow and gradual but uncertain process, which helps to explore and fill a number of gaps in our research field but can also lead to deadlocks. ***We want to better understand and master this process—i. e., analyzing the adjacent possible of HCI technology and methods—and introduce tools to support and extend it. This could help to make technology better suited to the exploration of the fundamentals of interaction, and to their integration into real systems, a way to ultimately improve interactive systems to be empowering tools.***

2.3.2. Computers vs Interactive Systems

In fact, today's interactive systems—e. g., desktop computers, mobile devices—share very similar layered architectures inherited from the first personal computers of the 1970s. This abstraction of resources provides developers with standard components (UI widgets) and high-level input events (mouse and keyboard) that obviously ease the development of common user interfaces for predictable and well-defined tasks and users' behaviors. But it does not favor the implementation of non-standard interaction techniques that could be better adapted to more particular contexts, to expressive and creative uses. Those often require to go deeper into the system layers, and to hack them until getting access to the required functionalities and/or data, which implies switching between programming paradigms and/or languages.

And these limitations are even more pervading as interactive systems have changed deeply in the last 20 years. They are no longer limited to a simple desktop or laptop computer with a display, a keyboard and a mouse. They are becoming more and more distributed and pervasive (e. g., mobile devices, Internet of Things). They are changing dynamically with recombinations of hardware and software (e. g., transition between multiple devices, modular interactive platforms for collaborative use). Systems are moving “out of the box” with Augmented Reality, and users are going “ inside of the box” with Virtual Reality. This is obviously raising new challenges in terms of human factors, usability and design, but it also deeply questions actual architectures.

2.3.3. *The Interaction Machine*

We believe that promoting digital devices to ***empowering tools*** requires ***better fundamental knowledge about interaction phenomena*** AND to ***revisit the architecture of interactive systems*** in order to support this knowledge. By following a comprehensive systems approach—encompassing human factors, hardware elements, and all software layers above—we want to define the founding principles of an *Interaction Machine*:

- a set of hardware and software requirements with associated specifications for interactive systems to be tailored to interaction by leveraging human skills;
- one or several implementations to demonstrate and validate the concept and the specifications in multiple contexts;
- guidelines and tools for designing and implementing interactive systems, based on these specifications and implementations.

To reach this goal, we will adopt an opportunistic and iterative strategy guided by the *designeering* approach, where the engineering aspect will be fueled by the interaction design and study aspect. We will address several fundamental problems of interaction related to our vision of “empowering tools”, which, in combination with state-of-the-art solutions, will instruct us on the requirements for the solutions to be supported in an interactive system. This consists in reifying the concept of the Interaction Machine into multiple contexts and for multiple problems, before converging towards a more unified definition of “what is an interactive system”, the ultimate Interaction Machine, which constitutes the main scientific and engineering challenge of our project.

3. Research Program

3.1. Introduction

Interaction is by nature a dynamic phenomenon that takes place between interactive systems and their users. Redesigning interactive systems to better account for interaction requires fine understanding of these dynamics from the user side so as to better handle them from the system side. In fact, layers of actual interactive systems abstract hardware and system resources from a system and programing perspective. Following our Interaction Machine concept, we are reconsidering these architectures from the user’s perspective, through different *levels of dynamics of interaction* (see Figure 1).

Considering phenomena that occur at each of these levels as well as their relationships will help us to acquire the necessary knowledge (Empowering Tools) and technological bricks (Interaction Machine) to reconcile the way interactive systems are designed and engineered with human abilities. Although our strategy is to investigate issues and address challenges for all of the three levels, our immediate priority is to focus on micro-dynamics since it concerns very fundamental knowledge about interaction and relates to very low-level parts of interactive systems, which is likely to influence our future research and developments at the other levels.

3.2. Micro-Dynamics

Micro-dynamics involve low-level phenomena and human abilities which are related to short time/instantness and to perception-action coupling in interaction, when the user has almost no control or consciousness of the action once it has been started. From a system perspective, it has implications mostly on input and output (I/O) management.

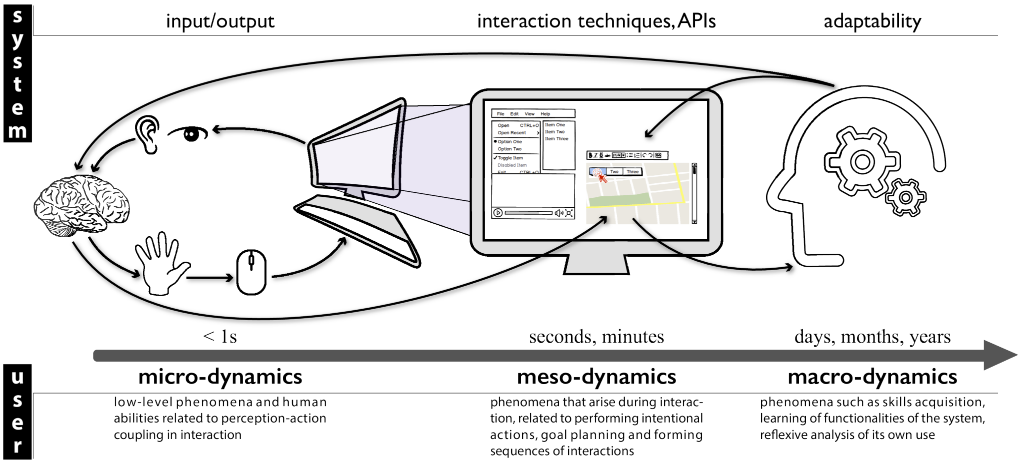

Figure 1. Levels of dynamics of interaction.

3.2.1. Transfer functions design and latency management

We have developed a recognized expertise in the characterization and the design of *transfer functions* [34], [45], i. e., the algorithmic transformations of raw user input for system use. Ideally, transfer functions should match the interaction context. Yet the question of how to maximize one or more criteria in a given context remains an open one, and on-demand adaptation is difficult because transfer functions are usually implemented at the lowest possible level to avoid latency. Latency has indeed long been known as a determinant of human performance in interactive systems [41] and recently regained attention with touch interactions [40]. These two problems require cross examination to improve performance with interactive systems: Latency can be a confounding factor when evaluating the effectiveness of transfer functions, and transfer functions can also include algorithms to compensate for latency.

We have recently proposed new cheap but robust methods for the measurement of end-to-end latency [2] and are currently working on compensation methods and the evaluation of their perceived side effects. Our goal is then to automatically adapt the transfer function to individual users and contexts of use while reducing latency in order to support stable and appropriate control. To achieve this, we will investigate combinations of low-level (embedded) and high-level (application) ways to take user capabilities and task characteristics into account and reduce or compensate for latency in different contexts, e. g., using a mouse or a touchpad, a touch-screen, an optical finger navigation device or a brain-computer interface. From an engineering perspective, this knowledge on low-level human factors will help us to rethink and redesign the I/O loop of interactive systems in order to better account for them and achieve more adapted and adaptable perception-action coupling.

3.2.2. Tactile feedback & haptic perception

We are also concerned with the physicality of human-computer interaction, with a focus on haptic perception and related technologies. For instance, when interacting with virtual objects such as software buttons on a touch surface, the user cannot feel the click sensation as with physical buttons. The tight coupling between how we perceive and how we manipulate objects is then essentially broken although this is instrumental for efficient direct manipulation. We have addressed this issue in multiple contexts by designing, implementing and evaluating novel applications of tactile feedback [5].

In comparison with many other modalities, one difficulty with tactile feedback is its diversity. It groups sensations of forces, vibrations, friction, or deformation. Although this is a richness, it also raises usability and technological challenges since each kind of haptic stimulation requires different kinds of actuators with their own parameters and thresholds. And results from one are hardly applicable to others. On a "knowledge" point of view, we want to better understand and empirically classify haptic variables and the kind of information they can represent (continuous, ordinal, nominal), their resolution, and their applicability to various contexts. From the "technology" perspective, we want to develop tools to inform and ease the design of haptic interactions taking best advantage of the different technologies in a consistent and transparent way.

3.3. Meso-Dynamics

Meso-dynamics relate to phenomena that arise during interaction, on a longer but still short time-scale. For users, it is related to performing intentional actions, to goal planning and tools selection, and to forming sequences of interactions based on a known set of rules or instructions. From the system perspective, it relates to how possible actions are exposed to the user and how they have to be executed (i. e., interaction techniques). It also has implication on the tools for designing and implementing those techniques (programming languages and APIs).

3.3.1. Interaction bandwidth and vocabulary

Interactive systems and their applications have an always-increasing number of available features and commands due to, e. g., the large amount of data to manipulate, increasing power and number of functionalities, or multiple contexts of use.

On the input side, we want to augment the *interaction bandwidth* between the user and the system in order to cope with this increasing complexity. In fact, most input devices capture only a few of the movements and actions the human body is capable of. Our arms and hands for instance have many degrees of freedom that are not fully exploited in common interfaces. We have recently designed new technologies to improve expressibility such as a bendable digitizer pen [36], or reliable technology for studying the benefits of finger identification on multi-touch interfaces [4].

On the output side, we want to expand users' *interaction vocabulary*. All of the features and commands of a system can not be displayed on screen at the same time and lots of *advanced* features are by default hidden to the users (e. g., hotkeys) or buried in deep hierarchies of command-triggering systems (e. g., menus). As a result, users tend to use only a subset of all the tools the system actually offers [44]. We will study how to help them to broaden their knowledge of available functions.

Through this "opportunistic" exploration of alternative and more expressive input methods and interaction techniques, we will particularly focus on the necessary technological requirements to integrate them into interactive systems, in relation with our redesign of the I/O stack at the micro-dynamics level.

3.3.2. Spatial and temporal continuity in interaction

At a higher-level, we will investigate how more expressive interaction techniques affect users' strategies when performing sequences of elementary actions and tasks. More generally, we will explore the "*continuity*" in interaction. Interactive systems have moved from one computer to multiple connected interactive devices (computer, tablets, phones, watches, etc.) that could also be augmented through a Mixed-Reality paradigm. This distribution of interaction raises new challenges from both usability and engineering perspectives that we clearly have to consider in our main objective of revisiting interactive systems [43]. It involves the simultaneous use of multiple devices and also the changes in the role of devices according to the location, the time, the task, and contexts of use: a tablet device can be used as the main device while traveling, and it becomes an input device or a secondary monitor for continuing the same task once in the office; a smart-watch can be used as a standalone device to send messages, but also as a remote controller for a wall-sized display. One challenge is then to design interaction techniques that support seamless and smooth transitions during these spatial and temporal changes of the system in order to maintain the continuity of uses and tasks, and how to integrate these principles in future interactive systems.

3.3.3. Expressive tools for prototyping, studying, and programming interaction

Current systems suffer from engineering issues that keep constraining and influencing how interaction is thought, designed, and implemented. Addressing the challenges we presented in this section and making the solutions possible require extended expressiveness, and researchers and designers must either wait for the proper toolkits to appear, or "hack" existing interaction frameworks, often bypassing existing mechanisms. For instance, numerous usability problems in existing interfaces stem from a common cause: the lack, or untimely discarding, of relevant information about how events are propagated and how changes come to occur in interactive environments. On top of our redesign of the I/O loop of interactive systems, we will investigate how to facilitate access to that information and also promote a more grounded and expressive way to describe and exploit input-to-output chains of events at every system level. We want to provide finer granularity and better-described connections between the *causes* of changes (e.g. input events and system triggers), their *context* (e.g. system and application states), their *consequences* (e.g. interface and data updates), and their *timing* [8]. More generally, a central theme of our Interaction Machine vision is to promote interaction as a first-class object of the system [33], and we will study alternative and better-adapted technologies for designing and programming interaction, such as we did recently to ease the prototyping of Digital Musical Instruments [1] or the programming of animations in graphical interfaces [10]. Ultimately, we want to propose a unified model of hardware and software scaffolding for interaction that will contribute to the design of our Interaction Machine.

3.4. Macro-Dynamics

Macro-dynamics involve longer-term phenomena such as skills acquisition, learning of functionalities of the system, reflexive analysis of its own use (e. g., when the user has to face novel or unexpected situations which require high-level of knowledge of the system and its functioning). From the system perspective, it implies to better support cross-application and cross-platform mechanisms so as to favor skill transfer. It also requires to improve the instrumentation and high-level logging capabilities to favor reflexive use, as well as flexibility and adaptability for users to be able to finely tune and shape their tools.

We want to move away from the usual binary distinction between "novices" and "experts" [3] and explore means to promote and assist digital skill acquisition in a more progressive fashion. Indeed, users have a permanent need to adapt their skills to the constant and rapid evolution of the tasks and activities they carry on a computer system, but also the changes in the software tools they use [47]. Software strikingly lacks powerful means of acquiring and developing these skills [3], forcing users to mostly rely on outside support (e. g., being guided by a knowledgeable person, following online tutorials of varying quality). As a result, users tend to rely on a surprisingly limited interaction vocabulary, or *make-do* with sub-optimal routines and tools [48]. Ultimately, the user should be able to master the interactive system to form durable and stabilized practices that would eventually become *automatic* and reduce the mental and physical efforts, making their interaction *transparent*.

In our previous work, we identified the fundamental factors influencing expertise development in graphical user interfaces, and created a conceptual framework that characterizes users' performance improvement with UIs [7], [3]. We designed and evaluated new command selection and learning methods to leverage user's digital skill development with user interfaces, on both desktop [6] and touch-based computers.

We are now interested in broader means to support the analytic use of computing tools:

- *to foster understanding of interactive systems.* As the digital world makes the shift to more and more complex systems driven by machine learning algorithms, we increasingly lose our comprehension of which process caused the system to respond in one way rather than another. We will study how novel interactive visualizations can help reveal and expose the "intelligence" behind, in ways that people better master their complexity.
- *to foster reflexion on interaction.* We will study how we can foster users' reflexion on their own interaction in order to encourage them to acquire novel digital skills. We will build real-time and off-line software for monitoring how user's ongoing activity is conducted at an application and system

level. We will develop augmented feedbacks and interactive history visualization tools that will offer contextual visualizations to help users to better understand and share their activity, compare their actions to that of others, and discover possible improvement.

- *to optimize skill-transfer and tool re-appropriation.* The rapid evolution of new technologies has drastically increased the frequency at which systems are updated, often requiring to relearn everything from scratch. We will explore how we can minimize the cost of having to appropriate an interactive tool by helping users to capitalize on their existing skills.

We plan to explore these questions as well as the use of such aids in several contexts like web-based, mobile, or BCI-based applications. Although, a core aspect of this work will be to design systems and interaction techniques that will be as little platform-specific as possible, in order to better support skill transfer. Following our Interaction Machine vision, this will lead us to rethink how interactive systems have to be engineered so that they can offer better instrumentation, higher adaptability, and fewer separation between applications and tasks in order to support reuse and skill transfer.

4. Application Domains

4.1. Application Domains

Loki works on fundamental and technological aspects of Human-Computer Interaction that can be applied to diverse application domains.
Our 2019 research involved desktop, augmented reality, touch-based, haptics, and BCI interfaces with notable applications to medicine (analysis of fine motor control for patients with Parkinson disease), digital humanities (interpretation of handwritten historical documents), as well as creativity support tools (production of illustrations, design of Digital Musical Instruments). Our technical work also contributes to the more general application domains of software engineering and systems' design.

5. Highlights of the Year

5.1. Highlights of the Year

Géry Casiez and Mathieu Nancel received a very selective Google Faculty Research Award for their project "Real-time Latency Measure and Compensation".

Mathieu Nancel contributed to the writing of the new NF Z71-300 French keyboard standard, and spoke at the official launch event in April 2019 at the National Assembly.

5.1.1. Personnel

Edward Lank joined Loki in October as part of the Inria International Chair program and will spend more than 50% of his time with us until 2023.

5.1.2. Awards

Best paper award from the ACM EICS conference to the paper "Polyphony: Programming Interfaces and Interactions with the Entity-Component-System Model", from T. Raffaillac & S. Huot .

Best paper award from the Francophone Conference on Human-Computer Interaction (IHM) to the paper "Reducing Error Aversion to Support Novice-to-Expert Transitions with FastTap", from A. Goguey, S. Malacria, A. Cockburn & C. Gutwin .
BEST PAPERS AWARDS :
[22]

T. RAFFAILLAC, S. HUOT.*Polyphony: Programming Interfaces and Interactions with the Entity-Component-System Model*, in "EICS 2019 - 11th ACM SIGCHI Symposium on Engineering Interactive Computing Systems", Valencia, Spain, June 2019, vol. 3 [*DOI :* 10.1145/3331150], https://hal.inria.fr/hal-02147180

[25]
A. GOGUEY, S. MALACRIA, A. COCKBURN, C. GUTWIN.*Reducing Error Aversion to Support Novice-to-Expert Transitions with FastTap*, in "Actes de la 31e conférence francophone sur l'Interaction Homme-Machine (IHM 2019)", Grenoble, France, ACM, 2019, p. 1:1-10 [*DOI :* 10.1145/3366550.3372247], https://hal.archives-ouvertes.fr/hal-02381584

6. New Software and Platforms

6.1. Polyphony

KEYWORDS: Human Computer Interaction - Toolkit - Engineering of Interactive Systems

FUNCTIONAL DESCRIPTION: Polyphony is an experimental toolkit demonstrating the use of Entity-Component-System (ECS) to design Graphical User Interfaces (GUI). It also extends the original ECS model to support advanced interfaces.

NEWS OF THE YEAR: Design and implementation of a first version of the toolkit and associated examples as a proof of concept.

- Participants: Thibault Raffaillac and Stéphane Huot
- Contact: Stéphane Huot
- Publications: Applying the Entity-Component-System Model to Interaction Programming - Polyphony: Programming Interfaces and Interactions with the Entity-Component-System Model
- URL: https://gitlab.inria.fr/Loki/PolyphonyECS

6.2. Esquisse

KEYWORDS: Vector graphics - 3D interaction - Human Computer Interaction

SCIENTIFIC DESCRIPTION: Trace figures are contour drawings of people and objects that capture the essence of scenes without the visual noise of photos or other visual representations. Their focus and clarity make them ideal representations to illustrate designs or interaction techniques. In practice, creating those figures is a tedious task requiring advanced skills, even when creating the figures by tracing outlines based on photos. To mediate the process of creating trace figures, we introduce the open-source tool Esquisse. Informed by our taxonomy of 124 trace figures, Esquisse provides an innovative 3D model staging workflow, with specific interaction techniques that facilitate 3D staging through kinematic manipulation, anchor points and posture tracking. Our rendering algorithm (including stroboscopic rendering effects) creates vector-based trace figures of 3D scenes. We validated Esquisse with an experiment where participants created trace figures illustrating interaction techniques, and results show that participants quickly managed to use and appropriate the tool.

FUNCTIONAL DESCRIPTION: Esquisse is an add-on for Blender that can be used to rapidly produce trace figures. It relies on a 3D model staging workflow, with specific interaction techniques that facilitate the staging through kinematic manipulation, anchor points and posture tracking. Staged 3D scenes can be exported to SVG thanks to Esquisse's dedicated rendering algorithm.

NEWS OF THE YEAR: First version of Esquisse, implementing both staging and vector rendering.

- Contact: Sylvain Malacria
- Publication: Esquisse: Using 3D Models Staging to Facilitate the Creation of Vector-based Trace Figures
- URL: https://github.com/LokiResearch/Esquisse

6.3. RayCursor

Source code for the pointing technique RayCursor

KEYWORDS: Virtual reality - Interaction technique

FUNCTIONAL DESCRIPTION: This is a Unity Project containing the source code and prefab for the pointing technique RayCursor to be easily integrated in other Unity Projects.

- Contact: Géry Casiez
- Publications: RayCursor: a 3D Pointing Facilitation Technique based on Raycasting - Improving Raycasting using Proximity Selection and Filtering
- URL: http://ns.inria.fr/loki/raycursor/

7. New Results

7.1. Introduction

According to our research program, in the next two to five years, we will study dynamics of interaction along three levels depending on interaction time scale and related user's perception and behavior: *Micro-dynamics*, *Meso-dynamics*, and *Macro-dynamics*. Considering phenomena that occur at each of these levels as well as their relationships will help us to acquire the necessary knowledge (Empowering Tools) and technological bricks (Interaction Machine) to reconcile the way interactive systems are designed and engineered with human abilities. Our strategy is to investigate issues and address challenges for all of the three levels of dynamics. Last year we focused on micro-dynamics since it concerns very fundamental knowledge about interaction and relates to very low-level parts of interactive systems. In 2019 we were able to build upon those results (micro), but also to enlarge the scope of our studies within larger interaction time scales, especially at the meso-dynamic level. Some of these results have also contributed to our objective of defining the basic principles of an Interaction Machine.

7.2. Micro-dynamics

Participants: Géry Casiez [contact person], Sylvain Malacria, Mathieu Nancel, Thomas Pietrzak.

7.2.1. Latency & Transfer functions

End-to-end latency in interactive systems is detrimental to performance and usability, and comes from a combination of hardware and software delays. While these delays are steadily addressed by hardware and software improvements, it is at a decelerating pace. In parallel, short-term input prediction has recently shown promising results to compensate for latency, in both research and industry.

In the context of the collaborative TurboTouch project, we proposed a method based on a frequency-domain approximation of a non-causal ideal predictor with a finite impulse response filter. Given a sufficiently rich dataset, the parameters of the filter can be either optimized off-line or tuned on-line with the proposed adaptive algorithm. The performance of the proposed solution is evaluated in an experimental study consisting of drawings on a touchscreen [13].

On the related topic of transfer functions, we proposed a switched dynamic model to model indirect pointing tasks with a computer mouse. The model contains a ballistic movement phase governed by a nonlinear model in Lurie form and a corrective movement phase described by a linear visual-feedback system. The stability of the model was evaluated and the derived model was then validated with experimental data acquired in a pointing task with a mouse. Numerical comparison to pointing models available in the literature is also provided [12].

7.2.2. 3D interaction

Raycasting is the most common target pointing technique in virtual reality environments. However, performance on small and distant targets is impacted by the accuracy of the pointing device and the user's motor skills. Current pointing facilitation techniques are currently only applied in the context of the virtual hand, i.e. for targets within reach. We proposed enhancements to Raycasting: filtering the ray, and adding a controllable cursor on the ray to select the nearest target (Figure 2). We ran a series of studies for the design of the visual feedforward, filtering technique, as well as a comparative study between different 3D pointing techniques. Our results show that highlighting the nearest target is one of the most efficient visual feedforward technique. We also show that filtering the ray reduces error rate in a drastic way. Finally we show the benefits of RayCursor compared to Raycasting and another technique from the literature [19], [14].

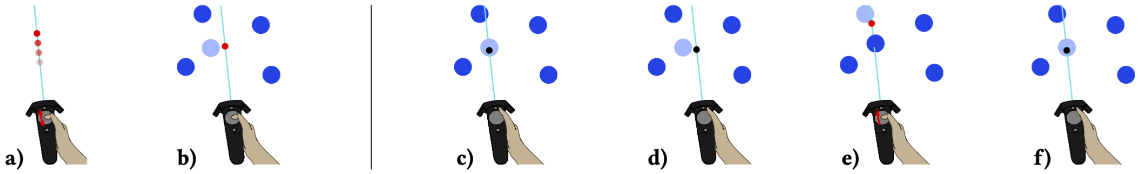

Figure 2. Illustration of manual RayCursor: a) the user controls a cursor along the ray using relative displacements of their thumb on the controller's touchpad; b) the target closest to the cursor is highlighted. Illustration of semi-auto RayCursor: c) by default, it works like Raycasting. The cursor (in black) is positioned at the intersection with a target; d) the target remains selected if the cursor moves out of the target, until it is closer to another target; e) the user can manually move the cursor using the controller's touchpad, to select another target (the cursor turns red to indicate manual mode); f) if the user does not touch the touchpad for 1s, the cursor returns to its behaviour described in c).

7.3. Meso-dynamics

Participants: Axel Antoine, Marc Baloup, Géry Casiez, Stéphane Huot, Edward Lank, Sylvain Malacria, Mathieu Nancel, Thomas Pietrzak [contact person], Thibault Raffaillac, Marcelo Wanderley.

7.3.1. Production of illustrative supports

Trace figures are contour drawings of people and objects that capture the essence of scenes without the visual noise of photos or other visual representations. Their focus and clarity make them ideal representations to illustrate designs or interaction techniques. In practice, creating those figures is a tedious task requiring advanced skills, even when creating the figures by tracing outlines based on photos. To mediate the process of creating trace figures, we introduce the open-source tool Esquisse (Figure 3). Informed by our taxonomy of 124 trace figures, Esquisse provides an innovative 3D model staging workflow, with specific interaction techniques that facilitate 3D staging through kinematic manipulation, anchor points and posture tracking. Our rendering algorithm (including stroboscopic rendering effects) creates vector-based trace figures of 3D scenes. We validated Esquisse with an experiment where participants created trace figures illustrating interaction techniques, and results show that participants quickly managed to use and appropriate the tool [18].

7.3.2. Impact of confirmation modes on expert interaction techniques adoption

Expert interaction techniques such as gestures or keyboard shortcuts are more efficient than traditional WIMP techniques because it is often faster to recall a command than to navigate to it. However, many users seem to be reluctant to switch to expert interaction. We hypothesized the cause might be the aversion to making errors. To test this, we designed two intermediate modes for the FastTap interaction technique, allowing quick confirmation of what the user has retrieved from memory, and quick adjustment if she made an error. We investigated the impact of these modes and of various error costs in a controlled study, and found that

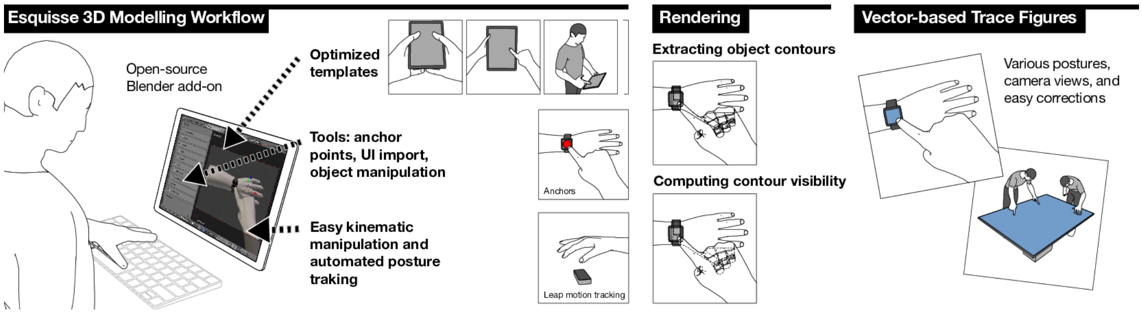

Figure 3. Workflow of Esquisse: (left) facilitating staging of 3D scenes with templates, anchor points, direct UI import, kinematic manipulation and automated posture tracking, (center) Esquisse's algorithm rendering the tracing images, and (right) final vector-based trace figures.

participants adopted the intermediate modes, that these modes reduced error rate when the cost of errors was high, and that they did not substantially change selection times. However, while it validates the design of our intermediate modes, we found no evidence of greater switch to memory-based interaction, suggesting that reducing error rate is not sufficient to motivate the adoption of expert use of techniques [25].

7.3.3. *Effect of the context on mobile interaction*

7.3.3.1. Pointing techniques for eyewear using a simulated pedestrian environment

Eyewear displays allow users to interact with virtual content displayed over real-world vision, in active situations like standing and walking. Pointing techniques for eyewear displays have been proposed, but their social acceptability, efficiency, and situation awareness remain to be assessed. Using a novel street-walking simulator, we conducted an empirical study of target acquisition while standing and walking under different levels of street crowdedness. Results showed that indirect touch was the most efficient and socially acceptable technique, and that in-air pointing was inefficient when walking. Interestingly, the eyewear displays did not improve situation awareness compared to the control condition [23].

7.3.3.2. Studying smartphone motion gestures in private or public contexts

We also investigated the effect of social exposure on smartphone motion gestures. We conducted a study where participants performed sets of motion gestures on a smartphone in both private and public locations. Using data from the smartphone's accelerometer, we found that the location had a significant effect on both the duration and intensity of the participants' gestures. We concluded that it may not be sufficient for gesture input systems to be designed and calibrated purely in private lab settings. Instead, motion gesture input systems for smartphones may need to be aware of the changing context of the device and to account for this in algorithms that interpret gestural input [26].

7.4. Macro-dynamics

Participants: Stéphane Huot, Sylvain Malacria [contact person], Nicole Pong.

7.4.1. *Awareness, usage and discovery of hidden controls*

Revealing a hidden widget with a dedicated sliding gesture is a common interaction design in today's handheld devices. Such "Swhidgets" (for swipe-revealed hidden widgets) provide a fast (and sometime unique) access to some commands. Interestingly, swhidgets do not follow conventional design guidelines in that they have no explicit signifiers, and users have to discover their existence before being able to use them. We conducted the first two studies specifically targeted to this type of interface design, investigating iOS users' experience with swhidgets. The first study conducted in a laboratory setting investigated which Swhidgets are spontaneously used by participants when prompted to perform certain operations on an iOS device. The second study conducted via an online survey platform, investigated which Swhidgets users reported to know and use. Combined, our studies provide the following main insights on awareness, usage and discovery of Swhidgets by middle-aged and technology-friendly users. Our results suggest that Swhidgets are moderately but unevenly known by participants, yet the awareness and the discovery issues of this design is worthy of further discussion [21].

7.5. Interaction Machine

Two of our contributions this year relate specifically to our Interaction Machine project.

7.5.1. Definition of Brain-Computer Interfaces

Regardless of the term used to designate them, Brain-Computer Interfaces are "Interfaces" between a user and a computer in the broad sense of the term. We provided a perspective to discuss how BCIs have been defined in the literature from the day the term was introduced by Jacques Vidal. From a Human-Computer Interaction perspective, we propose a new definition of Brain-Computer Interfaces as "any artificial system that transforms brain activity into input of a computer process" [24]. As they are interfaces, their definition should not include the finality and objective of the system they are used to interact with. To illustrate this, we compared BCIs with other widely used Human-Computer Interfaces, and draw analogies in their conception and purpose. This definition would help better encompassing for such interfaces in systems design, and more generally inform on how to better manage diverse forms of input in an Interaction Machine.

7.5.2. Software architecture for interactive systems

On the software engineering side, we have proposed a new Graphical User Interface (GUI) and Interaction framework based on the Entity-Component-System model (ECS) [22]. In this model, interactive elements (Entities) are characterized only by their data (Components). Behaviors are managed by continuously running processes (Systems) which select entities by the Components they possess. This model facilitates the handling of behaviors and promotes their reuse. It provides developers with a simple yet powerful composition pattern to build new interactive elements with Components. It materializes interaction devices as Entities and interaction techniques as a sequence of Systems operating on them. We have implemented these principles in the Polyphony toolkit in order to experiment the ECS model in the context of GUIs programming. It has proven to be useful and efficient for modeling standard interaction techniques, and we are now exploring its benefits for prototyping and implementing more advanced methods in a modular way. It also raises some interesting challenges about performance and scalability that we will explore further.

7.5.3. From the dynamics of interaction to an Interaction Machine

Several of our new results this year also informed our global objective of building an Interaction Machine. At the micro-dynamics level, as last year, our work on prediction algorithms and transfer functions highlighted the need for accessing low-level input data and to have flexible input management to be able to reliably predict current finger position and compensate for latency. Our work on new selection methods in 3D also highlighted the importance of easing the combination of input events from multiple sources and of data filtering to achieve better interaction. As it also leverages the real time aspect of the perception-action coupling for efficient interaction, it also confirms the need for efficient and low-latency input management stacks. These results give us the first leads to redefine input management and input events propagation in order to better account for human factors in interactive systems, and to extend the possibilities for designing more efficient and expressive interaction methods.

At the meso-dynamics level, our studies on the adoption of expert interaction techniques and of the impact of the context in performing interaction gestures highlighted the need for both adaptable and adaptive systems (e.g. context-based calibration of gesture recognition algorithms), which require more modular and flexible system architectures in order to enable real-time parametrization or even switching interaction techniques. These results also resonate with those at the micro-dynamics level, since they suggest strong links between users' behaviors and strategies (meso) and their low-level perception mechanisms (micro) that should be better taken into account in the design of interactive systems.

These conclusions and observations will be the basis for our investigations on the topic next year. We will in particular focus on the redefinition of the input stack and on applying the ECS model to the whole architecture of an interactive system.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

A research agreement between Synaptics Incorporated (San Jose, California) and Inria/Loki has been signed in September 2019, for a duration of nine months. The goal is to conduct joint studies on the impact of touchpads' characteristics (size, resolution) on the quality of interaction and users' performance.

8.2. Bilateral Grants with Industry

Géry Casiez and Mathieu Nancel have been awarded a Google Faculty Research Award for their project "Real-time Latency Measure and Compensation".

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. GeneaLire (CPER MAuVE, 2018-2020)

Participants: Stéphane Huot, Thomas Pietrzak [contact person].

Interactive tools for the interpretation of manuscripts

The goal of this project is to design, implement and evaluate interactive tools for helping transcription of scanned handwritten documents. Current solutions focus on automatic recognition, with recent advances thanks to deep learning methods. However these solutions still require a significant learning base that has to be made by hand. Not only this means that part of the work cannot be done automatically, but it also means that this technique is not a solution for small collections of documents. The tools we propose to create will ingeniously take advantage of interactive and automatic techniques. The interactive tools include a text selection technique [32], as well as advanced annotation techniques that will support collaborative work. This tool will be invaluable for bootstrapping the transcription of large collections, as well as helping transcribing small collections. We will use user-centered design, in order to make sure the tool fits historians and genealogists activities and workflow.

Partners: Inria Saclay's AVIZ team, École Polytechnique de l'Université de Tours, Laboratoire de Démographie et d'Histoire Sociale at l'École des hautes études en sciences sociales, and Geneanet.

9.2. National Initiatives

9.2.1. ANR

9.2.1.1. TurboTouch (PRC, 2014-2019)

Participants: Géry Casiez [contact person], Sylvain Malacria, Mathieu Nancel, Thomas Pietrzak.

High-performance touch interactions

Touch-based interactions with computing systems are greatly affected by two interrelated factors: the transfer functions applied on finger movements, and latency. This project aims at transforming the design of touch transfer functions from black art to science to support high-performance interactions. We are working on the precise characterization of the functions used and the latency observed in current touch systems. We are developing a testbed environment to support multidisciplinary research on touch transfer functions and will use this testbed to design latency reduction and compensation techniques, and new transfer functions.

Partners: Inria Lille's VALSE team (formerly NON-A) and the "Perceptual-motor behavior group" from the Institute of Movement Sciences.

Web site: http://mjolnir.lille.inria.fr/turbotouch/

Related publications in 2019: [13], [12]

9.2.1.2. Causality (JCJC, 2019-2023)

Participants: Géry Casiez, Stéphane Huot, Sylvain Malacria, Mathieu Nancel [contact person], Philippe Schmid.

Integrating Temporality and Causality to the Design of Interactive Systems

The project addresses a fundamental limitation in the way interfaces and interactions are designed and even thought about today, an issue we call *procedural information loss*: once a task has been completed by a computer, significant information that was used or produced while processing it is rendered inaccessible regardless of the multiple other purposes it could serve. It hampers the identification and solving of identifiable usability issues, as well as the development of new and beneficial interaction paradigms. We will explore, develop, and promote finer granularity and better-described connections between the causes of those changes, their context, their consequences, and their timing. We will apply it to facilitate the real-time detection, disambiguation, and solving of frequent timing issues related to human reaction time and system latency; to provide broader access to all levels of input data, therefore reducing the need to "hack" existing frameworks to implement novel interactive systems; and to greatly increase the scope and expressiveness of command histories, allowing better error recovery but also extended editing capabilities such as reuse and sharing of previous actions.

Web site: http://loki.lille.inria.fr/causality/

9.2.1.3. Discovery (JCJC, 2020-2024)

Participant: Sylvain Malacria [contact person].

Promoting and improving discoverability in interactive systems

This project addresses a fundamental limitation in the way interactive systems are usually designed, as in practice they do not tend to foster the discovery of their input methods (operations that can be used to communicate with the system) and corresponding features (commands and functionalities that the system supports). Its objective is to provide generic methods and tools to help the design of discoverable interactive systems: we will define validation procedures that can be used to evaluate the discoverability of user interfaces, design and implement novel UIs that foster input method and feature discovery, and create a design framework of discoverable user interfaces. This project investigates, but is not limited to, the context of touch-based interaction and will also explore two critical timings when the user might trigger a reflective practice on the available inputs and features: while the user is carrying her task (discovery in-action); and after having carried her task by having informed reflection on her past actions (discovery on-action). This dual investigation will reveal more generic and context-independent properties that will be summarized in a comprehensive framework of discoverable interfaces. Our ambition is to trigger a significant change in the way all interactive systems and interaction techniques, existing and new, are thought, designed, and implemented with both performance and discoverability in mind.

Web site: http://ns.inria.fr/discovery/

Related publications in 2019: [21].

9.2.2. Inria Project Labs

9.2.2.1. BCI-LIFT (2015-2019)

Participant: Géry Casiez [contact person].

Brain Computer Interfaces: Learning, Interaction, Feedback, Training

The goal of this large-scale initiative is to design a new generation of non-invasive Brain-Computer Interfaces (BCI) that are easier to appropriate, more efficient, and suited for a larger number of people.

Partners: Inria's ATHENA, NEUROSYS, POTIOC, HYBRID & DEMAR teams, *Centre de Recherche en Neurosciences de Lyon* (INSERM) and INSA Rouen.

Web site: https://bci-lift.inria.fr/

Related publication in 2019: [24]

9.2.2.2. AVATAR (2018-2022)

Participants: Géry Casiez, Stéphane Huot, Thomas Pietrzak [contact person].

The next generation of our virtual selves in digital worlds

This project aims at delivering the next generation of virtual selves, or *avatars*, in digital worlds. In particular, we want to push further the limits of perception and interaction through our avatars to obtain avatars that are better embodied and more interactive. Loki's contribution in this project consists in designing novel 3D interaction paradigms for avatar-based interaction and to design new multi-sensory feedbacks to better feel our interactions through our avatars.

Partners: Inria's GRAPHDECO, HYBRID, MIMETIC, MORPHEO & POTIOC teams, Mel Slater (Event Lab, University Barcelona, Spain), Technicolor and Faurecia.

Web site: https://avatar.inria.fr/

Related publication in 2019: [19], [14]

9.2.3. Others

9.2.3.1. ParkEvolution (Carnot Inria - Carnot STAR, 2015-2019)

Participant: Géry Casiez [contact person].

Longitudinal analysis of fine motor control for patients with Parkinson disease

This project studies the fine motor control of patients with Parkinson disease in an ecological environment, at home, without the presence of experimenters. Through longitudinal studies, we collect raw information from pointing devices to create a large database of pointing behavior data. From the analysis of this big dataset, the project aims at inferring the individual's disease progression and influence of treatments.

Partners: the "Perceptual-motor behavior group" from the Institute of Movement Sciences and Hôpital de la Timone.

Web site: http://parkevolution.org/

9.2.3.2. IRDICS (Projets Exploratoires Premier Soutien CNRS, 2018-2019)

Participants: Géry Casiez, Stéphane Huot, Sylvain Malacria, Thomas Pietrzak [contact person].

Interface de recueil de données imparfaites pour le crowd-sourcing

Many crowdsourcing studies involve asking hundreds of participants to answer questionnaires. There is typically a trade-off between precision and certitude of participants. Usually, investigators prefer participants to be certain, at the cost of precision. The idea is that the lack of precision can be compensated by the high number of answers. In this project we are interested in studying this trade-off. We performed a first study, in which we asked participants to rate their confidence in their answer. In the next studies, we will allow participants to give several answers, but make sure the right answer is among them. In the last study, participants will be able to rank their answers based on confidence.

Partners: IRISA's DRUID team.

Related publication in 2019: [31]

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

Andy Cockburn, University of Canterbury, Christchurch, NZ [25], [23]

Carl Gutwin, University of Saskatchewan, Saskatoon, CA [25]

Nicolai Marquardt, University College London, London, UK [18]

Antti Oulasvirta, Aalto University, Helsinki, FI

Daniel Vogel, University of Waterloo, Waterloo, CA

Audrey Girouard, Carleton University, Ottawa, CA

9.3.2. Participation in Other International Programs

9.3.2.1. Inria International Chairs

Expert interaction with devices for musical expression
Marcelo M. Wanderley – *Professor at Schulich School of Music/IDMIL, McGill University (Canada)*

Period: 2017 - 2021

The main topic of this project is the expert interaction with devices for musical expression and consists of two main directions: *the design of digital musical instruments* (DMIs) and *the evaluation of interactions with such instruments*. It will benefit from the unique, complementary expertise available at the Loki Team, including the design and evaluation of interactive systems, the definition and implementation of software tools to track modifications of, visualize and haptically display data, as well as the study of expertise development within human-computer interaction contexts. The project's main goal is to bring together advanced research on devices for musical expression (IDMIL – McGill) and cutting-edge research in Human-computer interaction (Loki Team).

Rich, Reliable Interaction in Ubiquitous Environments
Edward Lank – *Professor at Cheriton School of Computer Science, University of Waterloo (Canada)*

Period: 2019 - 2023

The objectives of the research program are:

1. Designing Rich Interactions for Ubiquitous and Augmented Reality Environments
2. Designing Mechanisms and Metaphors for Novices, Experts, and the Novice to Expert Transition
3. Integrating Intelligence with Human Action in Richly Augmented Environments.

9.3.2.2. Université de Lille - International Associate Laboratory

Reappearing Interfaces in Ubiquitous Environments (Réapp)
with Edward Lank, Daniel Vogel & Keiko Katsuragawa at University of Waterloo (Canada) - Cheriton School of Computer Science

Duration: 2019 - 2023

The LIA Réapp is an International Associated Laboratory between Loki and Cheriton School of Computer Science from the University of Waterloo in Canada. It is funded by the University of Lille to ease shared student supervision and regular inter-group contacts. The University of Lille will also provide a grant for a co-tutelle PhD thesis between the two universities.

We are at the dawn of the next computing paradigm where everything will be able to sense human input and augment its appearance with digital information without using screens, smartphones, or special glasses—making user interfaces simply disappear. This introduces many problems for users, including the discoverability of commands and use of diverse interaction techniques, the acquisition of expertise, and the balancing of trade-offs between inferential (AI) and explicit (user-driven) interactions in aware environments. We argue that interfaces must reappear in an appropriate way to make ubiquitous environments useful and usable. This project tackles these problems, addressing (1) the study of human factors related to ubiquitous and augmented reality environments, and the development of new interaction techniques helping to make interfaces reappear; (2) the improvement of transition between novice and expert use and optimization of skill transfer; and, last, (3) the question of delegation in smart interfaces, and how to adapt the trade-off between implicit and explicit interaction.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

Edward Lank, Professor at the University of Waterloo, who has been awarded an Inria International Chair in our team in 2019, spent 4 months in our group this year (September to December).

Marcelo M. Wanderley, Professor at McGill University, who has been awarded an Inria International Chair in our team in 2017, spent 2 months in our group this year (July to August).

9.4.1.1. Internships

Carola Trahms, PhD student, Christian-Albrecht University of Kiel, Sep. 2019.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organization

10.1.1.1. General Chair, Scientific Chair

- HAID 2019: Thomas Pietrzak, Marcelo Wanderley

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- IHM (AFIHM): Thomas Pietrzak

10.1.2.2. Member of the Conference Program Committees

- CHI (ACM): Géry Casiez, Stéphane Huot, Thomas Pietrzak
- IHM (AFIHM): Mathieu Nancel, Thibault Raffaillac (Work in Progress)
- WebMedia (SBC): Raiza Hanada
- ISS (ACM): Sylvain Malacria

10.1.2.3. Reviewer

- CHI (ACM): Géry Casiez, Stéphane Huot, Sylvain Malacria, Mathieu Nancel, Thomas Pietrzak
- UIST (ACM): Géry Casiez, Sylvain Malacria, Mathieu Nancel, Thomas Pietrzak
- IDC (ACM): Géry Casiez
- DIS (ACM): Stéphane Huot, Mathieu Nancel
- GI: Géry Casiez, Mathieu Nancel
- ISS (ACM): Mathieu Nancel, Thomas Pietrzak
- SUI (ACM): Mathieu Nancel

10.1.3. Journal

10.1.3.1. Reviewer - Reviewing Activities

- Transactions on Computer-Human Interaction (ACM): Géry Casiez, Thomas Pietrzak, Mathieu Nancel
- Intl. Journal of Human-Computer Studies (Elsevier): Sylvain Malacria
- User Modeling and User-Adapted Interaction (Springer): Sylvain Malacria

10.1.4. Invited Talks

- *Who [was|is|will be] in Control?: Some Control-Related Challenges in Interacting with Complex Systems*, 2nd JST-ANR Joint Symposium: Symbiotic Interaction, Tokyo, Japan: Stéphane Huot

10.1.5. Leadership within the Scientific Community

Association Francophone d'Interaction Homme-Machine (AFIHM):

- Géry Casiez: member of the scientific council since March 2019
- Stéphane Huot: member of the scientific council until March 2019

10.1.6. Scientific Expertise

- Agence Nationale de la Recherche: Stéphane Huot (vice-chair for the CES33 "Interaction and Robotics" committee)
- CN35 AFNOR normalization committee about normalizing the French keyboard: Mathieu Nancel (publication of the NF Z71-300 standard)

10.1.7. Research Administration

10.1.7.1. For Inria

- International relations working group (COST-GTRI): Stéphane Huot (member)

10.1.7.2. For Inria Lille – Nord Europe

- "Bureau du comité des équipes projets" (BCEP): Stéphane Huot (member)
- "Commission des Emplois de recherche du centre Inria Lille – Nord Europe" (CER): Sylvain Malacria (member)
- "Comité opérationnel d'évaluation des risques légaux et éthiques" (COERLE, the Inria Ethics board): Stéphane Huot (local correspondent)
- "Commission de développement technologique" (CDT): Mathieu Nancel (member)

10.1.7.3. For the CRIStAL lab of Université de Lille & CNRS

- Direction board: Géry Casiez
- Computer Science PhD recruiting committee: Géry Casiez (member)

10.1.7.4. For the Université de Lille

- Coordinator for internships at IUT A: Géry Casiez
- Computer Science Department council: Thomas Pietrzak

10.1.8. Hiring committees

- Université de Valenciennes hiring committee for a Computer Science Professor position: Géry Casiez (member)
- ENAC hiring committee for a Computer Science faculty position: Stéphane Huot (member)

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

DUT Informatique: Géry Casiez (38h), Stéphane Huot (28h), Axel Antoine (28h), Thibault Raffaillac (56h) *IHM*, 1st year, IUT A de Lille - Université de Lille

DUT Informatique: Axel Antoine (32h), CDIN-CALC, 1st year, IUT A de Lille - Université de Lille

DUT Informatique: Marc Baloup, Algorithme et Programmation, 21.5h, 1st year, IUT A de Lille - Université de Lille

DUT Informatique: Marc Baloup, CDIN Web, 11h, 1st year, IUT A de Lille - Université de Lille

DUT Informatique: Thibault Raffaillac, Codage et Systèmes d'exploitation, 64h, 1st year, IUT A de Lille - Université de Lille

DUT Informatique: Thibault Raffaillac, Algorithmique avancée, 64h, 1st year, IUT A de Lille - Université de Lille

Licence Informatique: Thomas Pietrzak, *Logique*, 36h, L3, Université de Lille

Licence Informatique: Thomas Pietrzak, *Image et Interaction 2D*, 10.5h, L3, Université de Lille

Cursus ingénieur: Sylvain Malacria (10h), *3DETech*, IMT Lille-Douai

Master Informatique: Thomas Pietrzak (18h), *NIHM*, M2, Université de Lille

Master Informatique: Sylvain Malacria (12h), *NIHM*, M2, Université de Lille

Master Informatique: Thomas Pietrzak (34.5h), Sylvain Malacria (34.5), *IHM*, M1, Université de Lille

Master Informatique: Mathieu Nancel, *Evaluation*, 8h, M2, Université de Lille

10.2.2. Supervision

PhD: Thibault Raffaillac, *Languages and System Infrastructure for Interaction*, defended in Dec. 2019, advised by Stéphane Huot

PhD: Hakim Si Mohammed, *Improving Interaction Based on a Brain-Computer Interface*, defended in Dec. 2019, advised by Anatole Lecuyer, Ferran Argelaguet, Géry Casiez & Nicolas Roussel (in Rennes)

PhD: Jeronimo Barbosa, *Design and Evaluation of Digital Musical Instruments*, McGill University, defended in May 2019, advised by Marcelo Wanderley & Stéphane Huot (in Montréal)

PhD in progress: Grégoire Richard, *Touching avatars : le rôle du retour haptique dans les interactions avec les avatars en réalité virtuelle*, started Oct. 2019, advised by Géry Casiez & Thomas Pietrzak

PhD in progress: Philippe Schmid, *Command History as a Full-fledged Interactive Object*, started Oct. 2019, advised by Mathieu Nancel & Stéphane Huot

PhD: Damien Masson, *Supporting Interactivity with Static Content*, University of Waterloo, started in January 2019, advised by Edward Lank, Géry Casiez& Sylvain Malacria (in Waterloo)

PhD in progress: Marc Baloup, *Interaction with avatars in immersive virtual environments*, started Oct. 2018, advised by Géry Casiez & Thomas Pietrzak

PhD in progress: Axel Antoine, *Helping Users with Interactive Strategies*, started Oct. 2017, advised by Géry Casiez & Sylvain Malacria

PhD in progress: Nicole Ke Chen Pong, *Understanding and Improving Users Interactive Vocabulary*, started Oct. 2016, advised by Nicolas Roussel, Sylvain Malacria & Stéphane Huot

10.2.3. Juries

Hugo Romat (PhD, Université Paris-Saclay): Stéphane Huot, president

Emmanouil Giannisakis (PhD, Université Paris-Saclay): Géry Casiez, reviewer

Marie-Éléonore Kessaci (HDR, Université de Lille): Géry Casiez, president

10.2.4. Mid-term evaluation committees

Antonin Durey (PhD, Université de Lille): Géry Casiez

Théo Voillemin (PhD, Université de Lille): Géry Casiez

Vikas Mishra (PhD, Université de Lille): Géry Casiez

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

AIRLab selection committee for the funding of art and science projects, Stéphane Huot (representative for Inria Lille – Nord Europe)

10.3.2. Articles and contents

The Mother Of All Demos, l'autre révolution de 1968, interview for Carré, Petit, Utile : Le programme radio des gens du numérique on Radio <FMR>, Stéphane Huot

web series *Dopamine* on Arte.tv, scientific expertise, Stéphane Huot

10.3.3. Interventions

Intervention during the official launch event of the new NF Z71-300 French keyboard standard at the Assemblée Nationale: Mathieu Nancel

11. Bibliography

Major publications by the team in recent years

[1] F. Calegario, M. Wanderley, S. Huot, G. Cabral, G. Ramalho.*A method and toolkit for digital musical instruments: generating ideas and prototypes*, in "IEEE MultiMedia", January 2017, vol. 24, n^{o} 1, p. 63-71, https://doi.org/10.1109/MMUL.2017.18

[2] G. Casiez, S. Conversy, M. Falce, S. Huot, N. Roussel.*Looking through the eye of the mouse: a simple method for measuring end-to-end latency using an optical mouse*, in "Proceedings of UIST'15", ACM, November 2015, p. 629-636, http://dx.doi.org/10.1145/2807442.2807454

[3] A. Cockburn, C. Gutwin, J. Scarr, S. Malacria.*Supporting novice to expert transitions in user interfaces*, in "ACM Computing Surveys", November 2014, vol. 47, n^{o} 2, http://dx.doi.org/10.1145/2659796

[4] A. Goguey, D. Vogel, F. Chevalier, T. Pietrzak, N. Roussel, G. Casiez.*Leveraging finger identification to integrate multi-touch command selection and parameter manipulation*, in "International Journal of Human-Computer Studies", March 2017, vol. 99, p. 21-36, http://dx.doi.org/10.1016/j.ijhcs.2016.11.002

[5] A. Gupta, T. Pietrzak, N. Roussel, R. Balakrishnan.*Direct manipulation in tactile displays*, in "Proceedings of CHI'16", ACM, May 2016, p. 3683-3693, http://dx.doi.org/10.1145/2858036.2858161

[6] S. Malacria, G. Bailly, J. Harrison, A. Cockburn, C. Gutwin.*Promoting hotkey use through rehearsal with ExposeHK*, in "Proceedings of CHI'13", ACM, April 2013, p. 573-582, http://doi.acm.org/10.1145/2470654.2470735

[7] S. Malacria, J. Scarr, A. Cockburn, C. Gutwin, T. Grossman.*Skillometers: reflective widgets that motivate and help users to improve performance*, in "Proceedings of UIST'13", ACM, October 2013, p. 321-330, http://doi.acm.org/10.1145/2501988.2501996

[8] M. NANCEL, A. COCKBURN.*Causality: a conceptual model of interaction history*, in "Proceedings of CHI'14", ACM, April 2014, p. 1777-1786, http://dx.doi.org/10.1145/2556288.2556990

[9] M. NANCEL, D. VOGEL, B. DE ARAÙJO, R. JOTA, G. CASIEZ.*Next-point prediction metrics for perceived spatial errors*, in "Proceedings of UIST'16", ACM, October 2016, p. 271-285, http://dx.doi.org/10.1145/2984511.2984590

[10] T. RAFFAILLAC, S. HUOT, S. DUCASSE.*Turning Function Calls Into Animations*, in "Proceedings of EICS'17", ACM, June 2017, p. 81-86, http://doi.acm.org/10.1145/3102113.3102134

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] J. BARBOSA.*Direct Manipulation for Musical Interfaces Based on Visual Software: Design and Evaluation*, McGill University, May 2019, https://hal.inria.fr/tel-02436801

Articles in International Peer-Reviewed Journal

[12] S. ARANOVSKIY, R. USHIROBIRA, D. EFIMOV, G. CASIEZ.*A switched dynamic model for pointing tasks with a computer mouse*, in "Asian Journal of Control", 2019, forthcoming [*DOI :* 10.1002/ASJC.2063], https://hal-centralesupelec.archives-ouvertes.fr/hal-02083760

[13] S. ARANOVSKIY, R. USHIROBIRA, D. EFIMOV, G. CASIEZ.*An adaptive FIR filter for trajectory prediction and latency reduction in direct Human-Computer interactions*, in "Control Engineering Practice", 2019, forthcoming [*DOI :* 10.1016/J.CONENGPRAC.2019.07.011], https://hal.inria.fr/hal-02189181

[14] M. BALOUP, T. PIETRZAK, G. CASIEZ.*Improving Raycasting using Proximity Selection and Filtering*, in " Journal d'Interaction Personne-Système", December 2019, vol. 8, n^{o} 1, p. 61 - 83, Special issue : the best of IHM'2018, https://hal.archives-ouvertes.fr/hal-02158113

[15] A. CHALBI, J. RITCHIE, D. PARK, J. CHOI, N. ROUSSEL, N. ELMQVIST, F. CHEVALIER.*Common Fate for Animated Transitions in Visualization*, in "IEEE Transactions on Visualization and Computer Graphics", 2019, forthcoming [*DOI :* 10.1109/TVCG.2019.2934288], https://hal.inria.fr/hal-02308930

[16] S. MAALEJ, A. KRUSZEWSKI, L. BELKOURA.*Derivative-Based Sampled Data Control for Continuous Linear Parameter Varying System With Unknown Parameters*, in "Journal of Dynamic Systems, Measurement, and Control", August 2019, vol. 141, n^{o} 8 [*DOI :* 10.1115/1.4042947], https://hal.archives-ouvertes.fr/hal-02403902

[17] J. D. MATHEW, S. HUOT, B. F. G. KATZ.*Comparison of Spatial and Temporal Interaction Techniques for 3D Audio Trajectory Authoring*, in "Journal on Multimodal User Interfaces", November 2019 [*DOI :* 10.1007/S12193-019-00314-X], https://hal.inria.fr/hal-02375493

International Conferences with Proceedings

[18] A. ANTOINE, S. MALACRIA, N. MARQUARDT, G. CASIEZ.*Esquisse: Using 3D Models Staging to Facilitate the Creation of Vector-based Trace Figures*, in "INTERACT 2019 - The 17th IFIP TC.13 International Conference on Human-Computer Interaction", Paphos, Cyprus, September 2019, p. 496-

516, https://youtu.be/X2UdHXrvUg0 [*DOI :* 10.1007/978-3-030-29384-0_30], https://hal.inria.fr/hal-02293837

[19] M. BALOUP, T. PIETRZAK, G. CASIEZ.*RayCursor: a 3D Pointing Facilitation Technique based on Raycasting*, in "CHI 2019 - Proceedings of the ACM Conference on Human Factors in Computing Systems", Glasgow, United Kingdom, 2019 [*DOI :* 10.1145/3290605.3300331], https://hal.inria.fr/hal-02015844

[20] J. HENDERSON, J. AVERY, L. GRISONI, E. LANK.*Leveraging Distal Vibrotactile Feedback for Target Acquisition*, in "ACM CHI 2019 Conference", Glasgow, United Kingdom, ACM Press, May 2019, p. 1-11 [*DOI :* 10.1145/3290605.3300715], https://hal.archives-ouvertes.fr/hal-02436742

[21] N. K. C. PONG, S. MALACRIA.*Awareness, Usage and Discovery of Swipe-revealed Hidden Widgets in iOS*, in "ISS 2019 - Proceedings of the ACM Conference on Interactive Surfaces and Spaces", Daejeon, South Korea, November 2019, https://hal.inria.fr/hal-02293836

[22] *Best Paper*
T. RAFFAILLAC, S. HUOT.*Polyphony: Programming Interfaces and Interactions with the Entity-Component-System Model*, in "EICS 2019 - 11th ACM SIGCHI Symposium on Engineering Interactive Computing Systems", Valencia, Spain, June 2019, vol. 3 [*DOI :* 10.1145/3331150], https://hal.inria.fr/hal-02147180.

[23] Q. ROY, C. ZAKARIA, S. PERRAULT, M. NANCEL, W. KIM, A. MISRA, A. COCKBURN.*A Comparative Study of Pointing Techniques for Eyewear Using a Simulated Pedestrian Environment*, in "INTERACT 2019 - Proceedings of the IFIP TC13 International Conference on Human-Computer Interaction", Paphos, Cyprus, September 2019 [*DOI :* 10.31219/OSF.IO/KTUY4], https://hal.archives-ouvertes.fr/hal-02272539

[24] H. SI-MOHAMMED, G. CASIEZ, F. ARGELAGUET SANZ, N. ROUSSEL, A. LÉCUYER.*Defining Brain-Computer Interfaces: A Human-Computer Interaction Perspective*, in "GBCIC 2019 - 8th Graz Brain-Computer Interface Conference", Graz, Austria, September 2019, p. 1-4, https://hal.inria.fr/hal-02163771

National Conferences with Proceeding

[25] *Best Paper*
A. GOGUEY, S. MALACRIA, A. COCKBURN, C. GUTWIN.*Reducing Error Aversion to Support Novice-to-Expert Transitions with FastTap*, in "Actes de la 31e conférence francophone sur l'Interaction Homme-Machine (IHM 2019)", Grenoble, France, ACM, 2019, p. 1:1-10 [*DOI :* 10.1145/3366550.3372247], https://hal.archives-ouvertes.fr/hal-02381584.

[26] E. LANK, P. MEYER, D. MASSON, H. RATEAU.*Measuring the Effect of Public vs Private Context on the Kinematics of Smartphone Motion Gestures*, in "Actes de la 31e conférence francophone sur l'Interaction Homme-Machine (IHM 2019)", Grenoble, France, ACM, 2019, p. 14:1-6 [*DOI :* 10.1145/3366550.3372260], https://hal.archives-ouvertes.fr/hal-02381598

Scientific Books (or Scientific Book chapters)

[27] J. MALLOCH, J. GARCIA, M. M. WANDERLEY, W. MACKAY, M. BEAUDOUIN-LAFON, S. HUOT.*A Design Workbench for Interactive Music Systems*, in "New Directions in Music and Human-Computer Interaction", S. HOLLAND, T. MUDD, K. WILKIE-MCKENNA, A. MCPHERSON, M. M. WANDERLEY (editors), Springer

Series on Cultural Computing, Springer, February 2019, p. 23-40 [*DOI :* 10.1007/978-3-319-92069-6_2], https://hal.inria.fr/hal-02056992

Other Publications

[28] R. BLANCH, C. COUTRIX, T. PIETRZAK. , J. COUTAZ, S. DUPUY-CHESSA, S. KUBICKI (editors)*Actes de la 31e conférence francophone sur l'Interaction Homme-Machine*, December 2019, Avant-propos des actes de la 31e conférence francophone sur l'Interaction Homme-Machine, https://hal.archives-ouvertes.fr/hal-02390557

[29] G. RICHARD.*Study of the impact of haptic feedback on the sense of embodiment*, Arts et Métiers ParisTech, Paris, September 2019, https://hal.inria.fr/hal-02382924

[30] P. SCHMID.*A simulation platform to study interaction temporality*, University of Lille, September 2019, https://hal.inria.fr/hal-02393133

[31] C. THIERRY, Y. LE GALL, G. CASIEZ, J.-C. DUBOIS, S. MALACRIA, A. MARTIN, T. PIETRZAK.*Interface for collecting answers to equivocal questions for crowdsourcing*, June 2019, Workshop Visualization & HCI for Crowd-Sourcing, https://hal.inria.fr/hal-02402853

References in notes

[32] E. ALAWOE, T. PIETRZAK, S. HUOT.*Outil de sélection de texte manuscrit sur des documents numérisés*, in "SIFED 2018 - Symposium International Francophone sur l'Ecrit et le Document", Tours, France, May 2018, https://hal.archives-ouvertes.fr/hal-01899373

[33] M. BEAUDOUIN-LAFON.*Designing interaction, not interfaces*, in "Proceedings of AVI'04", ACM, 2004, p. 15-22, http://doi.acm.org/10.1145/989863.989865

[34] G. CASIEZ, N. ROUSSEL.*No more bricolage! Methods and tools to characterize, replicate and compare pointing transfer functions*, in "Proceedings of UIST'11", ACM, October 2011, p. 603-614, http://dx.doi.org/10.1145/2047196.2047276

[35] D. C. ENGELBART.*Augmenting human intellect: a conceptual framework*, Stanford Research Institute, October 1962, n^o AFOSR-3233, http://www.dougengelbart.org/pubs/augment-3906.html

[36] N. FELLION, T. PIETRZAK, A. GIROUARD.*FlexStylus: leveraging flexion input for pen interaction*, in "Proceedings of UIST'17", ACM, October 2017, p. 375-385

[37] S. GREENBERG.*Toolkits and Interface Creativity*, in "Multimedia Tools Appl.", February 2007, vol. 32, n^o 2, p. 139–159, http://dx.doi.org/10.1007/s11042-006-0062-y

[38] J. GRUDIN.*A moving target: the evolution of Human-Computer Interaction*, in "The Human Computer Interaction handbook (3rd edition)", J. A. JACKO (editor), CRC Press, May 2012, p. xxvii-lxi, http://research.microsoft.com/en-us/um/people/jgrudin/publications/history/HCIhandbook3rd.pdf

[39] S. HUOT.*Designeering interaction: a missing link in the evolution of Human-Computer Interaction*, Université Paris-Sud, France, May 2013, 205 pages, Habilitation à diriger des recherches, https://hal.inria.fr/tel-00823763

[40] R. JOTA, A. NG, P. DIETZ, D. WIGDOR.*How fast is fast enough? A study of the effects of latency in direct-touch pointing tasks*, in "Proceedings of CHI'13", ACM, April 2013, p. 2291-2300, http://doi.acm.org/10.1145/2470654.2481317

[41] I. S. MACKENZIE, C. WARE.*Lag as a determinant of human performance in interactive systems*, in "Proceedings of CHI'93", ACM, April 1993, p. 488-493, http://doi.acm.org/10.1145/169059.169431

[42] W. E. MACKAY.*Responding to cognitive overload: coadaptation between users and technology*, in "Intellectica", 2000, vol. 30, n^{o} 1, p. 177-193, ARCo, http://intellectica.org/SiteArchives/archives/n30/30_06_Mackay.pdf

[43] N. MARQUARDT, S. GREENBERG.*Proxemic Interactions: From Theory to Practice*, Synthesis Lectures on Human-Centered Informatics, Morgan & Claypool, 2015, https://books.google.fr/books?id=2dPtBgAAQBAJ

[44] J. MATEJKA, W. LI, T. GROSSMAN, G. FITZMAURICE.*CommunityCommands: command recommendations for software applications*, in "Proceedings of UIST'09", ACM, October 2009, p. 193-202, http://dx.doi.org/10.1145/1622176.1622214

[45] M. NANCEL, E. PIETRIGA, O. CHAPUIS, M. BEAUDOUIN-LAFON.*Mid-air pointing on ultra-walls*, in "ACM ToCHI", October 2015, vol. 22, n^{o} 5, http://dx.doi.org/10.1145/2766448

[46] M. PANTIC, A. PENTLAND, A. NIJHOLT, T. S. HUANG.*Human computing and machine understanding of human behavior: a survey*, in "Artifical intelligence for human computing", T. S. HUANG, A. NIJHOLT, M. PANTIC, A. PENTLAND (editors), LNCS, Springer, 2007, vol. 4451, p. 47-71, http://dx.doi.org/10.1007/978-3-540-72348-6_3

[47] V. RAMESH, C. HSU, M. AGRAWALA, B. HARTMANN.*ShowMeHow: Translating User Interface Instructions Between Applications*, in "Proceedings of UIST'11", ACM, October 2011, p. 127-134, http://doi.acm.org/10.1145/2047196.2047212

[48] J. SCARR, A. COCKBURN, C. GUTWIN, P. QUINN.*Dips and Ceilings: understanding and Supporting Transitions to Expertise in User Interfaces*, in "Proceedings of CHI'11", ACM, May 2011, p. 2741-2750, http://doi.acm.org/10.1145/1978942.1979348

[49] T. WINOGRAD, F. FLORES.*Understanding computers and cognition: a new foundation for design*, Addison-Wesley, May 1986, https://dl.acm.org/doi/book/10.5555/5245

Project-Team MAGNET

Machine Learning in Information Networks

IN COLLABORATION WITH: Centre de Recherche en Informatique, Signal et Automatique de Lille

IN PARTNERSHIP WITH:
CNRS

Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Data and Knowledge Representation and Processing

Table of contents

Project-Team MAGNET

Creation of the Team: 2013 January 01, updated into Project-Team: 2016 May 01

Keywords:

Computer Science and Digital Science:

A3.1. - Data
A3.1.3. - Distributed data
A3.1.4. - Uncertain data
A3.4. - Machine learning and statistics
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.4. - Optimization and learning
A3.5. - Social networks
A3.5.1. - Analysis of large graphs
A3.5.2. - Recommendation systems
A4.8. - Privacy-enhancing technologies
A9.4. - Natural language processing

Other Research Topics and Application Domains:

B1. - Life sciences
B1.1.10. - Systems and synthetic biology
B2. - Health
B2.2.4. - Infectious diseases, Virology
B2.3. - Epidemiology
B2.4.1. - Pharmaco kinetics and dynamics
B2.4.2. - Drug resistance
B5.10. - Biotechnology
B6.3. - Network functions
B7.1.2. - Road traffic
B8.3. - Urbanism and urban planning
B9.5.1. - Computer science
B9.5.4. - Chemistry
B9.5.6. - Data science
B9.6.8. - Linguistics
B9.6.10. - Digital humanities
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

Aurelien Bellet [Inria, Researcher]
Pascal Denis [Inria, Researcher]
Jan Ramon [Inria, Senior Researcher]

Faculty Members

Marc Tommasi [Team leader, Université de Lille, Professor, HDR]
Mikaela Keller [Université de Lille, Associate Professor]
Fabio Vitale [Université de Lille, Associate Professor]

External Collaborator

Remi Gilleron [Université de Lille, HDR]

Technical Staff

William de Vazelhes [Inria, Engineer, until Aug 2019]
Pradipta Deb [Inria, Engineer, from Oct 2019]
Carlos Zubiaga Pena [Inria, Engineer, until Jan 2019]

PhD Students

Mahsa Asadi [Inria]
Moitree Basu [Inria, from Oct 2019]
Mathieu Dehouck [Université de Lille, from Feb 2019 until Jul 2019]
Onkar Pandit [Inria]
Arijus Pleska [Inria]
César Sabater [Inria]
Brij Mohan Lal Srivastava [Inria]
Mariana Vargas Vieyra [Inria]

Post-Doctoral Fellow

Mohamed Maouche [Inria, from Nov 2019]

Administrative Assistant

Julie Jonas [Inria]

2. Overall Objectives

2.1. Presentation

MAGNET is a research group that aims to design new machine learning based methods geared towards mining information networks. Information networks are large collections of interconnected data and documents like citation networks and blog networks among others. Our goal is to propose new prediction methods for texts and networks of texts based on machine learning algorithms in graphs. Such algorithms include node and link classification, link prediction, clustering and probabilistic modeling of graphs. We aim to tackle real-world problems such as browsing, monitoring and recommender systems, and more broadly information extraction in information networks. Application domains cover natural language processing, social networks for cultural data and e-commerce, and biomedical informatics.

3. Research Program

3.1. Introduction

The main objective of MAGNET is to develop original machine learning methods for networked data in order to build applications like browsing, monitoring and recommender systems, and more broadly information extraction in information networks. We consider information networks in which the data consist of both feature vectors and texts. We model such networks as (multiple) (hyper)graphs wherein nodes correspond to entities (documents, spans of text, users, ...) and edges correspond to relations between entities (similarity, answer, co-authoring, friendship, ...). Our main research goal is to propose new online and batch learning algorithms for various problems (node classification / clustering, link classification / prediction) which exploit the relationships between data entities and, overall, the graph topology. We are also interested in searching for the best hidden graph structure to be generated for solving a given learning task. Our research will be based on generative models for graphs, on machine learning for graphs and on machine learning for texts. The challenges are the dimensionality of the input space, possibly the dimensionality of the output space, the high level of dependencies between the data, the inherent ambiguity of textual data and the limited amount of human labeling. An additional challenge will be to design scalable methods for large information networks. Hence, we will explore how sampling, randomization and active learning can be leveraged to improve the scalability of the proposed algorithms.

Our research program is organized according to the following questions:

1. How to go beyond vectorial classification models in Natural Language Processing (NLP) tasks?
2. How to adaptively build graphs with respect to the given tasks? How to create networks from observations of information diffusion processes?
3. How to design methods able to achieve a good trade-off between predictive accuracy and computational complexity?
4. How to go beyond strict node homophilic/similarity assumptions in graph-based learning methods?

3.2. Beyond Vectorial Models for NLP

One of our overall research objectives is to derive graph-based machine learning algorithms for natural language and text information extraction tasks. This section discusses the motivations behind the use of graph-based ML approaches for these tasks, the main challenges associated with it, as well as some concrete projects. Some of the challenges go beyond NLP problems and will be further developed in the next sections. An interesting aspect of the project is that we anticipate some important cross-fertilizations between NLP and ML graph-based techniques, with NLP not only benefiting from but also pushing ML graph-based approaches into new directions.

Motivations for resorting to graph-based algorithms for texts are at least threefold. First, online texts are organized in networks. With the advent of the web, and the development of forums, blogs, and micro-blogging, and other forms of social media, text productions have become strongly connected. Interestingly, NLP research has been rather slow in coming to terms with this situation, and most of the literature still focus on document-based or sentence-based predictions (wherein inter-document or inter-sentence structure is not exploited). Furthermore, several multi-document tasks exist in NLP (such as multi-document summarization and cross-document coreference resolution), but most existing work typically ignore document boundaries and simply apply a document-based approach, therefore failing to take advantage of the multi-document dimension [38], [41].

A second motivation comes from the fact that most (if not all) NLP problems can be naturally conceived as graph problems. Thus, NLP tasks often involve discovering a relational structure over a set of text spans (words, phrases, clauses, sentences, etc.). Furthermore, the *input* of numerous NLP tasks is also a graph; indeed, most end-to-end NLP systems are conceived as pipelines wherein the output of one processor is in the input of the next. For instance, several tasks take POS tagged sequences or dependency trees as input. But this structured input is often converted to a vectorial form, which inevitably involves a loss of information.

Finally, graph-based representations and learning methods appear to address some core problems faced by NLP, such as the fact that textual data are typically not independent and identically distributed, they often live on a manifold, they involve very high dimensionality, and their annotations is costly and scarce. As such, graph-based methods represent an interesting alternative to, or at least complement, structured prediction methods (such as CRFs or structured SVMs) commonly used within NLP. Graph-based methods, like label propagation, have also been shown to be very effective in semi-supervised settings, and have already given some positive results on a few NLP tasks [21], [43].

Given the above motivations, our first line of research will be to investigate how one can leverage an underlying network structure (e.g., hyperlinks, user links) between documents, or text spans in general, to enhance prediction performance for several NLP tasks. We think that a “network effect”, similar to the one that took place in Information Retrieval (with the PageRank algorithm), could also positively impact NLP research. A few recent papers have already opened the way, for instance in attempting to exploit Twitter follower graph to improve sentiment classification [42].

Part of the challenge here will be to investigate how adequately and efficiently one can model these problems as instances of more general graph-based problems, such as node clustering/classification or link prediction discussed in the next sections. In a few cases, like text classification or sentiment analysis, graph modeling appears to be straightforward: nodes correspond to texts (and potentially users), and edges are given by relationships like hyperlinks, co-authorship, friendship, or thread membership. Unfortunately, modeling NLP problems as networks is not always that obvious. From the one hand, the right level of representation will probably vary depending on the task at hand: the nodes will be sentences, phrases, words, etc. From the other hand, the underlying graph will typically not be given a priori, which in turn raises the question of how we construct it. A preliminary discussion of the issue of optimal graph construction for semi-supervised learning in NLP is given in [21], [46]. We identify the issue of adaptive graph construction as an important scientific challenge for machine learning on graphs in general, and we will discuss it further in Section 3.3.

As noted above, many NLP tasks have been recast as structured prediction problems, allowing to capture (some of the) output dependencies. How to best combine structured output and graph-based ML approaches is another challenge that we intend to address. We will initially investigate this question within a semi-supervised context, concentrating on graph regularization and graph propagation methods. Within such approaches, labels are typically binary or in a small finite set. Our objective is to explore how one propagates an exponential number of *structured labels* (like a sequence of tags or a dependency tree) through graphs. Recent attempts at blending structured output models with graph-based models are investigated in [43], [31]. Another related question that we will address in this context is how does one learn with *partial labels* (like partially specified tag sequence or tree) and use the graph structure to complete the output structure. This last question is very relevant to NLP problems where human annotations are costly; being able to learn from partial annotations could therefore allow for more targeted annotations and in turn reduced costs [33].

The NLP tasks we will mostly focus on are coreference resolution and entity linking, temporal structure prediction, and discourse parsing. These tasks will be envisioned in both document and cross-document settings, although we expect to exploit inter-document links either way. Choices for these particular tasks is guided by the fact that they are still open problems for the NLP community, they potentially have a high impact for industrial applications (like information retrieval, question answering, etc.), and we already have some expertise on these tasks in the team (see for instance [32], [28], [30]). As a midterm goal, we also plan to work on tasks more directly relating to micro-blogging, such as sentiment analysis and the automatic thread structuring of technical forums; the latter task is in fact an instance of rhetorical structure prediction [45]. We have already initiated some work on the coreference resolution with graph-based learning, by casting the problem as an instance of spectral clustering [30].

3.3. Adaptive Graph Construction

In most applications, edge weights are computed through a complex data modeling process and convey crucially important information for classifying nodes, making it possible to infer information related to each data sample even exploiting the graph topology solely. In fact, a widespread approach to several classification problems is to represent the data through an undirected weighted graph in which edge weights quantify the similarity between data points. This technique for coding input data has been applied to several domains, including classification of genomic data [40], face recognition [29], and text categorization [34].

In some cases, the full adjacency matrix is generated by employing suitable similarity functions chosen through a deep understanding of the problem structure. For example for the TF-IDF representation of documents, the affinity between pairs of samples is often estimated through the cosine measure or the χ^2 distance. After the generation of the full adjacency matrix, the second phase for obtaining the final graph consists in an edge sparsification/reweighting operation. Some of the edges of the clique obtained in the first step are pruned and the remaining ones can be reweighted to meet the specific requirements of the given classification problem. Constructing a graph with these methods obviously entails various kinds of loss of information. However, in problems like node classification, the use of graphs generated from several datasets can lead to an improvement in accuracy ([47], [22], [23]). Hence, the transformation of a dataset into a graph may, at least in some cases, partially remove various kinds of irregularities present in the original datasets, while keeping some of the most useful information for classifying the data samples. Moreover, it is often possible to accomplish classification tasks on the obtained graph using a running time remarkably lower than is needed by algorithms exploiting the initial datasets, and a suitable sparse graph representation can be seen as a compressed version of the original data. This holds even when input data are provided in a online/stream fashion, so that the resulting graph evolves over time.

In this project we will address the problem of adaptive graph construction towards several directions. The first one is about how to choose the best similarity measure given the objective learning task. This question is related to the question of metric and similarity learning ([24], [25]) which has not been considered in the context of graph-based learning. In the context of structured prediction, we will develop approaches where output structures are organized in graphs whose similarity is given by top-k outcomes of greedy algorithms.

A different way we envision adaptive graph construction is in the context of semi-supervised learning. Partial supervision can take various forms and an interesting and original setting is governed by two currently studied applications: detection of brain anomaly from connectome data and polls recommendation in marketing. Indeed, for these two applications, a partial knowledge of the information diffusion process can be observed while the network is unknown or only partially known. An objective is to construct (or complete) the network structure from some local diffusion information. The problem can be formalized as a graph construction problem from partially observed diffusion processes. It has been studied very recently in [36]. In our case, the originality comes either from the existence of different sources of observations or from the large impact of node contents in the network.

We will study how to combine graphs defined by networked data and graphs built from flat data to solve a given task. This is of major importance for information networks because, as said above, we will have to deal with multiple relations between entities (texts, spans of texts, ...) and also use textual data and vectorial data.

3.4. Prediction on Graphs and Scalability

As stated in the previous sections, graphs as complex objects provide a rich representation of data. Often enough the data is only partially available and the graph representation is very helpful in predicting the unobserved elements. We are interested in problems where the complete structure of the graph needs to be recovered and only a fraction of the links is observed. The link prediction problem falls into this category. We are also interested in the recommendation and link classification problems which can be seen as graphs where the structure is complete but some labels on the links (weights or signs) are missing. Finally we are also interested in labeling the nodes of the graph, with class or cluster memberships or with a real value, provided that we have (some information about) the labels for some of the nodes.

The semi-supervised framework will be also considered. A midterm research plan is to study how graph regularization models help for structured prediction problems. This question will be studied in the context of NLP tasks, as noted in Section 3.2, but we also plan to develop original machine learning algorithms that have a more general applicability. Inputs are networks whose nodes (texts) have to be labeled by structures. We assume that structures lie in some manifold and we want to study how labels can propagate in the network. One approach is to find a smooth labeling function corresponding to an harmonic function on both manifolds in input and output.

Scalability is one of the main issues in the design of new prediction algorithms working on networked data. It has gained more and more importance in recent years, because of the growing size of the most popular networked data that are now used by millions of people. In such contexts, learning algorithms whose computational complexity scales quadratically, or slower, in the number of considered data objects (usually nodes or edges, depending on the task) should be considered impractical.

These observations lead to the idea of using graph sparsification techniques in order to work on a part of the original network for getting results that can be easily extended and used for the whole original input. A sparsified version of the original graph can often be seen as a subset of the initial input, i.e. a suitably selected input subgraph which forms the training set (or, more in general, it is included in the training set). This holds even for the active setting. A simple example could be to find a spanning tree of the input graph, possibly using randomization techniques, with properties such that we are allowed to obtain interesting results for the initial graph dataset. We have started to explore this research direction for instance in [44].

At the level of mathematical foundations, the key issue to be addressed in the study of (large-scale) random networks also concerns the segmentation of network data into sets of independent and identically distributed observations. If we identify the data sample with the whole network, as it has been done in previous approaches [35], we typically end up with a set of observations (such as nodes or edges) which are highly interdependent and hence overly violate the classic i.i.d. assumption. In this case, the data scale can be so large and the range of correlations can be so wide, that the cost of taking into account the whole data and their dependencies is typically prohibitive. On the contrary, if we focus instead on a set of subgraphs independently drawn from a (virtually infinite) target network, we come up with a set of independent and identically distributed observations—namely the subgraphs themselves, where subgraph sampling is the underlying ergodic process [26]. Such an approach is one principled direction for giving novel statistical foundations to random network modeling. At the same time, because one shifts the focus from the whole network to a set of subgraphs, complexity issues can be restricted to the number of subgraphs and their size. The latter quantities can be controlled much more easily than the overall network size and dependence relationships, thus allowing to tackle scalability challenges through a radically redesigned approach.

Another way to tackle scalability problems is to exploit the inherent decentralized nature of very large graphs. Indeed, in many situations very large graphs are the abstract view of the digital activities of a very large set of users equipped with their own device. Nowadays, smartphones, tablets and even sensors have storage and computation power and gather a lot of data that serve to analytics, prediction, suggestion and personalized recommendation. Gathering all user data in large data centers is costly because it requires oversized infrastructures with huge energy consumption and large bandwidth networks. Even though cloud architectures can optimize such infrastructures, data concentration is also prone to security leaks, lost of privacy and data governance for end users. The alternative we have started to develop in Magnet is to devise decentralized, private and personalized machine learning algorithms so that they can be deployed in the personal devices. The key challenges are therefore to learn in a collaborative way in a network of learners and to preserve privacy and control on personal data.

3.5. Beyond Homophilic Relationships

In many cases, algorithms for solving node classification problems are driven by the following assumption: linked entities tend to be assigned to the same class. This assumption, in the context of social networks, is known as homophily ([27], [37]) and involves ties of every type, including friendship, work, marriage, age, gender, and so on. In social networks, homophily naturally implies that a set of individuals can be parted into subpopulations that are more cohesive. In fact, the presence of homogeneous groups sharing common interests is a key reason for affinity among interconnected individuals, which suggests that, in spite of its simplicity, this principle turns out to be very powerful for node classification problems in general networks.

Recently, however, researchers have started to consider networked data where connections may also carry a negative meaning. For instance, disapproval or distrust in social networks, negative endorsements on the Web. Although the introduction of signs on graph edges appears like a small change from standard weighted graphs, the resulting mathematical model, called signed graphs, has an unexpectedly rich additional complexity. For example, their spectral properties, which essentially all sophisticated node classification algorithms rely on, are different and less known than those of graphs. Signed graphs naturally lead to a specific inference problem that we have discussed in previous sections: link classification. This is the problem of predicting signs of links in a given graph. In online social networks, this may be viewed as a form of sentiment analysis, since we would like to semantically categorize the relationships between individuals.

Another way to go beyond homophily between entities will be studied using our recent model of hypergraphs with bipartite hyperedges [39]. A bipartite hyperedge connects two ends which are disjoint subsets of nodes. Bipartite hyperedges is a way to relate two collections of (possibly heterogeneous) entities represented by nodes. In the NLP setting, while hyperedges can be used to model bags of words, bipartite hyperedges are associated with relationships between bags of words. But each end of bipartite hyperedges is also a way to represent complex entities, gathering several attribute values (nodes) into hyperedges viewed as records. Our hypergraph notion naturally extends directed and undirected weighted graph. We have defined a spectral theory for this new class of hypergraphs and opened a way to smooth labeling on sets of nodes. The weighting scheme allows to weigh the participation of each node to the relationship modeled by bipartite hyperedges accordingly to an equilibrium condition. This condition provides a competition between nodes in hyperedges and allows interesting modeling properties that go beyond homophily and similarity over nodes (the theoretical analysis of our hypergraphs exhibits tight relationships with signed graphs). Following this competition idea, bipartite hyperedges are like matches between two teams and examples of applications are team creation. The basic tasks we are interested in are hyperedge classification, hyperedge prediction, node weight prediction. Finally, hypergraphs also represent a way to summarize or compress large graphs in which there exists highly connected couples of (large) subsets of nodes.

4. Application Domains

4.1. Application Domains

Our main targeted applications are browsing, monitoring, recommending and mining in information networks. The learning tasks considered in the project such as node clustering, node and link classification and link prediction are likely to yield important improvements in these applications. Application domains cover social networks for cultural data and e-commerce, and biomedical informatics.

We also target applications related to decentralized learning and privacy preserving systems when users or devices are interconnected in large networks. We develop solutions based on urban and mobility data where privacy is a specific requirement.

5. Highlights of the Year

5.1. Highlights of the Year

- Metric-Learn software has been included in the scikit-learn-contrib packages. It records more than 900 stars and 190 forks on GitHub. It is also used by 51 projects.
- AURÉLIEN BELLET has applied for a ERC Starting Grant on privacy-preserving decentralized machine learning.
- MATHIEU DEHOUCK has successfully defended his PhD dissertation on *Multi-Lingual Dependency Parsing: Word Representation and Joint Training for Syntactic Analysis*, and he is doing a post-doc at University of A Coruña (Spain) funded by ERC grant FASTPARSE.
- MARIANA VARGAS VIEYRA's work on probabilistic end-to-end graph-based semi-supervised learning was accepted as one of the 8 contributed talks (among 92 accepted submissions) as the NeurIPS'19 workshop on Graph Representation Learning [0].

6. New Software and Platforms

6.1. CoRTeX

Python library for noun phrase COreference Resolution in natural language TEXts

[0] https://grlearning.github.io/papers/

KEYWORD: Natural language processing

FUNCTIONAL DESCRIPTION: CoRTex is a LGPL-licensed Python library for Noun Phrase coreference resolution in natural language texts. This library contains implementations of various state-of-the-art coreference resolution algorithms, including those developed in our research. In addition, it provides a set of APIs and utilities for text pre-processing, reading the CONLL2012 and CONLLU annotation formats, and performing evaluation, notably based on the main evaluation metrics (MUC, B-CUBED, and CEAF). As such, CoRTex provides benchmarks for researchers working on coreference resolution, but it is also of interest for developers who want to integrate a coreference resolution within a larger platform. It currently supports use of the English or French language.

- Participant: Pascal Denis
- Partner: Orange Labs
- Contact: Pascal Denis
- URL: https://gitlab.inria.fr/magnet/CoRTeX

6.2. Mangoes

MAgnet liNGuistic wOrd vEctorS

KEYWORDS: Word embeddings - NLP

FUNCTIONAL DESCRIPTION: Process textual data and compute vocabularies and co-occurrence matrices. Input data should be raw text or annotated text. Compute word embeddings with different state-of-the art unsupervised methods. Propose statistical and intrinsic evaluation methods, as well as some visualization tools.

- Contact: Nathalie Vauquier
- URL: https://gitlab.inria.fr/magnet/mangoes

6.3. metric-learn

KEYWORDS: Machine learning - Python - Metric learning

FUNCTIONAL DESCRIPTION: Distance metrics are widely used in the machine learning literature. Traditionally, practicioners would choose a standard distance metric (Euclidean, City-Block, Cosine, etc.) using a priori knowledge of the domain. Distance metric learning (or simply, metric learning) is the sub-field of machine learning dedicated to automatically constructing optimal distance metrics.

This package contains efficient Python implementations of several popular metric learning algorithms.

- Partner: Parietal
- Contact: Aurélien Bellet
- URL: https://github.com/scikit-learn-contrib/metric-learn

6.4. MyLocalInfo

KEYWORDS: Privacy - Machine learning - Statistics

FUNCTIONAL DESCRIPTION: Decentralized algorithms for machine learning and inference tasks which (1) perform as much computation as possible locally and (2) ensure privacy and security by avoiding personal data leaves devices.

- Contact: Nathalie Vauquier
- URL: https://gitlab.inria.fr/magnet/mylocalinfo

7. New Results

7.1. Natural Language Processing

Multi-Lingual Dependency Parsing

In [1], MATHIEU DEHOUCK presents his work on Word Representation and Joint Training for Syntactic Analysis. Syntactic analysis is a key step in working with natural languages. With the advances in supervised machine learning, modern parsers have reached human performances. However, despite the intensive efforts of the dependency parsing community, the number of languages for which data have been annotated is still below the hundred, and only a handful of languages have more than ten thousands annotated sentences. In order to alleviate the lack of training data and to make dependency parsing available for more languages, previous research has proposed methods for sharing syntactic information across languages. By transferring models and/or annotations or by jointly learning to parse several languages at once, one can capitalise on languages grammatical similarities in order to improve their parsing capabilities. However, while words are a key source of information for mono-lingual parsers, they are much harder to use in multi-lingual settings because they vary heavily even between very close languages. Morphological features on the contrary, are much more stable across related languages than word forms and they also directly encode syntactic information. Furthermore, it is arguably easier to annotate data with morphological information than with complete dependency structures. With the increasing availability of morphologically annotated data using the same annotation scheme for many languages, it becomes possible to use morphological information to bridge the gap between languages in multi-lingual dependency parsing.

In his thesis, MATHIEU DEHOUCK has proposed several new approaches for sharing information across languages. These approaches have in common that they rely on morphology as the adequate representation level for sharing information. We therefore also introduce a new method to analyse the role of morphology in dependency parsing relying on a new measure of morpho-syntactic complexity. The first method uses morphological information from several languages to learn delexicalised word representations that can then be used as feature and improve mono-lingual parser performances as a kind of distant supervision. The second method uses morphology as a common representation space for sharing information during the joint training of model parameters for many languages. The training process is guided by the evolutionary tree of the various language families in order to share information between languages historically related that might share common grammatical traits. We empirically compare this new training method to independently trained models using data from the Universal Dependencies project and show that it greatly helps languages with few resources but that it is also beneficial for better resourced languages when their family tree is well populated. We eventually investigate the intrinsic worth of morphological information in dependency parsing. Indeed not all languages use morphology as extensively and while some use morphology to mark syntactic relations (via cases and persons) other mostly encode semantic information (such as tense or gender). To this end, we introduce a new measure of morpho-syntactic complexity that measures the syntactic content of morphology in a given corpus as a function of preferential head attachment. We show through experiments that this new measure can tease morpho-syntactic languages and morpho-semantic languages apart and that it is more predictive of parsing results than more traditional morphological complexity measures.

Modal sense classification with task-specific context embeddings Sense disambiguation of modal constructions is a crucial part of natural language understanding. Framed as a supervised learning task, this problem heavily depends on an adequate feature representation of the modal verb context. Inspired by recent work on general word sense disambiguation, we propose in [8] a simple approach of modal sense classification in which standard shallow features are enhanced with task-specific context embedding features. Comprehensive experiments show that these enriched contextual representations fed into a simple SVM model lead to significant classification gains over shallow feature sets.

Learning Rich Event Representations and Interactions for Temporal Relation Classification Most existing systems for identifying temporal relations between events heavily rely on hand-crafted features derived from event words and explicit temporal markers. Besides, less attention has been given to automatically learning con-textualized event representations or to finding complex interactions between events. In [9], we fill this gap in showing that a combination of rich event representations and interaction learning is essential to more accurate temporal relation classification. Specifically, we propose a method in which i) Recurrent Neural Networks (RNN) extract contextual information ii) character embeddings capture morpho-semantic features (e.g. tense, mood, aspect), and iii) a deep Convolutional Neural Network (CNN) finds out intricate interactions between events. We show that the proposed approach outperforms most existing systems on the commonly used dataset while using fully automatic feature extraction and simple local inference.

Phylogenetic Multi-Lingual Dependency Parsing Languages evolve and diverge over time. Their evolutionary history is often depicted in the shape of a phylogenetic tree. Assuming parsing models are representations of their languages grammars, their evolution should follow a structure similar to that of the phylo-genetic tree. In [7], drawing inspiration from multi-task learning, we make use of the phylogenetic tree to guide the learning of multilingual dependency parsers leverag-ing languages structural similarities. Experiments on data from the Universal Dependency project show that phylogenetic training is beneficial to low resourced languages and to well furnished languages families. As a side product of phylogenetic training, our model is able to perform zero-shot parsing of previously unseen languages.

7.2. Decentralized Learning

Trade-offs in Large-Scale Distributed Tuplewise Estimation and Learning The development of cluster computing frameworks has allowed practitioners to scale out various statistical estimation and machine learning algorithms with minimal programming effort. This is especially true for machine learning problems whose objective function is nicely separable across individual data points, such as classification and regression. In contrast, statistical learning tasks involving pairs (or more generally tuples) of data points-such as metric learning, clustering or ranking-do not lend themselves as easily to data-parallelism and in-memory computing. In [13], we investigate how to balance between statistical performance and computational efficiency in such distributed tuplewise statistical problems. We first propose a simple strategy based on occasionally repartitioning data across workers between parallel computation stages, where the number of repartition-ing steps rules the trade-off between accuracy and runtime. We then present some theoretical results highlighting the benefits brought by the proposed method in terms of variance reduction, and extend our results to design distributed stochastic gradient descent algorithms for tuplewise empirical risk minimization. Our results are supported by numerical experiments in pairwise statistical estimation and learning on synthetic and real-world datasets.

Who started this rumor? Quantifying the natural differential privacy guarantees of gossip protocols Gossip protocols, also called rumor spreading or epidemic protocols, are widely used to disseminate information in massive peer-to-peer networks. These protocols are often claimed to guarantee privacy because of the uncertainty they introduce on the node that started the dissemination. But is that claim really true? Can one indeed start a gossip and safely hide in the crowd? In [14], we study gossip protocols using a rigorous mathematical framework based on differential privacy to determine the extent to which the source of a gossip can be traceable. Considering the case of a complete graph in which a subset of the nodes are curious, we derive matching lower and upper bounds on differential privacy showing that some gossip protocols achieve strong privacy guarantees. Our results further reveal an interesting tension between privacy and dissemination speed: the standard "push" gossip protocol has very weak privacy guarantees, while the optimal guarantees are attained at the cost of a drastic increase in the spreading time. Yet, we show that it is possible to leverage the inherent randomness and partial observability of gossip protocols to achieve both fast dissemination speed and near-optimal privacy.

Fully Decentralized Joint Learning of Personalized Models and Collaboration Graphs In [15], we consider the fully decentralized machine learning scenario where many users with personal datasets collaborate to learn models through local peer-to-peer exchanges, without a central coordinator. We propose to train personalized models that leverage a collaboration graph describing the relationships between the users' personal tasks, which we learn jointly with the models. Our fully decentralized optimization procedure alternates between training nonlinear models given the graph in a greedy boosting manner, and updating the collaboration graph (with controlled sparsity) given the models. Throughout the process, users exchange messages only with a small number of peers (their direct neighbors in the graph and a few random users), ensuring that the procedure naturally scales to large numbers of users. We analyze the convergence rate, memory and communication complexity of our approach, and demonstrate its benefits compared to competing techniques on synthetic and real datasets.

Advances and Open Problems in Federated Learning Federated learning (FL) is a machine learning setting where many clients (e.g. mobile devices or whole organizations) collaboratively train a model under the orchestration of a central server (e.g. service provider), while keeping the training data decentralized. FL embodies the principles of focused data collection and minimization, and can mitigate many of the systemic privacy risks and costs resulting from traditional, centralized machine learning and data science approaches. Motivated by the explosive growth in FL research, we participated in a collaborative paper [18] that discusses recent advances and presents an extensive collection of open problems and challenges.

7.3. Privacy and Machine Learning

Private Protocols for U-Statistics in the Local Model and Beyond In [16], we study the problem of computing U-statistics of degree 2, i.e., quantities that come in the form of averages over pairs of data points, in the local model of differential privacy (LDP). The class of U-statistics covers many statistical estimates of interest, including Gini mean difference, Kendall's tau coefficient and Area under the ROC Curve (AUC), as well as empirical risk measures for machine learning problems such as ranking, clustering and metric learning. We first introduce an LDP protocol based on quantizing the data into bins and applying randomized response, which guarantees an ϵ-LDP estimate with a Mean Squared Error (MSE) of $O(1/\sqrt{n}\epsilon)$ under regularity assumptions on the U-statistic or the data distribution. We then propose a specialized protocol for AUC based on a novel use of hierarchical histograms that achieves MSE of $O(\alpha^3/n\epsilon^2)$ for arbitrary data distribution. We also show that 2-party secure computation allows to design a protocol with MSE of $O(1/n\epsilon^2)$, without any assumption on the kernel function or data distribution and with total communication linear in the number of users n. Finally, we evaluate the performance of our protocols through experiments on synthetic and real datasets.

Privacy-Preserving Adversarial Representation Learning in ASR: Reality or Illusion? In [11], we study Automatic Speech Recognition (ASR), a key technology in many services and applications. This typically requires user devices to send their speech data to the cloud for ASR decoding. As the speech signal carries a lot of information about the speaker, this raises serious privacy concerns. As a solution, an encoder may reside on each user device which performs local computations to anonymize the representation. In this paper, we focus on the protection of speaker identity and study the extent to which users can be recognized based on the encoded representation of their speech as obtained by a deep encoder-decoder architecture trained for ASR. Through speaker identification and verification experiments on the Librispeech corpus with open and closed sets of speakers, we show that the representations obtained from a standard architecture still carry a lot of information about speaker identity. We then propose to use adversarial training to learn representations that perform well in ASR while hiding speaker identity. Our results demonstrate that adversarial training dramatically reduces the closed-set classification accuracy, but this does not translate into increased open-set verification error hence into increased protection of the speaker identity in practice. We suggest several possible reasons behind this negative result.

Evaluating Voice Conversion-based Privacy Protection against Informed Attackers Speech signals are a rich source of speaker-related information including sensitive attributes like identity or accent. With a small amount of found speech data, such attributes can be extracted and modeled for malicious purposes like voice cloning, spoofing, etc. In [19], we investigate speaker anonymization strategies based on voice conversion. In contrast to prior evaluations, we argue that different types of attackers can be defined depending on the extent of their knowledge about the conversion scheme. We compare two frequency warping-based conversion methods and a deep learning based method in three attack scenarios. The utility of the converted speech is measured through the word error rate achieved by automatic speech recognition, while privacy protection is assessed by state-of-the-art speaker verification techniques (i-vectors and x-vectors). Our results show that voice conversion schemes are unable to effectively protect against an attacker that has extensive knowledge of the type of conversion and how it has been applied, but may provide some protection against less knowledgeable attackers.

7.4. Learning in Graphs

Correlation Clustering with Adaptive Similarity Queries In correlation clustering, we are given n objects together with a binary similarity score between each pair of them. The goal is to partition the objects into clusters so to minimise the disagreements with the scores. In [6], we investigate correlation clustering as an active learning problem: each similarity score can be learned by making a query, and the goal is to minimise both the disagreements and the total number of queries. On the one hand, we describe simple active learning algorithms, which provably achieve an almost optimal trade-off while giving cluster recovery guarantees, and we test them on different datasets. On the other hand, we prove information-theoretical bounds on the number of queries necessary to guarantee a prescribed disagreement bound. These results give a rich characterization of the trade-off between queries and clustering error.

Flattening a Hierarchical Clustering through Active Learning In [12], we investigate active learning by pairwise similarity over the leaves of trees originating from hierarchical clustering procedures. In the realizable setting, we provide a full characterization of the number of queries needed to achieve perfect reconstruction of the tree cut. In the non-realizable setting, we rely on known important-sampling procedures to obtain regret and query complexity bounds. Our algorithms come with theoretical guarantees on the statistical error and, more importantly, lend themselves to linear-time implementations in the relevant parameters of the problem. We discuss such implementations, prove running time guarantees for them, and present preliminary experiments on real-world datasets showing the compelling practical performance of our algorithms as compared to both passive learning and simple active learning baselines.

MaxHedge: Maximising a Maximum Online In [10], we introduce a new online learning framework where, at each trial, the learner is required to select a subset of actions from a given known action set. Each action is associated with an energy value, a reward and a cost. The sum of the energies of the actions selected cannot exceed a given energy budget. The goal is to maximise the cumulative profit, where the profit obtained on a single trial is defined as the difference between the maximum reward among the selected actions and the sum of their costs. Action energy values and the budget are known and fixed. All rewards and costs associated with each action change over time and are revealed at each trial only after the learner's selection of actions. Our framework encompasses several online learning problems where the environment changes over time; and the solution trades-off between minimising the costs and maximising the maximum reward of the selected subset of actions, while being constrained to an action energy budget. The algorithm that we propose is efficient, general and may be specialised to multiple natural online combinatorial problems.

Closed-loop cycles of experiment design, execution, and learning accelerate systems biology model development in yeast One of the most challenging tasks in modern science is the development of systems biology models: Existing models are often very complex but generally have low predictive performance. The construction of high-fidelity models will require hundreds/thousands of cycles of model improvement, yet few current systems biology research studies complete even a single cycle. In [2], we combined multiple software tools with integrated laboratory robotics to execute three cycles of model improvement of the prototypical eukaryotic cellular transformation, the yeast (Saccharomyces cerevisiae) diauxic shift. In the first cycle, a model outperforming the best previous diauxic shift model was developed using bioinformatic and systems biology tools. In the second cycle, the model was further improved using automatically planned experiments. In the third cycle, hypothesis-led experiments improved the model to a greater extent than achieved using high-throughput experiments. All of the experiments were formalized and communicated to a cloud laboratory automation system (Eve) for automatic execution, and the results stored on the semantic web for reuse. The final model adds a substantial amount of knowledge about the yeast diauxic shift: 92 genes (+45%), and 1 048 interactions (+147%). This knowledge is also relevant to understanding cancer, the immune system, and aging. We conclude that systems biology software tools can be combined and integrated with laboratory robots in closed-loop cycles.

7.5. Metric Learning

Metric learning is at the core of many algorithms for learning graphs. A new software has been published in the scikit-learn contrib repository (See the Software section).

Escaping the Curse of Dimensionality in Similarity Learning: Efficient Frank-Wolfe Algorithm and Generalization Bounds Similarity and metric learning provides a principled approach to construct a task-specific similarity from weakly supervised data. However, these methods are subject to the curse of dimensionality: as the number of features grows large, poor generalization is to be expected and training becomes intractable due to high computational and memory costs. In [3], we propose a similarity learning method that can efficiently deal with high-dimensional sparse data. This is achieved through a parameterization of similarity functions by convex combinations of sparse rank-one matrices, together with the use of a greedy approximate Frank-Wolfe algorithm which provides an efficient way to control the number of active features. We show that the convergence rate of the algorithm, as well as its time and memory complexity, are independent of the data dimension. We further provide a theoretical justification of our modeling choices through an analysis of the generalization error, which depends logarithmically on the sparsity of the solution rather than on the number of features. Our experiments on datasets with up to one million features demonstrate the ability of our approach to generalize well despite the high dimensionality as well as its superiority compared to several competing methods.

metric-learn: Metric Learning Algorithms in Python In [20], we present metric-learn, an open source Python package implementing supervised and weakly-supervised distance metric learning algorithms. As part of scikit-learn-contrib, it provides a unified interface compatible with scikit-learn which allows to easily perform cross-validation, model selection, and pipelining with other machine learning estimators. metric-learn is thoroughly tested and available on PyPi under the MIT licence.

7.6. Graph Algorithms

We collaborate with the Links project team on graph-based computations and evaluation in databases.

Dependency Weighted Aggregation on Factorized Databases In [17], we study a new class of aggregation problems, called dependency weighted aggregation. The underlying idea is to aggregate the answer tuples of a query while accounting for dependencies between them, where two tuples are considered dependent when they have the same value on some attribute. The main problem we are interested in is to compute the dependency weighted count of a conjunctive query. This aggregate can be seen as a form of weighted counting, where the weights of the answer tuples are computed by solving a linear program. This linear program enforces that dependent tuples are not over represented in the final weighted count. The dependency weighted count can be used to compute the s-measure, a measure that is used in data mining to estimate the frequency of a pattern in a graph database. Computing the dependency weighted count of a conjunctive query is NP-hard in general. In this paper, we show that this problem is actually tractable for a large class of structurally restricted conjunctive queries such as acyclic or bounded hypertree width queries. Our algorithm works on a factorized representation of the answer set, in order to avoid enumerating it exhaustively. Our technique produces a succinct representation of the weighting of the answers. It can be used to solve other dependency weighted aggregation tasks, such as computing the (dependency) weighted average of the value of an attribute in the answers set.

7.7. Learning and Speech Recognition

We have worked on privacy and machine learning for speech recognition (See Section 7.3). Additional results concern kernel method for speech recognition.

Kernel Approximation Methods for Speech Recognition In [4], we study the performance of kernel methods on the acoustic modeling task for automatic speech recognition, and compare their performance to deep neural networks (DNNs). To scale the kernel methods to large data sets, we use the random Fourier feature method of Rahimi and Recht (2007). We propose two novel techniques for improving the performance of kernel acoustic models. First, we propose a simple but effective feature selection method which reduces the number of random features required to attain a fixed level of performance. Second, we present a number of metrics which correlate strongly with speech recognition performance when computed on the heldout set; we attain improved performance by using these metrics to decide when to stop training. Additionally, we show that the linear bottleneck method of Sainath et al. (2013a) improves the performance of our kernel models significantly, in addition to speeding up training and making the models more compact. Leveraging these three methods, the kernel methods attain token error rates between 0.5% better and 0.1% worse than fully-connected DNNs across four speech recognition data sets, including the TIMIT and Broadcast News benchmark tasks.

8. Partnerships and Cooperations

8.1. Regional Initiatives

We participate to the *Data Advanced data science and technologies* project (CPER Data). This project is organized following three axes: internet of things, data science, high performance computing. MAGNET is involved in the data science axis to develop machine learning algorithms for big data, structured data and heterogeneous data. The project MyLocalInfo is an open API for privacy-friendly collaborative computing in the internet of things.

MAGNET also has various collaborations with research groups in linguistics and psycholinguistics at Université de Lille, in particular UMR STL (with an ongoing joint ANR project) and UMR SCALab (co-supervision of students).

8.2. National Initiatives

8.2.1. ANR Pamela (2016-2020)

Participants: MARC TOMMASI [correspondent], AURÉLIEN BELLET, RÉMI GILLERON, JAN RAMON, MAHSA ASADI

The Pamela project aims at developing machine learning theories and algorithms in order to learn local and personalized models from data distributed over networked infrastructures. Our project seeks to provide first answers to modern information systems built by interconnecting many personal devices holding private user data in the search of personalized suggestions and recommendations. More precisely, we will focus on learning in a collaborative way with the help of neighbors in a network. We aim to lay the first blocks of a scientific foundation for these new types of systems, in effect moving from graphs of data to graphs of data and learned models. We argue that this shift is necessary in order to address the new constraints arising from the decentralization of information that is inherent to the emergence of big data. We will in particular focus on the question of learning under communication and privacy constraints. A significant asset of the project is the quality of its industrial partners, Snips and Mediego, who bring in their expertise in privacy protection and distributed computing as well as use cases and datasets. They will contribute to translate this fundamental research effort into concrete outcomes by developing personalized and privacy-aware assistants able to provide contextualized recommendations on small devices and smartphones.

https://project.inria.fr/pamela/

8.2.2. ANR JCJC GRASP (2016-2020)

Participants: Pascal Denis [correspondent], Aurélien Bellet, Rémi Gilleron, Mikaela Keller, Marc Tommasi

The GRASP project aims at designing new graph-based Machine Learning algorithms that are better tailored to Natural Language Processing structured output problems. Focusing on semi-supervised learning scenarios, we will extend current graph-based learning approaches along two main directions: (i) the use of structured outputs during inference, and (ii) a graph construction mechanism that is more dependent on the task objective and more closely related to label inference. Combined, these two research strands will provide an important step towards delivering more adaptive (to new domains and languages), more accurate, and ultimately more useful language technologies. We will target semantic and pragmatic tasks such as coreference resolution, temporal chronology prediction, and discourse parsing for which proper Machine Learning solutions are still lacking.

https://project.inria.fr/grasp/

8.2.3. ANR DEEP-Privacy (2019-2023)

Participants: Marc Tommasi [correspondent], Aurélien Bellet, Pascal Denis, Jan Ramon, Brij Srivastava

DEEP-PRIVACY proposes a new paradigm based on a distributed, personalized, and privacy-preserving approach for speech processing, with a focus on machine learning algorithms for speech recognition. To this end, we propose to rely on a hybrid approach: the device of each user does not share its raw speech data and runs some private computations locally, while some cross-user computations are done by communicating through a server (or a peer-to-peer network). To satisfy privacy requirements at the acoustic level, the information communicated to the server should not expose sensitive speaker information.

8.2.4. ANR-NFS REM (2016-2020)

Participants: Pascal Denis [correspondent], Bo Li, Mathieu Dehouck

With colleagues from the linguistics departments at Université de Lille and University of Neuchâtel (Switzerland), Pascal Denis is a member of another ANR project (REM), funded through the bilateral ANR-NFS Scheme. This project, co-headed by I. Depreatere (Université de Lille) and M. Hilpert (Neufchâtel), proposes to reconsider the analysis of English modal constructions from a multidisciplinary perspective, combining insights from theoretical, psycho-linguistic, and computational approaches.

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

Participants: Aurelien Bellet, Marc Tommasi, Brij Mohan Lal Srivastava.

Program: H2020 ICT-29-2018 (RIA)

Project acronym: COMPRISE

Project title: Cost-effective, Multilingual, Privacy-driven voice-enabled Services

Duration: Dec 2018 - Nov 2021

Coordinator: Emmanuel Vincent [Inria Nancy - Grand Est]

Other partners: Inria Multispeech, Ascora GmbH, Netfective Technology SA, Rooter Analysis SL, Tilde SIA, University of Saarland

Abstract: COMPRISE will define a fully private-by-design methodology and tools that will reduce the cost and increase the inclusiveness of voice interaction technologies.

8.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: Bilateral ANR project with Luxembourg

Project acronym: SLANT

Project title: Spin and Bias in Language Analyzed in News and Texts

Duration: Dec 2019 - June 2023

Coordinator: Philippe Muller [Université Paul Sabatier]

Other partners: IRIT (Toulouse), SnT (Luxembourg)

Abstract: There is a growing concern about misinformation or biased information in public communication, whether in traditional media or social forums. While automating fact-checking has received a lot of attention, the problem of fair information is much larger and includes more insidious forms like biased presentation of events and discussion. The SLANT project aims at characterizing bias in textual data, either intended, in public reporting, or unintended in writing aiming at neutrality. An abstract model of biased interpretation using work on discourse structure, semantics and interpretation will be complemented and concretized by finding relevant lexical, syntactic, stylistic or rhetorical differences through an automated but explainable comparison of texts with different biases on the same subject, based on a dataset of news media coverage from a diverse set of sources. We will also explore how our results can help alter bias in texts or remove it from automated representations of texts.

8.4. International Initiatives

8.4.1. Inria International Labs

Inria@SiliconValley
Associate Team involved in the International Lab:

8.4.1.1. LEGO

Title: LEarning GOod representations for natural language processing

International Partner (Institution - Laboratory - Researcher):

University of Southern California (United States) - Theoretical and Empirical Data Science (TEDS) research group Department of Computer Science - Fei Sha

Start year: 2019

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

LEGO lies in the intersection of Machine Learning and Natural Language Processing (NLP). Its goal is to address the following challenges: what are the right representations for text data and how to learn them in a robust and transferable way? How to apply such representations to solve real-world NLP tasks, specifically in scenarios where linguistic resources are scarce? The past years have seen an increasing interest in learning continuous vectorial embeddings, which can be trained together with the prediction model in an end-to-end fashion, as in recent sequence-to-sequence neural models. However, they are unsuitable to low-resource languages as they require massive amounts of data to train. They are also very prone to overfitting, which makes them very brittle, and sensitive to bias present in the original text as well as to confounding factors such as author attributes. LEGO strongly relies on the complementary expertise of the two partners in areas such as representation learning, structured prediction, graph-based learning, multi-task/transfer learning, and statistical NLP to offer a novel alternative to existing techniques. Specifically, we propose to investigate the following two research directions: (a) optimize the representations to make them robust to bias and adversarial examples, and (b) learn transferable representations across languages and domains, in particular in the context of structured prediction problems for low-resource languages. We will demonstrate the usefulness of the proposed methods on several NLP tasks, including multilingual dependency parsing, machine translation, question answering and text summarization.

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

North-European Associate Team PAD-ML: Privacy-Aware Distributed Machine Learning.

International Partner: the PPDA team at the Alan Turing Institute.

Start year: 2018

In the context of increasing legislation on data protection (e.g., the recent GDPR), an important challenge is to develop privacy-preserving algorithms to learn from datasets distributed across multiple data owners who do not want to share their data. The goal of this joint team is to devise novel privacy-preserving, distributed machine learning algorithms and to assess their performance and guarantees in both theoretical and practical terms.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

Several international researchers have been invited to give a talk at the MAGNET seminar:

- A. Korba (University College London, UK): Two families of (non-parametric) methods for label ranking
- M. Perrot (Max Planck Institute, Germany): Comparison-Based Learning: Hierarchical Clustering and Classification

8.5.2. Visits to International Teams

8.5.2.1. Research Stays Abroad

- FABIO VITALE was on leave at Department of Computer Science of Sapienza University (Rome, Italy) in the Algorithms Randomization Computation group with Prof. Alessandro Panconesi and Prof. Flavio Chierichetti. His current work on machine learning in graphs and published the following papers [6], [12], [10].
- AURÉLIEN BELLET and CÉSAR SABATER visited the Alan Turing Institute (London, UK) for one week in March 2019. They worked with Adrià Gascón, Brooks Paige, Daphne Ezer and Matt Kusner on privacy-preserving machine learning and privacy attacks in genomics.

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Organisation

9.1.1.1. Member of the Organizing Committees

- Marc Tommasi and Aurélien Bellet co-organized the APVP conference on privacy [0].

9.1.2. Scientific Events: Selection

9.1.2.1. Member of the Conference Program Committees

- Marc Tommasi served as PC member for ICML'19, CAP'19, IJCAI'19 (Senior PC chair), ECML'19 (Area Chair), NeurIPS'19 (Area Chair), APVP'19.
- Jan Ramon served as PC member for AAAI'19, AISTATS'19 & '20, APVP'19, Bigdata'19, CIKM'19, DS'19, ECML/PKDD'19, GEM@ECML'19, IEEE-ICDM'19, ICML'19, IJCAI'19, ILP'19, LEG@ECML'19, LOD'19, MLG'19, NeurIPS'19, SDM'19 & '20.
- Pascal Denis served as PC member for ACL'19 and EMNLP'19.
- Aurélien Bellet served as Area Chair for ICML'19, and PC member for AISTATS'20, NeurIPS'19, PPML workshop at CSS'19, PPML workshop at NeurIPS'19, FL workshop at NeurIPS'19, PPAI workshop at AAAI '20, CAP'19.
- Mikaela Keller served as PC member for ICML'19, CAP'19 and ECML'19.
- Rémi Gilleron served as PC member for AISTATS'19 & '20, ICLR'19 & '20, ECML/PKDD'19, NeurIPS'19 and CAP'19.

9.1.3. Journal

9.1.3.1. Member of the Editorial Boards

- Jan Ramon was member of the editorial board of Data mining and knowledge discovery (DMKD), and of the editorial board of Machine learning journal (MLJ)

9.1.3.2. Reviewer - Reviewing Activities

- Marc Tommasi was reviewer for the ECML-PKDD Journal Track.
- Aurélien Bellet was reviewer for SIAM Journal on Mathematics of Data Science.

9.1.4. Invited Talks

- Aurélien Bellet gave invited talks at several workshops and research school: "Graph signals: learning and optimization perspectives" [0], "The power of graphs in machine learning and sequential decision making" [0], "CIFAR-UKRI-CNRS Workshop on AI & Society: From principles to practice" [0], "Workshop on Federated Learning and Analytics" [0], "International Workshop on Machine Learning & Artificial Intelligence" [0], "Kick-off seminar of the HumAIn Alliance in Artificial Intelligence" [0], "Research School on Uncertainty in Scientific Computing" [0], "GDR RSD and ASF Winter School on Distributed Systems and Networks" [0].
- Aurélien Bellet was invited to talk at the seminars of Thales Research & Technology, Miles team at Paris Dauphine.
- Mikaela Keller gave an invited talk on "Algorithmiques de graphes pour l'étude de réseaux sociaux" at "École d'été internationale : Formation de maîtrise et de doctorat en méthodes et outils numériques : l'analyse de réseaux en histoire" [0].

[0] https://project.inria.fr/apvp2019/
[0] https://graph-sig-2019.sciencesconf.org/
[0] https://graphpower.inria.fr/
[0] https://www.turing.ac.uk/events/cifar-ukri-cnrs-ai-society-principles-practice
[0] https://sites.google.com/view/federated-learning-2019/
[0] https://workshopmlai.wp.imt.fr/
[0] https://www.alliance-humain.fr/
[0] http://www.gdr-mascotnum.fr/etics.html
[0] https://sites.google.com/site/rsdwinterschool/
[0] http://www.grhs.uqam.ca/?portfolio=ecole-dete-internationale-formation-de-maitrise-et-de-doctorat-methodes-et-outils-numeriques-lanalyse-de-reseaux-en-histoire

9.1.5. Scientific Expertise

- AURÉLIEN BELLET was a member of the jury for the Gilles-Kahn PhD award of the French Society of Computer Science (SIF), sponsored by the French Academy of Sciences [0].
- JAN RAMON was a member of the jury for the European Young Researchers Award (EYRA) of Euroscience.
- PASCAL DENIS served as CoCNRS representative for the evaluation panel of Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur, HCERES [0].
- PASCAL DENIS acted as external expert for Innovis [0], Brussels Institute for Research and Innovation (Belgium).

9.1.6. Research Administration

- MIKAELA KELLER is member of the Conseil du laboratoire CRIStAL.
- MARC TOMMASI is co-head of the DatInG group (4 teams, 80 persons) and member of the Conseil Scientifique du laboratoire CRIStAL.
- PASCAL DENIS served as an elected member of the Comité National du CNRS, section 34 (Sciences du Langage).
- PASCAL DENIS served as a member of the CNRS Pre-GDR NLP Group.
- JAN RAMON was a member of the Inria-Lille committee emploi-recherche (CER).

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence SCE: FABIO VITALE, Apprentissage et émergence des comportements, 30h, L2, Université de Lille.

Licence MIASHS: FABIO VITALE, Algorithmique et programmation, 42h, L3, Université de Lille.

Licence SoQ: FABIO VITALE, Bases de données et SQL, 16h, L1, Université de Lille.

Licence MIASHS: FABIO VITALE, Introduction aux bases de données, 18h, L3, Université de Lille.

Licence SHS: MARC TOMMASI, Data Science, 24h, L2, Université de Lille.

Licence MIASHS: MARC TOMMASI, Codage et représentation de l'information, 48h, L1, Université de Lille.

Licence: MARC TOMMASI, Traitement de textes et tableur, 28h Université de Lille.

Licence: MARC TOMMASI, Humanités numériques - Découvrir et faire découvrir la programmation, 20h, Université de Lille

Master SHS: MARC TOMMASI, IT Tools for professional translators, 24h, M2 Université de Lille.

Licence SHS: MARC TOMMASI, Python Programming, 48h, L2 Université de Lille.

Licence MIASHS: MIKAELA KELLER, Python II, 40h, L2, Université de Lille.

Licence MIASHS: MIKAELA KELLER, Traitement de données, 42h, L2, Université de Lille.

Licence SoQ (SHS): MIKAELA KELLER, Algorithmique de graphes, 24h, L3, Université de Lille.

Master MIASHS: MIKAELA KELLER, Algorithmes fondamentaux de la fouille de données, 60h, M1, Université de Lille.

Master Computer Science: MIKAELA KELLER, Machine Learning, 24h, M1, Université de Lille.

Master Data Science: FABIO VITALE, Algorithms and their complexity, 30h, M1, Ecole Centrale de Lille.

[0] https://www.societe-informatique-de-france.fr/recherche/prix-de-these-gilles-kahn/
[0] https://www.hceres.fr/en
[0] https://innoviris.brussels/

Master Data Analysis & Decision Making: AURÉLIEN BELLET, Machine Learning, 12h, Ecole Centrale de Lille.

Master / Master Spécialisé Big Data: AURÉLIEN BELLET, Advanced Machine Learning, 15h, Télécom ParisTech.

Formation continue (Certificat d'Études Spécialisées Data Scientist): AURÉLIEN BELLET, Supervised Learning and Support Vector Machines, 14h, Télécom ParisTech.

Master Informatique: PASCAL DENIS, Foundations of Machine Learning, 46h, M1, Université de Lille.

9.2.2. Supervision

Postdoc: MOHAMED MAOUCHE, supervised by AURÉLIEN BELLET, MARC TOMMASI, Privacy attacks on representation learning for speech processing, since November 2019.

PhD: MATHIEU DEHOUCK, Multi-lingual Dependency Parsing: Word Representation and Joint Training for Syntactic Analysis, 2015-2019, PASCAL DENIS and MARC TOMMASI.

PhD in progress: ONKAR PANDIT, Graph-based Semi-supervised Linguistic Structure Prediction, since Dec. 2017, PASCAL DENIS, MARC TOMMASI and LIVA RALAIVOLA (University of Marseille).

PhD in progress: MARIANA VARGAS VIEYRA, Adaptive Graph Learning with Applications to Natural Language Processing, since Jan. 2018. PASCAL DENIS and AURÉLIEN BELLET and MARC TOMMASI.

PhD in progress: BRIJ SRIVASTAVA, Representation Learning for Privacy-Preserving Speech Recognition, since Oct 2018 AURÉLIEN BELLET and MARC TOMMASI and EMMANUEL VINCENT.

PhD in progress: MAHSA ASADI, On Decentralized Machine Learning, since Oct 2018. AURÉLIEN BELLET and MARC TOMMASI.

PhD in progress: NICOLAS CROSETTI, Privacy Risks of Aggregates in Data Centric-Workflows, since Oct 2018. FLORENT CAPELLI and SOPHIE TISON and JOACHIM NIEHREN and JAN RAMON.

PhD in progress: ROBIN VOGEL, Learning to rank by similarity and performance optimization in biometric identification, since 2017 (CIFRE thesis with IDEMIA and Télécom ParisTech). AURÉLIEN BELLET, STÉPHAN CLÉMENÇON and ANNE SABOURIN.

PhD in progress: MOITREE BASU, Integrated privacy-preserving AI, since 2019. JAN RAMON.

PhD in progress: CÉSAR SABATER, Privacy Preserving Machine Learning, since 2019. JAN RAMON.

Master internship in progress: PAUL MANGOLD, supervised by AURÉLIEN BELLET and MARC TOMMASI, since October 2019.

9.2.3. Juries

MARC TOMMASI was member of the PhD jury of Corentin Hardy (rapporteur), Guillaume Metzler (rapporteur), Ronan Fruit (President)

MARC TOMMASI was member of the Habilitation jury of Rémi Eyraud (rapporteur).

MARC TOMMASI was member of the recruitment committee of Assistant Professors in Computer Science at École centrale de Lille.

AURÉLIEN BELLET was member of the PhD jury of Alexandre Garcia (Télécom Paris).

MIKAELA KELLER was member of the PhD jury of Guillaume Lample (LIP6)

MIKAELA KELLER was member of the recruitment committee of Assistant Professors in Computer Science at Sorbonne Université (LIP6).

RÉMI GILLERON was head of the PhD jury of Lily Galois (Université de Lille)

9.3. Popularization

9.3.1. Internal or external Inria responsibilities

AURÉLIEN BELLET is member of the Operational Committee for the assesment of Legal and Ethical risks (COERLE).

PASCAL DENIS is administrator of Inria membership to Linguistic Data Consortium (LDC).

9.3.2. Articles and contents

AURÉLIEN BELLET and MARC TOMMASI provided expertise for the Dopamine TV program on Arte about new technologies [0].

AURÉLIEN BELLET co-authored an article in the French daily newspaper Libération on AI and transparency [0].

RÉMI GILLERON authored a book (in French) introducing Machine Learning [0].

RÉMI GILLERON gave a talk in order to popularize Artificial Intelligence.

9.3.3. Interventions

In educational institutions: MARC TOMMASI was facilitator for the event *Jouer à débattre*, about machine learning in high school (Tourcoing).

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] M. DEHOUCK.*Multi-Lingual Dependency Parsing : Word Representation and Joint Training for Syntactic Analysis*, Université de lille, May 2019, https://tel.archives-ouvertes.fr/tel-02197615

Articles in International Peer-Reviewed Journal

[2] A. COUTANT, K. ROPER, D. TREJO-BANOS, D. BOUTHINON, M. CARPENTER, J. GRZEBYTA, G. SANTINI, H. SOLDANO, M. ELATI, J. RAMON, C. ROUVEIROL, L. N. SOLDATOVA, R. D. KING.*Closed-loop cycles of experiment design, execution, and learning accelerate systems biology model development in yeast*, in "Proceedings of the National Academy of Sciences of the United States of America ", 2019, vol. 116, n^{o} 36, p. 18142-18147 [*DOI :* 10.1073/PNAS.1900548116], https://hal.sorbonne-universite.fr/hal-02297702

[3] K. LIU, A. BELLET.*Escaping the Curse of Dimensionality in Similarity Learning: Efficient Frank-Wolfe Algorithm and Generalization Bounds*, in "Neurocomputing", 2019, vol. 333, p. 185-199, https://hal.inria.fr/hal-02166425

[4] A. MAY, A. BAGHERI GARAKANI, Z. LU, D. GUO, K. LIU, A. BELLET, L. FAN, M. COLLINS, D. HSU, B. KINGSBURY, M. PICHENY, F. SHA.*Kernel Approximation Methods for Speech Recognition*, in "Journal of Machine Learning Research", 2019, vol. 20, p. 1 - 36, https://hal.inria.fr/hal-02166422

[0] https://www.arte.tv/fr/videos/RC-017841/dopamine/
[0] https://www.liberation.fr/debats/2019/04/07/grand-debat-et-ia-quelle-transparence-pour-les-donnees_1719944
[0] https://www.editions-ellipses.fr/accueil/4741-apprentissage-machine-cle-de-l-intelligence-artificielle-une-introduction-pour-non-specialistes-9782340028807.html

International Conferences with Proceedings

[5] M. Asadi, M. S. Talebi, H. Bourel, O.-A. Maillard.*Model-Based Reinforcement Learning Exploiting State-Action Equivalence*, in "ACML 2019, Proceedings of Machine Learning Research", Nagoya, Japan, 2019, vol. 101, p. 204 - 219, https://hal.archives-ouvertes.fr/hal-02378887

[6] M. Bressan, N. Cesa-Bianchi, A. Paudice, F. Vitale.*Correlation Clustering with Adaptive Similarity Queries*, in "Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, https://arxiv.org/abs/1905.11902 , https://hal.inria.fr/hal-02376961

[7] M. Dehouck, P. Denis.*Phylogenetic Multi-Lingual Dependency Parsing*, in "NAACL 2019 - Annual Conference of the North American Chapter of the Association for Computational Linguistics", Minneapolis, United States, Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, June 2019, https://hal.archives-ouvertes.fr/hal-02143747

[8] B. Li, M. Dehouck, P. Denis.*Modal sense classification with task-specific context embeddings*, in "ESANN 2019 - 27th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning", Bruges, Belgium, April 2019, https://hal.archives-ouvertes.fr/hal-02143762

[9] O. Pandit, P. Denis, L. Ralaivola.*Learning Rich Event Representations and Interactions for Temporal Relation Classification*, in "ESANN 2019 - 27th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning", Bruges, Belgium, April 2019, https://hal.archives-ouvertes.fr/hal-02265061

[10] S. Pasteris, F. Vitale, K. Chan, S. Wang, M. Herbster.*MaxHedge: Maximising a Maximum Online*, in "International Conference on Artificial Intelligence and Statistics", Naha, Okinawa, Japan, April 2019, https://arxiv.org/abs/1810.11843 [*DOI :* 10.11843], https://hal.inria.fr/hal-02376987

[11] B. M. L. Srivastava, A. Bellet, M. Tommasi, E. Vincent.*Privacy-Preserving Adversarial Representation Learning in ASR: Reality or Illusion?*, in "INTERSPEECH 2019 - 20th Annual Conference of the International Speech Communication Association", Graz, Austria, September 2019, https://hal.inria.fr/hal-02166434

[12] F. Vitale, A. Rajagopalan, C. Gentile.*Flattening a Hierarchical Clustering through Active Learning*, in "Conference on Neural Information Processing Systems", Vancouver, Canada, December 2019, https://hal.inria.fr/hal-02376981

[13] R. Vogel, A. Bellet, S. Clémençon, O. Jelassi, G. Papa.*Trade-offs in Large-Scale Distributed Tuplewise Estimation and Learning*, in "ECML PKDD 2019 - European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases", Würzburg, Germany, September 2019, https://hal.inria.fr/hal-02166428

Research Reports

[14] A. Bellet, R. Guerraoui, H. Hendrikx.*Who started this rumor? Quantifying the natural differential privacy guarantees of gossip protocols*, Inria, 2019, https://hal.inria.fr/hal-02166432

[15] V. Zantedeschi, A. Bellet, M. Tommasi.*Fully Decentralized Joint Learning of Personalized Models and Collaboration Graphs*, Inria, 2019, https://hal.inria.fr/hal-02166433

Other Publications

[16] J. Bell, A. Bellet, A. Gascón, T. Kulkarni.*Private Protocols for U-Statistics in the Local Model and Beyond*, October 2019, https://arxiv.org/abs/1910.03861 - working paper or preprint [*DOI :* 10.03861], https://hal.inria.fr/hal-02310236

[17] F. Capelli, N. Crosetti, J. Niehren, J. Ramon.*Dependency Weighted Aggregation on Factorized Databases*, January 2019, https://arxiv.org/abs/1901.03633 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01981553

[18] P. Kairouz, H. B. McMahan, B. Avent, A. Bellet, M. Bennis, A. N. Bhagoji, K. Bonawitz, Z. Charles, G. Cormode, R. Cummings, R. G. L. D'Oliveira, S. E. Rouayheb, D. Evans, J. Gardner, Z. Garrett, A. Gascón, B. Ghazi, P. B. Gibbons, M. Gruteser, Z. Harchaoui, C. He, L. He, Z. Huo, B. Hutchinson, J. Hsu, M. Jaggi, T. Javidi, G. Joshi, M. Khodak, J. Konečný, A. Korolova, F. Koushanfar, S. Koyejo, T. Lepoint, Y. Liu, P. Mittal, M. Mohri, R. Nock, A. Ozgür, R. Pagh, M. Raykova, H. Qi, D. Ramage, R. Raskar, D. Song, W. Song, S. U. Stich, Z. Sun, A. T. Suresh, F. Tramèr, P. Vepakomma, J. Wang, L. Xiong, Z. Xu, Q. Yang, F. X. Yu, H. Yu, S. Zhao.*Advances and Open Problems in Federated Learning*, December 2019, https://arxiv.org/abs/1912.04977 - working paper or preprint, https://hal.inria.fr/hal-02406503

[19] B. M. L. Srivastava, N. Vauquier, M. Sahidullah, A. Bellet, M. Tommasi, E. Vincent.*Evaluating Voice Conversion-based Privacy Protection against Informed Attackers*, November 2019, working paper or preprint, https://hal.inria.fr/hal-02355115

[20] W. de Vazelhes, C. Carey, Y. Tang, N. Vauquier, A. Bellet.*metric-learn: Metric Learning Algorithms in Python*, November 2019, https://arxiv.org/abs/1908.04710 - GitHub repository: https://github.com/scikit-learn-contrib/metric-learn, https://hal.inria.fr/hal-02376986

References in notes

[21] A. Alexandrescu, K. Kirchhoff.*Graph-based learning for phonetic classification*, in "IEEE Workshop on Automatic Speech Recognition & Understanding, ASRU 2007, Kyoto, Japan, December 9-13, 2007", 2007, p. 359-364

[22] M.-F. Balcan, A. Blum, P. P. Choi, J. Lafferty, B. Pantano, M. R. Rwebangira, X. Zhu.*Person Identification in Webcam Images: An Application of Semi-Supervised Learning*, in "ICML2005 Workshop on Learning with Partially Classified Training Data", 2005

[23] M. Belkin, P. Niyogi.*Towards a Theoretical Foundation for Laplacian-Based Manifold Methods*, in "Journal of Computer and System Sciences", 2008, vol. 74, n^o 8, p. 1289-1308

[24] A. Bellet, A. Habrard, M. Sebban.*A Survey on Metric Learning for Feature Vectors and Structured Data*, in "CoRR", 2013, vol. abs/1306.6709

[25] A. Bellet, A. Habrard, M. Sebban.*Metric Learning*, Morgan & Claypool Publishers, 2015

[26] P. J. BICKEL, A. CHEN.*A nonparametric view of network models and Newman–Girvan and other modularities*, in "Proceedings of the National Academy of Sciences", 2009, vol. 106, p. 21068–21073

[27] P. BLAU.*Inequality and Heterogeneity: A Primitive Theory of Social Structure*, MACMILLAN Company, 1977, http://books.google.fr/books?id=jvq2AAAAIAAJ

[28] C. BRAUD, P. DENIS.*Combining Natural and Artificial Examples to Improve Implicit Discourse Relation Identification*, in "coling", Dublin, Ireland, August 2014, https://hal.inria.fr/hal-01017151

[29] H. CHANG, D.-Y. YEUNG.*Graph Laplacian Kernels for Object Classification from a Single Example*, in "Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2", Washington, DC, USA, CVPR '06, IEEE Computer Society, 2006, p. 2011–2016, http://dx.doi.org/10.1109/CVPR.2006.128

[30] D. CHATEL, P. DENIS, M. TOMMASI.*Fast Gaussian Pairwise Constrained Spectral Clustering*, in "ECML/PKDD - 7th European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases", Nancy, France, September 2014, p. 242 - 257 [*DOI :* 10.1007/978-3-662-44848-9_16], https://hal.inria.fr/hal-01017269

[31] D. DAS, S. PETROV.*Unsupervised Part-of-Speech Tagging with Bilingual Graph-Based Projections*, in "ACL", 2011, p. 600-609

[32] P. DENIS, P. MULLER.*Predicting globally-coherent temporal structures from texts via endpoint inference and graph decomposition*, in "IJCAI-11 - International Joint Conference on Artificial Intelligence", Barcelone, Espagne, 2011, http://hal.inria.fr/inria-00614765

[33] E. R. FERNANDES, U. BREFELD.*Learning from Partially Annotated Sequences*, in "ECML/PKDD", 2011, p. 407-422

[34] A. B. GOLDBERG, X. ZHU.*Seeing stars when there aren't many stars: graph-based semi-supervised learning for sentiment categorization*, in "Proceedings of the First Workshop on Graph Based Methods for Natural Language Processing", Stroudsburg, PA, USA, TextGraphs-1, Association for Computational Linguistics, 2006, p. 45–52, http://dl.acm.org/citation.cfm?id=1654758.1654769

[35] A. GOLDENBERG, A. X. ZHENG, S. E. FIENBERG.*A Survey of Statistical Network Models*, Foundations and trends in machine learning, Now Publishers, 2010, http://books.google.fr/books?id=gPGgcOf95moC

[36] M. GOMEZ-RODRIGUEZ, J. LESKOVEC, A. KRAUSE.*Inferring networks of diffusion and influence*, in "Proc. of KDD", 2010, p. 1019-1028

[37] M. MCPHERSON, L. S. LOVIN, J. M. COOK.*Birds of a Feather: Homophily in Social Networks*, in "Annual Review of Sociology", 2001, vol. 27, n^{o} 1, p. 415–444, http://dx.doi.org/10.1146/annurev.soc.27.1.415

[38] A. NENKOVA, K. MCKEOWN.*A Survey of Text Summarization Techniques*, in "Mining Text Data", Springer, 2012, p. 43-76

[39] T. RICATTE, R. GILLERON, M. TOMMASI.*Hypernode Graphs for Spectral Learning on Binary Relations over Sets*, in "ECML/PKDD - 7th European Conference on Machine Learning and Principles and Practice

of Knowledge Discovery in Databases", Nancy, France, Machine Learning and Knowledge Discovery in Databases, September 2014, https://hal.inria.fr/hal-01017025

[40] H. SHIN, K. TSUDA, B. SCHÖLKOPF.*Protein functional class prediction with a combined graph*, in "Expert Syst. Appl.", March 2009, vol. 36, n^o 2, p. 3284–3292, http://dx.doi.org/10.1016/j.eswa.2008.01.006

[41] S. SINGH, A. SUBRAMANYA, F. C. N. PEREIRA, A. MCCALLUM.*Large-Scale Cross-Document Coreference Using Distributed Inference and Hierarchical Models*, in "ACL", 2011, p. 793-803

[42] M. SPERIOSU, N. SUDAN, S. UPADHYAY, J. BALDRIDGE.*Twitter Polarity Classification with Label Propagation over Lexical Links and the Follower Graph*, in "Proceedings of the First Workshop on Unsupervised Methods in NLP", Edinburgh, Scotland, 2011

[43] A. SUBRAMANYA, S. PETROV, F. C. N. PEREIRA.*Efficient Graph-Based Semi-Supervised Learning of Structured Tagging Models*, in "EMNLP", 2010, p. 167-176

[44] F. VITALE, N. CESA-BIANCHI, C. GENTILE, G. ZAPPELLA.*See the Tree Through the Lines: The Shazoo Algorithm*, in "Proc of NIPS", 2011, p. 1584-1592

[45] L. WANG, S. N. KIM, T. BALDWIN.*The Utility of Discourse Structure in Identifying Resolved Threads in Technical User Forums*, in "COLING", 2012, p. 2739-2756

[46] K. K. YUZONG LIU.*Graph-Based Semi-Supervised Learning for Phone and Segment Classification*, in "Proceedings of Interspeech", Lyon, France, 2013

[47] X. ZHU, Z. GHAHRAMANI, J. LAFFERTY.*Semi-Supervised Learning Using Gaussian Fields and Harmonic Functions*, in "Proc. of ICML", 2003, p. 912-919

Team MEPHYSTO

Quantitative methods for stochastic models in physics

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
Lille - Nord Europe

THEME
Numerical schemes and simulations

Table of contents

Team MEPHYSTO

Creation of the Team: 2017 October 01

Keywords:

Computer Science and Digital Science:

A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.2. - Stochastic Modeling
A6.1.4. - Multiscale modeling
A6.2.1. - Numerical analysis of PDE and ODE

Other Research Topics and Application Domains:

B9.5.2. - Mathematics

1. Team, Visitors, External Collaborators

Research Scientists

Guillaume Dujardin [Team leader, Inria, Researcher, HDR]
Marielle Simon [Inria, Researcher]

Faculty Members

Stephan de Bièvre [Université de Lille, Professor, HDR]
André de Laire [Université de Lille, Associate Professor]
Adrien Hardy [Université de Lille, Associate Professor]
Michele Triestino [Université de Bourgogne, Associate Professor, until Feb 2019]

Technical Staff

Alexandre Roget [Inria, Engineer, from Nov 2019]

PhD Students

Pierre Mennuni [Université de Lille, PhD Student, until Sep 2019]
Anthony Nahas [Université de Lille, PhD Student, from Oct 2019]

Administrative Assistant

Karine Lewandowski [Inria]

2. Overall Objectives

2.1. Overall Objectives

The MEPHYSTO team is a follow up of the MEPHYSTO project-team. Since the former scientific leader, Antoine Gloria, left in September 2017, the scientific objectives have been modified.

The MEPHYSTO team gathers mathematicians from different communities with the same motivation: to provide a better understanding of dynamical phenomena involving particles. These phenomena are described by fundamental models arising from several fields of physics. We focus on model derivation, study of stationary states and asymptotic behaviors, as well as links between different levels of description (e.g. micro and macro models) and numerical methods to simulate such models. Applications include nonlinear optics, thermodynamics and ferromagnetism.

3. Research Program

3.1. Time asymptotics: Stationary states, solitons, and stability issues

The team investigates the existence of solitons and their link with the global dynamical behavior for nonlocal problems such as that of the Gross–Pitaevskii (GP) equation which arises in models of dipolar gases. These models, in general, also introduce nonzero boundary conditions which constitute an additional theoretical and numerical challenge. Numerous results are proved for local problems, and numerical simulations allow to verify and illustrate them, as well as making a link with physics. However, most fundamental questions are still open at the moment for nonlocal problems.

The nonlinear Schrödinger (NLS) equation finds applications in numerous fields of physics. We concentrate, in a continued collaboration with our colleagues from the physics department (PhLAM) of the Université de Lille (UdL), in the framework of the Laboratoire d'Excellence CEMPI, on its applications in nonlinear optics and cold atom physics. Issues of orbital stability and modulational instability are central here.

Another typical example of problems that the team wishes to address concerns the Landau–Lifshitz (LL) equation, which describes the dynamics of the spin in ferromagnetic materials. This equation is a fundamental model in the magnetic recording industry [37] and solitons in magnetic media are of particular interest as a mechanism for data storage or information transfer [38]. It is a quasilinear PDE involving a function that takes values on the unit sphere $\mathbb{S}^2$ of $\mathbb{R}^3$. Using the stereographic projection, it can be seen as a quasilinear Schrödinger equation and the questions about the solitons, their dynamics and potential blow-up of solutions evoked above are also relevant in this context. This equation is less understood than the NLS equation: even the Cauchy theory is not completely done [36], [35]. In particular, the geometry of the target sphere imposes nonvanishing boundary conditions; even in dimension one, there are kink-type solitons having different limits at $\pm\infty$.

3.2. Derivation of macroscopic laws from microscopic dynamics

The team investigates, from a microscopic viewpoint, the dynamical mechanism at play in the phenomenon of relaxation towards thermal equilibrium for large systems of interacting particles. For instance, a first step consists in giving a rigorous proof of the fact that a particle repeatedly scattered by random obstacles through a Hamiltonian scattering process will eventually reach thermal equilibrium, thereby completing previous work in this direction by the team. As a second step, similar models as the ones considered classically will be defined and analysed in the quantum mechanical setting, and more particularly in the setting of quantum optics.

Another challenging problem is to understand the interaction of large systems with the boundaries, which is responsible for most energy exchanges (forcing and dissipation), even though it is concentrated in very thin layers. The presence of boundary conditions to evolution equations sometimes lacks understanding from a physical and mathematical point of view. In order to legitimate the choice done at the macroscopic level of the mathematical definition of the boundary conditions, we investigate systems of atoms (precisely chains of oscillators) with different local microscopic defects. We apply our recent techniques to understand how anomalous (in particular fractional) diffusive systems interact with the boundaries. For instance, the powerful tool given by Wigner functions that we already used has been successfully applied to the derivation of anomalous behaviors in open systems (for instance in [7]). The next step consists in developing an extension of that tool to deal with bounded systems provided with fixed boundaries. We also intend to derive anomalous diffusion by adding long range interactions to diffusive models. There are very few rigorous results in this direction. Finally, we aim at obtaining from a microscopic description the fractional porous medium equation (FPME), a nonlinear variation of the fractional diffusion equation, involving the fractional Laplacian instead of the usual one. Its rigorous study carries out many mathematical difficulties in treating at the same time the nonlinearity and fractional diffusion. We want to make PDE theorists and probabilists work together, in order to take advantage of the analytical results which went far ahead and are more advanced than the statistical physics theory.

3.3. Numerical methods: analysis and simulations

The team addresses both questions of precision and numerical cost of the schemes for the numerical integration of nonlinear evolution PDEs, such as the NLS equation. In particular, we aim at developing, studying and implementing numerical schemes with high order that are more efficient for these problems. We also want to contribute to the design and analysis of schemes with appropriate qualitative properties. These properties may as well be "asymptotic preserving" properties, energy-preserving properties, or convergence to an equilibrium properties. Other numerical goals of the team include the numerical simulation of standing waves of nonlinear nonlocal GP equations. We also keep on developing numerical methods to efficiently simulate and illustrate theoretical results on instability, in particular in the context of the modulational instability in optical fibers, where we study the influence of randomness in the physical parameters of the fibers.

The team also designs simulation methods to estimate the accuracy of the physical description via microscopic systems, by computing precisely the rate of convergence as the system size goes to infinity. One method under investigation is related to cloning algorithms, which were introduced very recently and turn out to be essential in molecular simulation.

4. Highlights of the Year

4.1. Highlights of the Year

The team has almost completed the process of creation of a new project-team named Paradyse (for PARticles And DYnamical SystEms), between Inria and the Laboratoire Paul Painlevé of the Université de Lille in 2019.

In 2019, the Mephysto team has been granted an Action de Développement Technologique (ADT) by Inria. This allowed the team to hire Alexandre Roget as an engineer for 2 years. The goal of this ADT is to develop software using mathematical techniques developed in the team, to be used by theoretical and experimental physics communities.

In 2019, the team also had individual successes that can be highlighted. Amongst others, M. Simon submitted an ERC Starting Grant project which was ranked A, and S. De Bièvre became Associate Editor of the Journal of Mathematical Physics.

5. New Results

5.1. Traveling waves for some nonlocal 1D Gross-Pitaevskii equations with nonzero conditions at infinity

The nonlocal Gross-Pitaevskii equation is a model that appears naturally in several areas of quantum physics, for instance in the description of superfluids and in optics when dealing with thermo-optic materials because the thermal nonlinearity is usually highly nonlocal. A. de Laire and P. Mennuni have considered a nonlocal family of Gross–Pitaevskii equations in dimension one, and they have provided in [27] conditions on the nonlocal interaction such that there is a branch of traveling waves solutions with nonvanishing conditions at infinity. Moreover, they showed that the branch is orbitally stable. In this manner, this result generalizes known properties for the contact interaction given by a Dirac delta function. Their proof relies on the minimization of the energy at fixed momentum.

5.2. Numerical simulation of traveling waves for some nonlocal Gross-Pitaevskii equations with nonzero conditions at infinity in dimensions 1 and 2

As a follow-up of the previous result, P. Mennuni and G. Dujardin carried out numerical simulations of traveling waves for some nonlocal nonlinear Gross-Pitaevskii equations with nonzero conditions at infinity in dimensions 1 and 2. Using a numerical analogue of the minimization of the energy at fixed momentum, they used gradient methods with nonuniform fast Fourier transforms (to deal with the nonlocal terms numerically) to carry out significant numerical simulations to illustrate numerically the theoretical results and to discuss the hypotheses numerically. These results can be found in P. Mennuni's PhD manuscript [10].

5.3. The cubic Schrödinger regime of the Landau-Lifshitz equation with a strong easy-axis anisotropy

It is well-known that the dynamics of biaxial ferromagnets with a strong easy-axis anisotropy is essentially governed by the cubic Schrödinger equation. A. de Laire and P. Gravejat provided in [26] a rigorous justification to this observation. More precisely, they showed the convergence of the solutions to the Landau-Lifshitz equation for biaxial ferromagnets towards the solutions to the cubic Schrödinger equation in the regime of an easy-axis anisotropy. This result holds for solutions to the Landau–Lifshitz equation in high order Sobolev spaces. By introducing high order energy quantities with good symmetrization properties, they derived the convergence from the consistency of the Landau–Lifshitz equation with the Sine-Gordon equation by using well-tailored energy estimates.

In this regime, they additionally classified the one-dimensional solitons of the Landau-Lifshitz equation and quantified their convergence towards the solitons of the one-dimensional cubic Schrödinger equation.

5.4. The Cauchy problem for the Landau–Lifshitz–Gilbert equation in BMO and self-similar solutions

A. de Laire and S. Gutierrez established in [22] a global well-posedness result for the Landau–Lifshitz equation with Gilbert damping, provided that the BMO semi-norm of the initial data is small. As a consequence, they deduced the existence of self-similar solutions in any dimension. Moreover, in the one-dimensional case, they characterized the self-similar solutions when the initial data is given by some step function and established their stability. They also showed the existence of multiple solutions if the damping is strong enough.

5.5. Microscopic derivation of moving interfaces problems

In [15], M. Simon and her coauthors derive the porous medium equation from an interacting particle system which belongs to the family of kinetically constrained lattice gases. It was already proved in the literature that the macroscopic density profile is governed by the porous medium equation for initial densities uniformly bounded away from 0 and 1. Here we consider the more general case where the density can take those extreme values. The solutions display a richer behavior, like moving interfaces, finite speed of propagation and breaking of regularity. Since standard techniques cannot be straightforwardly applied, we present a way to generalize the relative entropy method, by involving approximations of solutions to the hydrodynamic equation, instead of exact solutions.

In [16], M. Simon and her coauthors study the hydrodynamic limit for a similar one-dimensional exclusion process but with an even more restricting dynamical constraint: this process with degenerate jump rates admits transient states, which it eventually leaves to reach an ergodic component if the initial macroscopic density is larger than a critical value, or one of its absorbing states otherwise. They show that, for initial profiles smooth enough and uniformly larger than the critical density, the macroscopic density profile evolves under the diffusive time scaling according to a fast diffusion equation. The first step in the proof is to show that the system typically reaches an ergodic component in subdiffusive time.

These two macroscopic behaviors belong to the class of moving interfaces problems, which are particularly hard to derive from the microscopic point of view.

5.6. Towards the weak KPE universality conjecture

In [32], P. Gonçalves, N. Perkowski and M. Simon derive the KPZ equation with boundary conditions, from an interacting particle system in contact with stochastic reservoirs, and they legitimate the choice done at the macroscopic level for the KPZ equation from the microscopic description of the system. This is more subtle than expected, because the boundary conditions do not behave canonically. The main challenge is to clarify the link between the macroscopic boundary effects and their atomic description.

5.7. Joule effect in chains of oscillators

In physics, the rotor chain has been investigated as an example of a system with two conserved quantities (angular momentum and energy), for which the thermal conductivity is finite (and therefore energy diffuses). Numerics shows an unexpected behaviour of the chain when the latter is connected at the boundaries to two thermostats, and a mechanical force imposes an average angular momentum at one boundary: the stationary temperature profile coincides with the values of the thermostats, but in the middle of the chain it raises to a much higher value. This behaviour is related to the presence of two conserved quantities and is sometimes referred to as Joule effect. Since the rotor model is too difficult to be treated analytically, T. Komorowski, S. Olla and M. Simon investigate in [25] the harmonic chain of oscillators, perturbed with a stochastic noise, which makes the heat transport diffusive, namely: the noise destroys the conservation law of the total momentum, but keeps the other two conservation laws (energy and stretch) intact. The boundaries of the chain are connected to two Langevin thermostats and an external force acting on one boundary puts the system in a non-equilibrium stationary state. The authors rigorously derive the Joule effect for a particular value of the noise intensity.

5.8. Quantum optics

In [18], S. De Bièvre and his co-authors introduce a new measure of the nonclassicality of the quantum states of an optical field, the so-called “ordering sensitivity” of the state, that measures the fluctuations of its Wigner function. This work is prolonged in two subsequent papers. In [24], S. De Bièvre and the same co-authors investigate a new class of quantum states they call the “Thermal Difference States” that can be generated by parametric down conversion. They investigate in particular the degree to which such states are nonclassical. In [34], S. De Bièvre and his postdoc A. Hertz, re-interpret the ordering sensitivity in terms of another physical property of the quantum states of an optical field, namely their quadrature coherence scale. It is shown in particular that a large such coherence scale is responsible for very fast environmental decoherence of the state.

5.9. Orbital stability

In [19], S. De Bièvre and S. Rota Nodari consider the orbital stability of relative equilibria of Hamiltonian dynamical systems on Banach spaces, in the presence of a multi-dimensional invariance group for the dynamics. They prove a persistence result for such relative equilibria, present a generalization of the Vakhitov-Kolokolov slope condition to this higher dimensional setting, and show how it allows to prove the local coercivity of the Lyapunov function, which in turn implies orbital stability. The method is applied to study the orbital stability of relative equilibria of nonlinear Schrödinger and Manakov equations. A comparison of their approach to the one by Grillakis-Shatah-Strauss is provided.

5.10. Exponential time-decay for discrete Fokker–Planck equations

In the research direction exposed in Section 3.3, G. Dujardin and his coauthors proposed and studied in [21] several discrete versions of homogeneous and inhomogeneous one-dimensional Fokker-Planck equations. They proved in particular, for these discretizations of velocity and space, the exponential convergence to the equilibrium of the solutions, for time-continuous equations as well as for time-discrete equations. Their method uses new types of discrete Poincaré inequalities for a “two-direction” discretization of the derivative in velocity. For the inhomogeneous problem, they adapted for the very first time hypocoercive methods to the discrete level.

5.11. Energy preserving methods for nonlinear Schrödinger equations

G. Dujardin and his coauthors have revisited and extended relaxation methods for nonlinear Schrödinger equations (NLS). The classical relaxation method for NLS is an energy preserving method and a mass preserving method. Moreover, it is only linearly implicit. A first proof of the second order accuracy was achieved in [14]. Moreover, the method was extended to enable to treat noncubic nonlinearities, nonlocal nonlinearities, as well as rotation terms. The resulting methods are still energy preserving and mass preserving. Moreover, they are shown to have second order accuracy numerically. These new methods are compared with fully implicit, mass and energy preserving methods of Crank and Nicolson.

5.12. High order linearly implicit methods for evolution problems

In [31], I. Lacroix and G. Dujardin have developed a new class of numerical integration methods for evolution problems. This class contains methods of arbitrarily high order that only require the solution of a linear system per time step. For evolution ODEs (Cauchy problems), they give a constructive proof of existence for such arbitrarily high order methods. For evolution PDEs, they demonstrate numerically that these new methods can outperform high order methods from the literature on several test cases.

5.13. CLT for Circular beta-Ensembles at High Temperature

In [33], A. Hardy and G. Lambert have obtained a central limit theorem for the 2D Coulomb gas particle system constrained on a circle in the high temperature regime. An interesting feature is that the limiting variance interpolates between the Lebesgue L^2 norm, corresponding to the infinite temperature setting, and the Sobolev $H^{1/2}$ seminorm, corresponding to the zero temperature regime.

5.14. DLR equations and rigidity for the Sine-β process

The work [20] by A. Hardy and his collaborators, recently accepted for publication in Communications on Pure and Applied Mathematics, provides a “statistical physics” description of the sine-β process by means of Dobroshin-Lanford-Ruelle (DLR) equations. This basically allows to give a meaning to “the natural infinite configurations process on the real line in the 2D Coulomb interaction”, provided there is a unique solution to the DLR equation which turns out to be true in this setting.

6. Partnerships and Cooperations

6.1. National Initiatives

6.1.1. ANR

A. de Laire is a member of the ANR ODA project.

Title: Dispersive and random waves

ANR reference: ANR-18-CE40-0020-01

Coordinator: Nikolay Tzvetkov, Université de Cergy-Pontoise

A. Hardy is a member of the ANR BoB project.

Title: Inférence bayésienne à ressources limitées - données massives et modèles coûteux

Programme ANR: (DS0705) 2016

ANR reference: ANR-16-CE23-0003

Coordinator: R. Bardenet, CNRS & Université de Lille

Duration: October 2016 - October 2020

M. Simon has been a member of the ANR EDNHS project.

Title: Diffusion de l'énergie dans des système hamiltoniens bruités

Type: Défi de tous les savoirs (DS10) 2014

ANR reference: ANR-14-CE25-0011

Coordinator: C. Bernardin, Université de Nice

Duration: October 2014 - October 2019

6.2. European Initiatives

6.2.1. FP7 & H2020 Projects

M. Simon is a collaborator of the ERC Starting Grant HyLEF project.

Title: Hydrodynamic Limits and Equilibrium Fluctuations: universality from stochastic systems

Duration: May 2017 - April 2022

Coordinator: P. Gonçalves, Instituto Superior Técnico, Lisbon, Portugal

7. Dissemination

7.1. Promoting Scientific Activities

7.1.1. Scientific Events: Organisation

7.1.1.1. Member of the Organizing Committees

A. de Laire co-organized the "Journée des Doctorants en Mathématiques du Nord-Pas-de-Calais".

7.1.2. Journal

7.1.2.1. Member of the Editorial Boards

S. De Bièvre is Associate Editor of the Journal of Mathematical Physics since January 2019.

7.1.2.2. Reviewer - Reviewing Activities

M. Simon is reviewer for the main international peer-reviewed journals in probability and statistical physics. In 2019, G. Dujardin served as a reviewer for the numerical analysis journals ESAIM:M2AN, Numerische Mathematik and IMA Journal of Numerical Analysis.

7.1.3. Invited Talks

M. Simon has been invited speaker at several conferences, among them:

- *Interactions PDEs/Probability: particle systems, hyperbolic conservation laws*, CIRM, Marseille
- *1st SFB International Workshop Taming Complexity in Partial Differential Systems*, University of Vienna, Austria

7.2. Teaching - Supervision - Juries

7.2.1. Teaching

Licence: A. de Laire, Mathématiques fondamentales 1, 54 TD, L1, Université de Lille

Licence: G. Dujardin, Calcul différentiel et intégral, 60h, BA2, Université Libre de Bruxelles, Belgium.

Master: A. de Laire, "Etude de Problèmes Elliptiques", 60h, M1, Université de Lille

Master: A. Hardy, "Probabilité, modèles et applications" 60h, M1, Université de Lille

Master: A. Hardy, "Séries temporelles", 30h, M1, Université de Lille

Master: M. Simon, "Introduction à la physique statistique", 56h, M2, Université de Lille

Master: M. Simon, "Markov Chains and Applications", Université de Lille and École Centrale Lille

In addition, A. de Laire is in charge of the Master 2 of Applied Mathematics at Université de Lille.

7.2.2. Supervision

PhD: P. Mennuni, “Ondes progressives de l’équation de Gross–Pitaevskii non locale : analyse et simulations”, Université de Lille, defended on the November 4, 2019; advisors: S. De Bièvre, A. de Laire, G. Dujardin.

7.3. Popularization

7.3.1. Interventions

M. Simon participated in the programs “Chercheurs itinérants” and “Fête de la Science”, and gave several lectures aimed at high-school students.

8. Bibliography

Major publications by the team in recent years

[1] C. BELTRÁN, A. HARDY. *Energy of the Coulomb gas on the sphere at low temperature*, in "Archive for Rational Mechanics and Analysis", 2019, vol. 231, p. 2007-2017 [*DOI :* 10.1007/s00205-018-1316-3], https://arxiv.org/abs/1803.11018

[2] C. BERNARDIN, P. GONÇALVES, M. JARA, M. SIMON. *Interpolation process between standard diffusion and fractional diffusion*, in "Annales de l’IHP, Probabilités et Statistiques", 2018, vol. 54, n^{o} 3, p. 1731–1757

[3] C. BESSE, G. DUJARDIN, I. LACROIX-VIOLET. *High Order Exponential Integrators for Nonlinear Schrödinger Equations with Application to Rotating Bose–Einstein Condensates*, in "SIAM Journal on Numerical Analysis", 2017, vol. 55, n^{o} 3, p. 1387-1411, https://dx.doi.org/10.1137/15M1029047

[4] D. CHAFAÏ, A. HARDY, M. MAÏDA. *Concentration for Coulomb gases and Coulomb transport inequalities*, in "Journal of Functional Analysis", 2018, vol. 275, n^{o} 6, p. 1447-1483

[5] M. CONFORTI, A. MUSSOT, A. KUDLINSKI, S. ROTA NODARI, G. DUJARDIN, S. DE BIÈVRE, A. ARMAROLI, S. TRILLO. *Heteroclinic Structure of Parametric Resonance in the Nonlinear Schrödinger Equation*, in "Phys. Rev. Lett.", Jun 2016, vol. 117, 013901, https://link.aps.org/doi/10.1103/PhysRevLett.117.013901

[6] T. FRANCO, P. GONÇALVES, M. SIMON. *Crossover to the stochastic Burgers equation for the WASEP with a slow bond*, in "Communications in Mathematical Physics", 2016, vol. 346, n^{o} 3, p. 801–838

[7] M. JARA, T. KOMOROWSKI, S. OLLA. *Superdiffusion of Energy in a Chain of Harmonic Oscillators with Noise*, in "Communications in Mathematical Physics", Oct 2015, vol. 339, n^{o} 2, p. 407–453, https://doi.org/10.1007/s00220-015-2417-6

[8] A. DE LAIRE. *Minimal Energy for the Traveling Waves of the Landau–Lifshitz Equation*, in "SIAM Journal on Mathematical Analysis", 2014, vol. 46, n^{o} 1, p. 96-132, https://doi.org/10.1137/130909081

[9] A. DE LAIRE, P. GRAVEJAT. *The Sine–Gordon regime of the Landau–Lifshitz equation with a strong easy-plane anisotropy*, in "Annales de l’Institut Henri Poincaré C, Analyse non linéaire", 2018 [*DOI :* 10.1016/J.ANIHPC.2018.03.005], http://www.sciencedirect.com/science/article/pii/S029414491830026X

Publications of the year

Doctoral Dissertations and Habilitation Theses

[10] P. MENNUNI.*Traveling waves of the nonlocal Gross-Pitaevskii equation: analysis and simulations*, Université de Lille, November 2019, https://hal.archives-ouvertes.fr/tel-02395234

[11] M. SIMON.*Microscopic derivation of degenerated diffusion phenomena*, Université de Lille, December 2019, Habilitation à diriger des recherches, https://hal.inria.fr/tel-02399713

Articles in International Peer-Reviewed Journal

[12] R. BARDENET, A. HARDY.*Time-frequency transforms of white noises and Gaussian analytic functions*, in "Applied and Computational Harmonic Analysis", 2019, forthcoming [*DOI :* 10.1016/J.ACHA.2019.07.003], https://hal.archives-ouvertes.fr/hal-01855678

[13] A. BENOIT.*Geometric optics expansions for hyperbolic corner problems II : from weak stability to violent instability*, in "SIAM Journal on Mathematical Analysis", 2019, vol. 49, n^{o} 5, p. 3335–3395, forthcoming [*DOI :* 10.1137/16M1060145], https://hal.archives-ouvertes.fr/hal-01242899

[14] C. BESSE, S. DESCOMBES, G. DUJARDIN, I. LACROIX-VIOLET.*Energy preserving methods for nonlinear Schrödinger equations*, in "IMA Journal of Numerical Analysis", 2019, forthcoming [*DOI :* 10.1016/J.APNUM.2019.11.008], https://hal.archives-ouvertes.fr/hal-01951527

[15] O. BLONDEL, C. CANCÈS, M. SASADA, M. SIMON.*Convergence of a Degenerate Microscopic Dynamics to the Porous Medium Equation*, in "Annales de l'Institut Fourier", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-01710628

[16] O. BLONDEL, C. ERIGNOUX, M. SASADA, M. SIMON.*Hydrodynamic limit for a facilitated exclusion process*, in "Annales de l'IHP - Probabilités et Statistiques", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-01798051

[17] D. BONHEURE, J.-B. CASTERAS, T. GOU, L. JEANJEAN.*Normalized solutions to the mixed dispersion nonlinear Schrödinger equation in the mass critical and supercritical regime*, in "Transactions of the American Mathematical Society", March 2019, vol. 372, n^{o} 3, p. 2167-2212 [*DOI :* 10.1090/TRAN/7769], https://hal.archives-ouvertes.fr/hal-02367293

[18] S. DE BIÈVRE, D. B. HOROSHKO, G. PATERA, M. I. KOLOBOV.*Measuring nonclassicality of bosonic field quantum states via operator ordering sensitivity*, in "Physical Review Letters", February 2019 [*DOI :* 10.1103/PHYSREVLETT.122.080402], https://hal.archives-ouvertes.fr/hal-01865597

[19] S. DE BIÈVRE, S. ROTA NODARI.*Orbital stability via the energy-momentum method: the case of higher dimensional symmetry groups*, in "Archive for Rational Mechanics and Analysis", January 2019, vol. 231, n^{o} 1, p. 233-284, https://arxiv.org/abs/1605.02523 [*DOI :* 10.1007/S00205-018-1278-5], https://hal.archives-ouvertes.fr/hal-01312534

[20] D. DEREUDRE, A. HARDY, T. LEBLÉ, M. MAÏDA.*DLR equations and rigidity for the Sine-beta process*, in "Communications on Pure and Applied Mathematics", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-01954367

[21] G. DUJARDIN, F. HÉRAU, P. LAFITTE-GODILLON.*Coercivity, hypocoercivity, exponential time decay and simulations for discrete Fokker- Planck equations*, in "Numerische Mathematik", 2019, forthcoming [*DOI :* 10.1007/S00211-019-01094-Y], https://hal.archives-ouvertes.fr/hal-01702545

[22] S. GUTIÉRREZ, A. DE LAIRE.*The Cauchy problem for the Landau-Lifshitz-Gilbert equation in BMO and self-similar solutions*, in "Nonlinearity", 2019, vol. 32, n^o 7, p. 2522-2563 [*DOI :* 10.1088/1361-6544/AB1296], https://hal.archives-ouvertes.fr/hal-01948679

[23] W. HACHEM, A. HARDY, S. SHAMAI.*Mutual Information of Wireless Channels and Block-Jacobi Ergodic Operators*, in "IEEE Transactions on Information Theory", November 2019, vol. 65, n^o 11, https://hal.archives-ouvertes.fr/hal-01954454

[24] D. HOROSHKO, S. DE BIÈVRE, G. PATERA, M. I. KOLOBOV.*Thermal-difference states of light: Quantum states of heralded photons*, in "Physical Review A", November 2019, vol. 100, n^o 5 [*DOI :* 10.1103/PHYSREVA.100.053831], https://hal.archives-ouvertes.fr/hal-02399355

[25] T. KOMOROWSKI, S. OLLA, M. SIMON.*An open microscopic model of heat conduction: evolution and non-equilibrium stationary states*, in "Communications in Mathematical Sciences", 2019, forthcoming, https://hal.inria.fr/hal-02081200

[26] A. DE LAIRE, P. GRAVEJAT.*The cubic Schrödinger regime of the Landau-Lifshitz equation with a strong easy-axis anisotropy*, in "Revista Matemática Iberoamericana", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-01954762

[27] A. DE LAIRE, P. MENNUNI.*Traveling waves for some nonlocal 1D Gross-Pitaevskii equations with nonzero conditions at infinity*, in "Discrete and Continuous Dynamical Systems - Series A", January 2020, vol. 40, n^o 1, p. 635-682, https://arxiv.org/abs/1812.08713 , https://hal.archives-ouvertes.fr/hal-01962779

Conferences without Proceedings

[28] R. BARDENET, P. CHAINAIS, J. FLAMANT, A. HARDY.*A correspondence between zeros of time-frequency transforms and Gaussian analytic functions*, in "13th International Conference on Sampling Theory and Applications, SampTA 2019", Bordeaux, France, August 2019, https://hal.archives-ouvertes.fr/hal-02091672

Other Publications

[29] R. AHMED, C. BERNARDIN, P. GONÇALVES, M. SIMON.*A Microscopic Derivation of Coupled SPDE's with a KPZ Flavor*, October 2019, https://arxiv.org/abs/1910.03996 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02307963

[30] O. BLONDEL, C. ERIGNOUX, M. SASADA, M. SIMON.*Hydrodynamic limit for a facilitated exclusion process*, February 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02008606

[31] G. DUJARDIN, I. LACROIX-VIOLET.*High order linearly implicit methods for evolution equations: How to solve an ODE by inverting only linear systems*, November 2019, https://arxiv.org/abs/1911.06016 - working paper or preprint, https://hal.inria.fr/hal-02361814

[32] P. GONÇALVES, N. PERKOWSKI, M. SIMON.*Derivation of the stochastic Burgers equation with Dirichlet boundary conditions from the WASEP*, March 2019, https://arxiv.org/abs/1710.11011 - 69 pages, https://hal.inria.fr/hal-01626604

[33] A. HARDY, G. LAMBERT.*CLT for Circular beta-Ensembles at High Temperature*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02385025

[34] A. HERTZ, S. DE BIÈVRE.*Quadrature coherence scale driven fast decoherence of bosonic quantum field states*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02394344

References in notes

[35] I. BEJENARU, A. D. IONESCU, C. E. KENIG, D. TATARU.*Global Schrödinger maps in dimensions $d \geq 2$: small data in the critical Sobolev spaces*, in "Annals of Mathematics", 2011, p. 1443–1506

[36] R. L. JERRARD, D. SMETS.*On Schrödinger maps from T^1 to S^2*, in "Ann. Sci. ENS", 2012, vol. 45, p. 637-680

[37] D. WEI.*Micromagnetics and Recording Materials*, Springer–Verlag Berlin Heidelberg, 2012, http://dx.doi.org/10.1007/978-3-642-28577-6 , https://doi.org/10.1007/978-3-642-28577-6

[38] S. ZHANG, A. A. BAKER, S. KOMINEAS, T. HESJEDAL.*Topological computation based on direct magnetic logic communication*, in "Scientific Reports", 2015, vol. 5, http://dx.doi.org/10.1038/srep15773

Project-Team MODAL

MOdel for Data Analysis and Learning

IN COLLABORATION WITH: Laboratoire Paul Painlevé (LPP)

IN PARTNERSHIP WITH:
CNRS

Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Optimization, machine learning and statistical methods

Table of contents

Project-Team MODAL

Creation of the Team: 2010 September 01, updated into Project-Team: 2012 January 01

Keywords:

Computer Science and Digital Science:

A3.1.4. - Uncertain data
A3.2.3. - Inference
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4.1. - Supervised learning
A3.4.2. - Unsupervised learning
A3.4.5. - Bayesian methods
A3.4.7. - Kernel methods
A5.2. - Data visualization
A6.2.3. - Probabilistic methods
A6.2.4. - Statistical methods
A6.3.3. - Data processing
A9.2. - Machine learning

Other Research Topics and Application Domains:

B2.2.3. - Cancer
B9.5.6. - Data science
B9.6.3. - Economy, Finance
B9.6.5. - Sociology

1. Team, Visitors, External Collaborators

Research Scientists

Christophe Biernacki [Team leader, Inria, Senior Researcher, HDR]
Pascal Germain [Inria, Researcher, until Oct 2019]
Benjamin Guedj [Inria, Researcher]
Hemant Tyagi [Inria, Researcher]

Faculty Members

Alain Celisse [Université de Lille, Associate Professor, HDR]
Sophie Dabo-Niang [Université de Lille, Professor]
Philippe Heinrich [Université de Lille, Associate Professor]
Serge Iovleff [Université de Lille, Associate Professor]
Guillemette Marot [Université de Lille, Associate Professor]
Cristian Preda [Université de Lille, Professor, HDR]
Vincent Vandewalle [Université de Lille, Associate Professor]

External Collaborators

Jean Francois Bouin [DiagRAMS, from Jul 2019]
Margot Correard [DiagRAMS, from Jul 2019]

Technical Staff

Iheb Eladib [Inria, Engineer, from Sep 2019]

Olivier Gauriau [Inria, Engineer, from Sep 2019]
Vincent Kubicki [Inria, Engineer, until Jul 2019]
Arthur Talpaert [Inria, Engineer, from Oct 2019]
Jean Francois Bouin [Inria, Engineer, from Apr 2019 until Jun 2019]
Margot Correard [Inria, Engineer, from Apr 2019 until Jun 2019]
Quentin Grimonprez [Inria, Engineer, from Oct 2019]

PhD Students
Filippo Antonazzo [Inria, from Oct 2019]
Yaroslav Averyanov [Inria]
Maxime Baelde [A-Volute]
Anne-Lise Bedenel [meilleureassurance.com]
Felix Biggs [University College London, UK, from Sep 2019]
Adrien Ehrhardt [CACF, until Mar 2019]
Arthur Leroy [Université de Paris]
Le Li [iAdvize, Université d'Angers]
Antoine Vendeville [University College London, UK, from Sep 2019]
Luxin Zhang [Worldline, from Feb 2019]
Wilfried Heyse [Inria, PhD Student, from Sep 2019]

Post-Doctoral Fellows
Florent Dewez [Inria, from Feb 2019]
Fabien Laporte [Ecole polytechnique, until Jul 2019]
Vera Shalaeva [Inria]

Visiting Scientists
Mihai Cucuringu [University of Oxford, UK, Associate Professor, until Jan 2019]
Abdou Ka Diongue [Gaston Berger University, Senegal, Professor, from Jun 2019 until Jul 2019]
Apoorv Vikram Singh [Independent, Visiting Researcher, from Oct 2019]
Vlad Barbu [Université de Rouen, Associate Professor, from Sep 2019]
Seydou-Nourou Sylla [Independent, Engineer, from May 2019 until Jun 2019]

Administrative Assistant
Anne Rejl [Inria]

2. Overall Objectives

2.1. Context

In several respects, modern society has strengthened the need for statistical analysis, even if other related names are sometimes preferably used depending on methods, communities and applications, as data analysis, machine learning or artificial intelligence. The genesis comes from the easier availability of data thanks to technological breakthroughs (storage, transfer, computing), and are now so widespread that they are no longer limited to large human organizations. The more or less conscious goal of such data availability is the expectation to improving the quality of "since the dawn of time" statistical stories which are namely discovering new knowledge or doing better predictions. These both central tasks can be referred respectively as unsupervised learning or supervised learning, even if it is not limited to them or other names exist depending on communities. Somewhere, it pursues the following hope: "more data for better quality and more numerous results".

However, today's data are increasingly complex. They gather mixed type features (for instance continuous data mixed with categorical data), missing or partially missing items (like intervals) and numerous variables (high dimensional situation). As a consequence, the target "better quality and more numerous results" of the previous adage (both words are important: "better quality" and also "more numerous") could not be reached through a somewhat "handwork" way, but should inevitably rely on some theoretical formalization and guarantee. Indeed, data can be so numerous and so complex (data can live in quite abstract spaces) that the "empirical" statistician is quickly outdated. However, data being subject by nature to randomness, the probabilistic framework is a very sensible theoretical environment to serve as a general guide for modern statistical analysis.

2.2. Goals

Modal is a project-team working on today's complex data sets (mixed data, missing data, high-dimensional data), for classical statistical targets (unsupervised learning, supervised learning, regression,...) with approaches relying on the probabilistic framework. This latter can be tackled through both model-based methods (as mixture models for a generic tool) and model-free methods (as probabilistic bounds on empirical quantities). Furthermore, Modal is connected to the real world by applications, typically with biological ones (some members have this skill) but many other ones are also considered since the application coverage of the Modal methodology is very large. It is also important to note that, in return, applications are often real opportunities for initiating academic questioning for the statistician (case of the Bilille platform and some bilateral contracts of the team).

From the academic communities point of view, Modal can be seen as belonging simultaneously to both the statistical learning and machine learning ones, as attested by its publications. Somewhere it is the opportunity to make a bridge between these two stochastic communities around a common but large probabilistic framework.

3. Research Program

3.1. Research axis 1: Unsupervised learning

Scientific locks related to unsupervised learning are numerous, concerning the clustering outcome validity, the ability to manage different kinds of data, the missing data questioning, the dimensionality of the data set,... Many of them are addressed by the team, leading to publication achievements, often with a specific package delivery (sometimes upgraded as a software or even as a platform grouping several software). Because of the variety of the scope, it involves nearly all the permanent team members, often with PhD students and some engineers. The related works are always embedded inside a probabilistic framework, typically model-based approaches but also model-free ones like PAC-Bayes (PAC stands for Probably Approximately Correct), because such a mathematical environment offers both a well-posed problem and a rigorous answer.

3.2. Research axis 2: Performance assessment

One main concern of the Modal team is to provide theoretical justifications on the procedures which are designed. Such guarantees are important to avoid misleading conclusions resulting from any unsuitable use. For example, one ingredient in proving these guarantees is the use of the PAC framework, leading to finite-sample concentration inequalities. More precisely, contributions to PAC learning rely on the classical empirical process theory and the PAC-Bayesian theory. The Modal team exploits such non-asymptotic tools to analyze the performance of iterative algorithms (such as gradient descent), cross-validation estimators, online change-point detection procedures, ranking algorithms, matrix factorization techniques and clustering methods, for instance. The team also develops some expertise on the formal dynamic study of algorithms related to mixture models (important models used in the previous unsupervised setting), like degeneracy for EM algorithm or also label switching for Gibbs algorithm.

3.3. Research axis 3: Functional data

Mainly due to technological advances, functional data are more and more widespread in many application domains. Functional data analysis (FDA) is concerned with the modeling of data, such as curves, shapes, images or a more complex mathematical object, though as smooth realizations of a stochastic process (an infinite dimensional data object valued in a space of eventually infinite dimension; space of squared integrable functions,...). Time series are an emblematic example even if it should not be limited to them (spectral data, spatial data,...). Basically, FDA considers that data correspond to realizations of stochastic processes, usually assumed to be in a metric, semi-metric, Hilbert or Banach space. One may consider, functional independent or dependent (in time or space) data objects of different types (qualitative, quantitative, ordinal, multivariate, time-dependent, spatial-dependent,...). The last decade saw a dynamic literature on parametric or non-parametric FDA approaches for different types of data and applications to various domains, such as principal component analysis, clustering, regression and prediction.

3.4. Research axis 4: Applications motivating research

The fourth axis consists in translating real application issues into statistical problems raising new (academic) challenges for models developed in Modal team. Cifre Phds in industry and interdisciplinary projects with research teams in Health and Biology are at the core of this objective. The main originality of this objective lies in the use of statistics with complex data, including in particular ultra-high dimension problems. We focus on real applications which cannot be solved by classical data analysis.

4. Application Domains

4.1. Economic world

The Modal team applies it research to the economic world through CIFRE Phd supervision such as CACF (credit scoring), A-Volute (expert in 3D sound), Meilleur Taux (insurance comparator), Worldline. It also has several contracts with companies such as COLAS, Nokia-Apsys/Airbus.

4.2. Biology

The second main application domain of the team is the biology. Members of the team are involved in the supervision and scientific animation of the bilille platform, the bioinformatics and bioanalysis platform and OncoLille project of Lille.

5. Highlights of the Year

5.1. Highlights of the Year

- Benjamin Guedj gave (with John Shawe-Taylor) a plenary tutorial of 2 hours for opening the ICML 2019 (Longbeach, California, USA – June 2019).
- Official creation in July 2019 of a startup DiagRAMS using MODAL's technology (MixtComp software) for predictive maintenance.
- Benjamin Guedj has received two best reviewer awards (top 5% of reviewers) for ICML 2019 and NeurIPS 2019, the flagship conferences in machine learning. Pascal Germain received the best reviewer award (top 5% of reviewers) for NeurIPS 2019.

5.1.1. More relevant results in 2019.

While Section 7 contains a complete list of results for 2019, the important results which were published in peer-reviewed international conferences/journals are described in Sections 7.2, 7.3, 7.13, 7.16, 7.17, 7.18, 7.19, 7.21, 7.22, 7.27, 7.28, 7.32, 7.37, 7.43 and 7.47.

6. New Software and Platforms

6.1. MixtComp

Mixture Computation

KEYWORDS: Clustering - Statistics - Missing data

FUNCTIONAL DESCRIPTION: MixtComp (Mixture Computation) is a model-based clustering package for mixed data originating from the Modal team (Inria Lille). It has been engineered around the idea of easy and quick integration of all new univariate models, under the conditional independence assumption. New models will eventually be available from researches, carried out by the Modal team or by other teams. Currently, central architecture of MixtComp is built and functionality has been field-tested through industry partnerships. Three basic models (Gaussian, multinomial, Poisson) are implemented, as well as two advanced models (Ordinal and Rank). MixtComp has the ability to natively manage missing data (completely or by interval). MixtComp is used as an R package, but its internals are coded in C++ using state of the art libraries for faster computation.

- Participants: Christophe Biernacki, Etienne Goffinet, Matthieu Marbac-Lourdelle, Quentin Grimonprez, Serge Iovleff and Vincent Kubicki
- Contact: Christophe Biernacki
- URL: https://cran.r-project.org/web/packages/RMixtComp/index.html

6.2. BlockCluster

Block Clustering

KEYWORDS: Statistic analysis - Clustering package

SCIENTIFIC DESCRIPTION: Simultaneous clustering of rows and columns, usually designated by biclustering, co-clustering or block clustering, is an important technique in two way data analysis. It consists of estimating a mixture model which takes into account the block clustering problem on both the individual and variables sets. The blockcluster package provides a bridge between the C++ core library and the R statistical computing environment. This package allows to co-cluster binary, contingency, continuous and categorical data-sets. It also provides utility functions to visualize the results. This package may be useful for various applications in fields of Data mining, Information retrieval, Biology, computer vision and many more.

FUNCTIONAL DESCRIPTION: BlockCluster is an R package for co-clustering of binary, contingency and continuous data based on mixture models.

RELEASE FUNCTIONAL DESCRIPTION: Initialization strategy enhanced

- Participants: Christophe Biernacki, Gilles Celeux, Parmeet Bhatia, Serge Iovleff, Vincent Brault and Vincent Kubicki
- Partner: Université de Technologie de Compiègne
- Contact: Serge Iovleff
- URL: http://cran.r-project.org/web/packages/blockcluster/index.html

6.3. CloHe

Clustering of Mixed data

KEYWORDS: Classification - Clustering - Missing data

FUNCTIONAL DESCRIPTION: Software of classification for mixed data with missing values with application to multispectral satellite image time-series

- Partners: CNRS - INRA
- Contact: Serge Iovleff
- URL: https://modal.lille.inria.fr/CloHe/

6.4. PACBayesianNMF

KEYWORDS: Statistics - Machine learning

FUNCTIONAL DESCRIPTION: Implementing NMF with a PAC-Bayesian approach relying upon block gradient descent

- Participants: Benjamin Guedj and Astha Gupta
- Contact: Benjamin Guedj
- URL: https://github.com/astha736/PACbayesianNMF

6.5. pycobra

KEYWORDS: Statistics - Data visualization - Machine learning

SCIENTIFIC DESCRIPTION: pycobra is a python library for ensemble learning, which serves as a toolkit for regression, classification, and visualisation. It is scikit-learn compatible and fits into the existing scikit-learn ecosystem.

pycobra offers a python implementation of the COBRA algorithm introduced by Biau et al. (2016) for regression.

Another algorithm implemented is the EWA (Exponentially Weighted Aggregate) aggregation technique (among several other references, you can check the paper by Dalalyan and Tsybakov (2007).

Apart from these two regression aggregation algorithms, pycobra implements a version of COBRA for classification. This procedure has been introduced by Mojirsheibani (1999).

pycobra also offers various visualisation and diagnostic methods built on top of matplotlib which lets the user analyse and compare different regression machines with COBRA. The Visualisation class also lets you use some of the tools (such as Voronoi Tesselations) on other visualisation problems, such as clustering.

- Participants: Bhargav Srinivasa Desikan and Benjamin Guedj
- Contact: Benjamin Guedj
- Publication: Pycobra: A Python Toolbox for Ensemble Learning and Visualisation
- URL: https://github.com/bhargavvader/pycobra

6.6. STK++

Statistical ToolKit

KEYWORDS: Statistics - Linear algebra - Framework - Learning - Statistical learning

FUNCTIONAL DESCRIPTION: STK++ (Statistical ToolKit in C++) is a versatile, fast, reliable and elegant collection of C++ classes for statistics, clustering, linear algebra, arrays (with an API Eigen-like), regression, dimension reduction, etc. The library is interfaced with lapack for many linear algebra usual methods. Some functionalities provided by the library are available in the R environment using rtkpp and rtkore.

STK++ is suitable for projects ranging from small one-off projects to complete data mining application suites.

- Participant: Serge Iovleff
- Contact: Serge Iovleff
- URL: http://www.stkpp.org

6.7. rtkore

STK++ core library integration to R using Rcpp

KEYWORDS: C++ - Data mining - Clustering - Statistics - Regression

FUNCTIONAL DESCRIPTION: STK++ (http://www.stkpp.org) is a collection of C++ classes for statistics, clustering, linear algebra, arrays (with an Eigen-like API), regression, dimension reduction, etc. The integration of the library to R is using Rcpp. The rtkore package includes the header files from the STK++ core library. All files contain only templated classes or inlined functions. STK++ is licensed under the GNU LGPL version 2 or later. rtkore (the stkpp integration into R) is licensed under the GNU GPL version 2 or later. See file LICENSE.note for details.

- Participant: Serge Iovleff
- Contact: Serge Iovleff
- URL: https://cran.r-project.org/web/packages/rtkore/index.html

6.8. MixAll

Clustering using Mixture Models

KEYWORDS: Clustering - Clustering package - Generative Models

FUNCTIONAL DESCRIPTION: MixAll is a model-based clustering package for modelling mixed data sets. It has been engineered around the idea of easy and quick integration of any kind of mixture models for any kind of data, under the conditional independence assumption. Currently five models (Gaussian mixtures, categorical mixtures, Poisson mixtures, Gamma mixtures and kernel mixtures) are implemented. MixAll has the ability to natively manage completely missing values when assumed as random. MixAll is used as an R package, but its internals are coded in C++ as part of the STK++ library (www.stkpp.org) for faster computation.

RELEASE FUNCTIONAL DESCRIPTION: clusterPredict allow to predcit membership for a new data set using a previously estimated model

- Participant: Serge Iovleff
- Partner: Université Lille 1
- Contact: Serge Iovleff
- URL: https://cran.r-project.org/web/packages/MixAll/

6.9. simerge

Statistical Inference for the Management of Extrem Risks, Genetics and Global epidemiology

KEYWORD: Biclustering

FUNCTIONAL DESCRIPTION: Allows to perform Co-Clustering on binary (Bernoulli) and counting variables (Poisson) using co-variables.

- Partner: Inria
- Contact: Serge Iovleff

6.10. MixtComp.V4

KEYWORDS: Clustering - Statistics - Missing data - Mixed data

FUNCTIONAL DESCRIPTION: MixtComp (Mixture Computation) is a model-based clustering package for mixed data originating from the Modal team (Inria Lille). It has been engineered around the idea of easy and quick integration of all new univariate models, under the conditional independence assumption. New models will eventually be available from researches, carried out by the Modal team or by other teams. Currently, central architecture of MixtComp is built and functionality has been field-tested through industry partnerships. Five basic models (Gaussian, Multinomial, Poisson, Weibull, NegativeBinomial) are implemented, as well as two advanced models (Functional and Rank). MixtComp has the ability to natively manage missing data (completely or by interval). MixtComp is used as an R package, but its internals are coded in C++ using state of the art libraries for faster computation.

RELEASE FUNCTIONAL DESCRIPTION: - New I/O system - Replacement of regex library - Improvement of initialization - Criteria for stopping the algorithm - Added management of partially missing data for several models - User documentation - Adding user features in R

- Participants: Christophe Biernacki, Vincent Kubicki, Matthieu Marbac-Lourdelle, Serge Iovleff, Quentin Grimonprez and Etienne Goffinet
- Partners: Université de Lille - CNRS
- Contact: Christophe Biernacki

6.11. MASSICCC

Massive Clustering with Cloud Computing

KEYWORDS: Statistic analysis - Big data - Machine learning - Web Application

SCIENTIFIC DESCRIPTION: The web application let users use several software packages developed by Inria directly in a web browser. Mixmod is a classification library for continuous and categorical data. MixtComp allows for missing data and a larger choice of data types. BlockCluster is a library for co-clustering of data. When using the web application, the user can first upload a data set, then configure a job using one of the libraries mentioned and start the execution of the job on a cluster. The results are then displayed directly in the browser allowing for rapid understanding and interactive visualisation.

FUNCTIONAL DESCRIPTION: The MASSICCC web application offers a simple and dynamic interface for analysing heterogeneous data with a web browser. Various software packages for statistical analysis are available (Mixmod, MixtComp, BlockCluster) which allow for supervised and supervised classification of large data sets.

- Contact: Christophe Biernacki
- URL: https://massiccc.lille.inria.fr

6.12. Platforms

6.12.1. MASSICCC Platform

MASSICCC is a demonstration platform giving access through a SaaS (service as a software) concept to data analysis libraries developed at Inria. It allows obtaining results either directly through a website specific display (specific and interactive visual outputs) or through an R data object download. It started in October 2015 for two years and is common to the Modal team (Inria Lille) and the Select team (Inria Saclay). In 2016, two packages have been integrated: Mixmod and MixtComp (see the specific section about MixtComp). In 2017, the BlockCluster package has been integrated and also a particular attention to provide meaningful graphical outputs (for Mixmod, MixtComp and BlockCluster) directly in the web platform itself has led to some specific developments. In 2019, a new version of the MixtComp software has been developed.

7. New Results

7.1. Axis 1: Data Units Selection in Statistics

Participant: Christophe Biernacki.

Usually, the data unit definition is fixed by the practitioner but it can happen that he/she hesitates between several data unit options. In this context, it is highlighted that it is possible to embed data unit selection into a classical model selection principle. The problem is introduced in a regression context before to focus on the model-based clustering and co-clustering context, for data of different kinds (continuous, count, categorical). This work was published in an international journal in 2018 and leads to a keynote as an invited speaker to the 12th Scientific Meeting Classification and Data Analysis Group Cassino (CLADAG 2019) in Italy [41].

It is a joint work with Alexandre Lourme from University of Bordeaux.

7.2. Axis 1: Model-Based Co-clustering for Ordinal Data of different dimensions

Participant: Christophe Biernacki.

This work has been motivated by a psychological survey on women affected by a breast tumor. Patients replied at different moments of their treatment to questionnaires with answers on ordinal scale. The questions relate to aspects of their life called dimensions. To assist the psychologists in analyzing the results, it is useful to emphasize a structure in the dataset. The clustering method achieves that by creating groups of individuals that are depicted by a representative of the group. From a psychological position, it is also useful to observe how questions may be grouped. This is why a clustering should also be performed on the features, which is called a co-clustering problem. However, gathering questions that are not related to the same dimension does not make sense from a psychologist stance. Therefore, the present work corresponds to perform a constrained co-clustering method aiming to prevent questions from different dimensions from getting assembled in a same column-cluster. In addition, evolution of co-clusters along time has been investigated. The method relies on a constrained Latent Block Model embedding a probability distribution for ordinal data. Parameter estimation relies on a Stochastic EM-algorithm associated to a Gibbs sampler, and the ICL-BIC criterion is used for selecting the numbers of co-clusters. The resulting work is now accepted in an international journal [28]. The related R package ordinalClust has been also written and has led to a specific preprint [73] now submitted to an international journal.

This is joint work with Margot Selosse (PhD student) and Julien Jacques, both from University of Lyon 2, and Florence Cousson-Gélie from University Paul Valéry Montpellier 3.

7.3. Axis 1: Model-based co-clustering for mixed type data

Participant: Christophe Biernacki.

Over decades, a lot of studies have shown the importance of clustering to emphasize groups of observations. More recently, due to the emergence of high-dimensional datasets with a huge number of features, co-clustering techniques have emerged and proposed several methods for simultaneously producing groups of observations and features. By synthesizing the dataset in blocks (the crossing of a row-cluster and a column-cluster), this technique can sometimes summarize better the data and its inherent structure. The Latent Block Model (LBM) is a well-known method for performing a co-clustering. However, recently, contexts with features of different types (here called mixed type datasets) are becoming more common. Unfortunately, the LBM is not directly applicable on this kind of dataset. The present work extends the usual LBM to the so-called Multiple Latent Block Model (MLBM) which is able to handle mixed type datasets. The inference is done through a Stochastic EM-algorithm embedding a Gibbs sampler and model selection criterion is defined to choose the number of row and column clusters. This method was successfully used on simulated and real datasets. This work is now accepted in an international journal [29].

This is joint work with Margot Selosse (PhD student) and Julien Jacques, both from University of Lyon 2.

7.4. Axis 1: Relaxing the Identically Distributed Assumption in Gaussian Co-Clustering for High Dimensional Data

Participant: Christophe Biernacki.

A co-clustering model for continuous data that relaxes the identically distributed assumption within blocks of traditional co-clustering is presented. The proposed model, although allowing more flexibility, still maintains the very high degree of parsimony achieved by traditional co-clustering. A stochastic EM algorithm along with a Gibbs sampler is used for parameter estimation and an ICL criterion is used for model selection. Simulated and real datasets are used for illustration and comparison with traditional co-clustering. This work has been submitted to an international journal [63].

This is a joint work with Michael Gallaugher (PhD student) and Paul McNicholas, both from McMaster University (Canada). Michael Gallaugher visited Modal for three months in 2018.

7.5. Axis 1: Gaussian-based visualization of Gaussian and non-Gaussian model-based clustering

Participants: Christophe Biernacki, Vincent Vandewalle.

A generic method is introduced to visualize in a Gaussian-like way, and onto R^2, results of Gaussian or non-Gaussian model-based clustering. The key point is to explicitly force a spherical Gaussian mixture visualization to inherit from the within cluster overlap which is present in the initial clustering mixture. The result is a particularly user-friendly draw of the clusters, allowing any practitioner to have a thorough overview of the potentially complex clustering result. An entropic measure allows us to inform of the quality of the drawn overlap, in comparison to the true one in the initial space. The proposed method is illustrated on four real data sets of different types (categorical, mixed, functional and network) and is implemented on the R package ClusVis. This work is now in minor revision for an international journal [54]. It has also led to an invited talk to an international conference [42], and several other invitations (the workshop “Advances in data science for big and complex data” at Université Paris-Dauphine in January and the seminary of the Probability and Statistics team of the University Nice Sophia-Antipolis in November).

This is a joint work with Matthieu Marbac from ENSAI.

7.6. Axis 1: Co-clustering: A versatile way to perform clustering

Participant: Christophe Biernacki.

Standard model-based clustering is known to be very efficient for low dimensional data sets, but it fails for properly addressing high dimension (HD) ones, where it suffers from both statistical and computational drawbacks. In order to counterbalance this curse of dimensionality, some proposals have been made to take into account redundancy and features utility, but related models are not suitable for too many variables. We advocate that the latent bloc model, a probabilistic model for co-clustering, is of particular interest to perform HD clustering of individuals even if it is not its primary function. We illustrate in an empirical manner the trade-off bias-variance of the co-clustering strategy in scenarii involving HD fundaments (correlated variables, irrelevant variables) and show the ability of co-clustering to outperform simple mixture row-clustering. An early version of this work has been presented to an national conference with international audience [46].

We also co-organized a special session to an international conference [45] to discuss the potential links between deterministic methods for co-clustering (based on a metric and computer science procedure) or probabilistic methods for co-clustering (mainly based on mixture models). It was the opportunity to gather related communities which are often distinct.

All are joint works with Christine Keribin from Université Paris-Sud.

7.7. Axis 1: Dealing with missing data in model-based clustering through a MNAR model

Participants: Christophe Biernacki, Fabien Laporte.

Since the 90s, model-based clustering is largely used to classify data. Nowadays, with the increase of available data, missing values are more frequent. Traditional ways to deal with them consist in obtaining a filled data set, either by discarding missing values or by imputing them. In the first case, some information is lost; in the second case, the final clustering purpose is not taken into account through the imputation step. Thus, both solutions risk to blur the clustering estimation result. Alternatively, we defend the need to embed the missingness mechanism directly within the clustering modeling step. There exists three types of missing data: missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR). In all situations logistic regression is proposed as a natural and flexible candidate model. In particular, its flexibility property allows us to design some meaningful parsimonious variants, as dependency on missing values or dependency on the cluster label. In this unified context, standard model selection criteria can be used to select between such different missing data mechanisms, simultaneously with the number of clusters. Practical interest of our proposal is illustrated on data derived from medical studies suffering from many missing data. This work has been presented as an invited speaker to an international conference [31]. It has also been presented at a national conference with international audience [47] and as a poster to the international Working Group on Model-Based Clustering [69]. Currently, a preprint is being finalized for submission to an international journal.

It is a joint work with Gilles Celeux from Inria Saclay and Julie Josse from Ecole Polytechnique.

7.8. Axis 1: Organized Co-Clustering for textual data synthesis

Participant: Christophe Biernacki.

Recently, different studies have demonstrated the interest of co-clustering, which simultaneously produces clusters of lines and columns. The present work introduces a novel co-clustering model for parsimoniously summarizing textual data in documents × terms format. Besides highlighting homogeneous coclusters - as other existing algorithms do - we also distinguish noisy coclusters from significant ones, which is particularly useful for sparse documents × term matrices. Furthermore, our model proposes a structure among the significant coclusters and thus obtains a better interpretability to the user. By forcing a structure through row-clusters and column-clusters, this approach is competitive in terms of documents clustering, and offers user-friendly results. The algorithm derived for the proposed method is a Stochastic EM algorithm embedding a Gibbs sampling step and the Poisson distribution. A paper is currently in revision in an international journal [72].

This is joint work with Margot Selosse (PhD student) and Julien Jacques, both from University of Lyon 2.

7.9. Axis 1: Model-Based Co-clustering with Co-variables

Participant: Serge Iovleff.

This work has been motivated by an epidemiological and genetic survey of malaria disease in Senegal. Data were collected between 1990 and 2008. It is based on a latent block model taking into account the problem of grouping variables and clustering individuals by integrating information given by a set of co-variables. Numerical experiments on simulated data sets and an application on real genetic data highlight the interest of this approach. An article has been submitted to *Journal of Classification* and should incorporate “Major Revisions”.

7.10. Axis 1: Linking canonical and spectral clustering

Participants: Christophe Biernacki, Vincent Vandewalle.

It is a recent work aiming at defining a mathematical bridge between classical model-based clustering and classical spectral clustering. Interest of such a prospect is to be able to compare both methods through the rigorous scheme of model selection paradigm. It is still an ongoing work, with several short working papers.

It is a joint work with Alexandre Lourme from University of Bordeaux.

7.11. Axis 1: Predictive clustering

Participants: Christophe Biernacki, Vincent Vandewalle.

Many data, for instance in biostatistics, contain some sets of variables which permit evaluating unobserved traits of the subjects (e.g., we ask question about how many pizzas, hamburgers, chips... are eaten to know how healthy are the food habits of the subjects). Moreover, we often want to measure the relations between these unobserved traits and some target variables (e.g., obesity). Thus, a two-steps procedure is often used: first, a clustering of the observations is performed on the sets of variables related to the same topic; second, the predictive model is fitted by plugging the estimated partitions as covariates. Generally, the estimated partitions are not exactly equal to the true ones. We investigate the impact of these measurement errors on the estimators of the regression parameters, and we explain when this two-steps procedure is consistent. We also present a specific EM algorithm which simultaneously estimates the parameters of the clustering and predictive models. It is an ongoing work.

It is a joint work with Matthieu Marbac from ENSAI and Mohammed Sedki from University Paris-Sud.

7.12. Axis 1: Ranking and synchronization from pairwise measurements via SVD

Participant: Hemant Tyagi.

Given a measurement graph $G = ([n], E)$ and an unknown signal $r \in R^n$, we investigate algorithms for recovering r from pairwise measurements of the form $r_i - r_j$; $\{i, j\} \in E$. This problem arises in a variety of applications, such as ranking teams in sports data and time synchronization of distributed networks. Framed in the context of ranking, the task is to recover the ranking of n teams (induced by r) given a small subset of noisy pairwise rank offsets. We propose a simple SVD-based algorithmic pipeline for both the problem of time synchronization and ranking. We provide a detailed theoretical analysis in terms of robustness against both sampling sparsity and noise perturbations with outliers, using results from matrix perturbation and random matrix theory. Our theoretical findings are complemented by a detailed set of numerical experiments on both synthetic and real data, showcasing the competitiveness of our proposed algorithms with other state-of-the-art methods.

This is joint work with Alexandre d'Aspremont (CNRS & ENS, Paris) and Mihai Cucuringu (University of Oxford, UK) and is available as a preprint [61].

7.13. Axis 1: SPONGE: A generalized eigenproblem for clustering signed networks

Participant: Hemant Tyagi.

We introduce a principled and theoretically sound spectral method for k-way clustering in signed graphs, where the affinity measure between nodes takes either positive or negative values. Our approach is motivated by social balance theory, where the task of clustering aims to decompose the network into disjoint groups, such that individuals within the same group are connected by as many positive edges as possible, while individuals from different groups are connected by as many negative edges as possible. Our algorithm relies on a generalized eigenproblem formulation inspired by recent work on constrained clustering. We provide theoretical guarantees for our approach in the setting of a signed stochastic block model, by leveraging tools from matrix perturbation theory and random matrix theory. An extensive set of numerical experiments on both synthetic and real data shows that our approach compares favorably with state-of-the-art methods for signed clustering, especially for large number of clusters and sparse measurement graphs.

This is joint work with Mihai Cucuringu (University of Oxford, UK), Peter Davies (University of Warwick, UK) and Aldo Glielmo (Imperial College, London, UK) and was mostly done while Hemant Tyagi was affiliated to the Alan Turing Institute. It was published in the proceedings of an international conference [32].

7.14. Axis 2: Multi-kernel unmixing and super-resolution using the Modified Matrix Pencil method

Participant: Hemant Tyagi.

Consider L groups of point sources or spike trains, with the l^{th} group represented by $x_l(t)$. For a function $g : R \to R$, let $g_l(t) = g(t/\mu_l)$ denote a point spread function with scale $\mu_l > 0$, and with $\mu_1 < \cdots < \mu_L$. With $y(t) = \sum_{l=1}^{L} (g_l \star x_l)(t)$, our goal is to recover the source parameters given samples of y, or given the Fourier samples of y. This problem is a generalization of the usual super-resolution setup wherein $L = 1$; we call this the multi-kernel unmixing super-resolution problem. Assuming access to Fourier samples of y, we derive an algorithm for this problem for estimating the source parameters of each group, along with precise non-asymptotic guarantees. Our approach involves estimating the group parameters sequentially in the order of increasing scale parameters, i.e., from group 1 to L. In particular, the estimation process at stage $1 \leq l \leq L$ involves (i) carefully sampling the tail of the Fourier transform of y, (ii) a *deflation* step wherein we subtract the contribution of the groups processed thus far from the obtained Fourier samples, and (iii) applying Moitra's modified Matrix Pencil method on a deconvolved version of the samples in (ii).

This is joint work with Stephane Chretien (National Physical Laboratory, UK & Alan Turing Institute, London) and was mostly done while Hemant Tyagi was affiliated to the Alan Turing Institute. It is currently under revision in an international journal and is available as a preprint [56].

7.15. Axis 2: Provably robust estimation of modulo 1 samples of a smooth function with applications to phase unwrapping

Participant: Hemant Tyagi.

Consider an unknown smooth function $f : [0,1]^d \to R$, and assume we are given n noisy mod 1 samples of f, i.e., $y_i = (f(x_i) + \eta_i) \mod 1$, for $x_i \in [0,1]^d$, where η_i denotes the noise. Given the samples $(x_i, y_i)_{i=1}^n$, our goal is to recover smooth, robust estimates of the clean samples $f(x_i) \mod 1$. We formulate a natural approach for solving this problem, which works with angular embeddings of the noisy mod 1 samples over the unit circle, inspired by the angular synchronization framework. This amounts to solving a smoothness regularized least-squares problem – a quadratically constrained quadratic program (QCQP) – where the variables are constrained to lie on the unit circle. Our proposed approach is based on solving its relaxation, which is a *trust-region sub-problem* and hence solvable efficiently. We provide theoretical guarantees demonstrating its robustness to noise for adversarial, as well as random Gaussian and Bernoulli noise models. To the best of our knowledge, these are the first such theoretical results for this problem. We demonstrate the robustness and efficiency of our proposed approach via extensive numerical simulations on synthetic data, along with a simple least-squares based solution for the unwrapping stage, that recovers the original samples of f (up to a global shift). It is shown to perform well at high levels of noise, when taking as input the denoised modulo 1 samples. Finally, we also consider two other approaches for denoising the modulo 1 samples that leverage tools from Riemannian optimization on manifolds, including a Burer-Monteiro approach for a semidefinite programming relaxation of our formulation. For the two-dimensional version of the problem, which has applications in synthetic aperture radar interferometry (InSAR), we are able to solve instances of real-world data with a million sample points in under 10 seconds, on a personal laptop.

This is joint work with Mihai Cucuringu (University of Oxford, UK) and was mostly done while Hemant Tyagi was affiliated to the Alan Turing Institute. It has been accepted to appear (after minor revision) in an international journal, and is available as a preprint [60].

7.16. Axis 2: Learning general sparse additive models from point queries in high dimensions

Participant: Hemant Tyagi.

We consider the problem of learning a d-variate function f defined on the cube $[-1,1]^d \subset R^d$, where the algorithm is assumed to have black box access to samples of f within this domain. Denote $S_r; r = 1, \cdots, r_0$ to be sets consisting of unknown r-wise interactions amongst the coordinate variables. We then focus on the setting where f has an additive structure, i.e., it can be represented as

$$f = \sum_{j \in \S_1} \phi_j + \sum_{j \in S_2} \phi_j + \cdots + \sum_{j \in S_{r_0}} \phi_j,$$

where each ϕ_j; $j \in S_r$ is at most r-variate for $1 \le r \le r_0$. We derive randomized algorithms that query f at carefully constructed set of points, and exactly recover each S_r with high probability. In contrary to the previous work, our analysis does not rely on numerical approximation of derivatives by finite order differences.

This is joint work with Jan Vybiral (Czech Technical University, Prague) and was mostly done while Hemant Tyagi was affiliated to the Alan Turing Institute. It has now been published in an international journal [30].

7.17. Axis 2: Sparse non-negative super-resolution - simplified and stabilized

Participant: Hemant Tyagi.

The convolution of a discrete measure, $x = \sum_{i=1}^{k} a_i \delta_{t_i}$, with a local window function, $\phi(s-t)$, is a common model for a measurement device whose resolution is substantially lower than that of the objects being observed. Super-resolution concerns localising the point sources with an accuracy beyond the essential support of $\phi(s-t)$, typically from m noisy samples of the convolution output. We consider the setting of x being non-negative and seek to characterise all non-negative measures approximately consistent with the samples. We first show that x is the unique non-negative measure consistent with the samples provided the samples are exact, and $m \ge 2k+1$ samples are available, and $\phi(s-t)$ generates a Chebyshev system. This is independent of how close the sample locations are and *does not rely on any regulariser beyond non-negativity*; as such, it extends and clarifies the work by Schiebinger et al. and De Castro et al., who achieve the same results but require a total variation regulariser, which we show is unnecessary. Moreover, we establish stability results in the setting where the samples are corrupted with noise. The main innovation of these results is that non-negativity alone is sufficient to localise point sources beyond the essential sensor resolution.

This is joint work with Armin Eftekhari (EPFL, Switzerland), Jared Tanner (University of Oxford, UK), Andrew Thompson (National Physical Laboratory, UK), Bogdan Toader (University of Oxford, UK) and was mostly done while Hemant Tyagi was affiliated to the Alan Turing Institute. It has now been published in an international journal [24].

7.18. Axis 2: Pseudo-Bayesian learning with kernel Fourier transform as prior

Participant: Pascal Germain.

We revisit the kernel random Fourier features (RFF) method through the lens of the PAC-Bayesian theory. While the primary goal of RFF is to approximate a kernel, we look at the Fourier transform as a prior distribution over trigonometric hypotheses. It naturally suggests learning a posterior on these hypotheses. We derive generalization bounds that are optimized by learning a pseudo-posterior obtained from a closed-form expression, and corresponding learning algorithms.

This joint work with Emilie Morvant from Université Jean Monnet de Saint-Etienne (France), and Gaël Letarte from Université Laval (Québec, Canada) has been initiated in 2018 when Gaël Letarte was doing an internship at Inria, and led to a publication in the proceedings of AISTATS 2019 conference [36]. The same work has been prensented as a poster in the “Workshop on Machine Learning with guarantees @ NeurIPS 2019”.

An extension of this work, co-authored with Léo Gautheron, Amaury Habrard, Marc Sebban, and Valentina Zantedeschi – all from Université Jean Monnet de Saint-Etienne – has been presented at the national conference CAp 2019 [44]. It is also the topic of a technical report [64].

7.19. Axis 2: PAC-Bayesian binary activated deep neural networks

Participant: Pascal Germain, Benjamin Guedj

We present a comprehensive study of multilayer neural networks with binary activation, relying on the PAC-Bayesian theory. Our contributions are twofold: (i) we develop an end-to-end framework to train a binary activated deep neural network, overcoming the fact that binary activation function is non-differentiable; (ii) we provide nonvacuous PAC-Bayesian generalization bounds for binary activated deep neural networks. Noteworthy, our results are obtained by minimizing the expected loss of an architecture-dependent aggregation of binary activated deep neural networks. The performance of our approach is assessed on a thorough numerical experiment protocol on real-life datasets. This work has been published in the proceedings of NeurIPS 2019 conference [35].

It is a joint work with Gaël Letarte and François Laviolette, from Université Laval (Québec, Canada).

7.20. Axis 2: Improved PAC-Bayesian Bounds for Linear Regression

Participant: Pascal Germain, Vera Shalaeva

We improve the PAC-Bayesian error bound for linear regression provided in the literature. The improvements are two-fold. First, the proposed error bound is tighter, and converges to the generalization loss with a well-chosen temperature parameter. Second, the error bound also holds for training data that are not independently sampled. In particular, the error bound applies to certain time series generated by well-known classes of dynamical models, such as ARX models.

It is a joint work with Mihaly Petreczky and Alireza Fakhrizadeh Esfahani from Université de Lille. It has been accepted for publication as part of the AAAI 2020 conference [38].

7.21. Axis 2: Multiview Boosting by controlling the diversity and the accuracy of view-specific voters

Participant: Pascal Germain

We present a comprehensive study of multilayer neural networks with binary activation, relying on the PAC-Bayesian We propose a boosting based multiview learning algorithm which iteratively learns i) weights over view-specific voters capturing view-specific information; and ii) weights over views by optimizing a PAC-Bayes multiview C-Bound that takes into account the accuracy of view-specific classifiers and the diversity between the views. We derive a generalization bound for this strategy following the PAC-Bayes theory which is a suitable tool to deal with models expressed as weighted combination over a set of voters.

It is a joint work with Emilie Morvant from Université Jean Monnet de Saint-Etienne and with Massih-Reza Amini of Université de Grenoble, and with Anil Goyal affiliated to both institutions. This work has been published in the journal Neurocomputing [26].

7.22. Axis 2: PAC-Bayes and Domain Adaptation

Participant: Pascal Germain

In machine learning, Domain Adaptation (DA) arises when the distribution generating the test (target) data differs from the one generating the learning (source) data. It is well known that DA is a hard task even under strong assumptions, among which the covariate-shift where the source and target distributions diverge only in their marginals, i.e. they have the same labeling function. Another popular approach is to consider a hypothesis class that moves closer the two distributions while implying a low-error for both tasks. This is a VC-dim approach that restricts the complexity of a hypothesis class in order to get good generalization. Instead, we propose a PAC-Bayesian approach that seeks for suitable weights to be given to each hypothesis in order to build a majority vote. We prove a new DA bound in the PAC-Bayesian context. This leads us to design the first DA-PAC-Bayesian algorithm based on the minimization of the proposed bound. Doing so, we seek for a ρ-weighted majority vote that takes into account a trade-off between three quantities. The first two quantities being, as usual in the PAC-Bayesian approach, (a) the complexity of the majority vote (measured by a Kullback-Leibler divergence) and (b) its empirical risk (measured by the ρ-average errors on the source sample). The third quantity is (c) the capacity of the majority vote to distinguish some structural difference between the source and target samples.

This work has been published in the journal Neurocomputing [25].

It is a joint work with Emilie Morvant and Amaury Habrard from Université Jean Monnet de Saint-Etienne (France), and with François Laviolette from Université Laval (Québec, Canada).

7.23. Axis 2: Interpreting Neural Networks as Majority Votes through the PAC-Bayesian Theory

Participant: Pascal Germain, Paul Viallard

We propose a PAC-Bayesian theoretical study of the two-phase learning procedure of a neural network introduced by Kawaguchi et al. (2017). In this procedure, a network is expressed as a weighted combination of all the paths of the network (from the input layer to the output one), that we reformulate as a PAC-Bayesian majority vote. Starting from this observation, their learning procedure consists in (1) learning "prior" network for fixing some parameters, then (2) learning a "posterior" network by only allowing a modification of the weights over the paths of the prior network. This allows us to derive a PAC-Bayesian generalization bound that involves the empirical individual risks of the paths (known as the Gibbs risk) and the empirical diversity between pairs of paths. Note that similarly to classical PAC-Bayesian bounds, our result involves a KL-divergence term between a "prior" network and the "posterior" network. We show that this term is computable by dynamic programming without assuming any distribution on the network weights.

This early result has been accepted as a poster presentation in the international workshop "Workshop on Machine Learning with guarantees @ NeurIPS 2019" [50].

This is a joint work with researchers from Université Jean Monnet de Saint-Etienne: Amaury Habrard, Emilie Morvant, and Rémi Emonet.

7.24. Axis 2: Still no free lunches: the price to pay for tighter PAC-Bayes bounds

Participant: Benjamin Guedj

"No free lunch" results state the impossibility of obtaining meaningful bounds on the error of a learning algorithm without prior assumptions and modelling. Some models are expensive (strong assumptions, such as as subgaussian tails), others are cheap (simply finite variance). As it is well known, the more you pay, the more you get: in other words, the most expensive models yield the more interesting bounds. Recent advances in robust statistics have investigated procedures to obtain tight bounds while keeping the cost minimal. The present paper explores and exhibits what the limits are for obtaining tight PAC-Bayes bounds in a robust setting for cheap models, addressing the question: is PAC-Bayes good value for money?

Joint work with Louis Pujol (Université Paris-Saclay). Available as a preprint: [68]

7.25. Axis 2: PAC-Bayesian Contrastive Unsupervised Representation Learning

Participant: Benjamin Guedj, Pascal Germain

Contrastive unsupervised representation learning (CURL) is the state-of-the-art technique to learn representations (as a set of features) from unlabelled data. While CURL has collected several empirical successes recently, theoretical understanding of its performance was still missing. In a recent work, Arora et al. (2019) provide the first generalisation bounds for CURL, relying on a Rademacher complexity. We extend their framework to the flexible PAC-Bayes setting, allowing to deal with the non-iid setting. We present PAC-Bayesian generalisation bounds for CURL, which are then used to derive a new representation learning algorithm. Numerical experiments on real-life datasets illustrate that our algorithm achieves competitive accuracy, and yields generalisation bounds with non-vacuous values.

Joint work with Kento Nozawa (University of Tokyo & RIKEN). Available as a preprint: [71]

7.26. Axis 2: Sequential Learning of Principal Curves: Summarizing Data Streams on the Fly

Participant: Benjamin Guedj

When confronted with massive data streams, summarizing data with dimension reduction methods such as PCA raises theoretical and algorithmic pitfalls. Principal curves act as a nonlinear generalization of PCA and the present paper proposes a novel algorithm to automatically and sequentially learn principal curves from data streams. We show that our procedure is supported by regret bounds with optimal sublinear remainder terms. A greedy local search implementation (called `slpc`, for Sequential Learning Principal Curves) that incorporates both sleeping experts and multi-armed bandit ingredients is presented, along with its regret computation and performance on synthetic and real-life data.

Joint work with Le Li (Université d'Angers & iAdvize). Available as a preprint: [67]

7.27. Axis 2: PAC-Bayes Un-Expected Bernstein Inequality

Participant: Benjamin Guedj

We present a new PAC-Bayesian generalization bound. Standard bounds contain a $\sqrt{L_n \cdot KL/n}$ complexity term which dominates unless L_n, the empirical error of the learning algorithm's randomized predictions, vanishes. We manage to replace L_n by a term which vanishes in many more situations, essentially whenever the employed learning algorithm is sufficiently stable on the dataset at hand. Our new bound consistently beats state-of-the-art bounds both on a toy example and on UCI datasets (with large enough n). Theoretically, unlike existing bounds, our new bound can be expected to converge to 0 faster whenever a Bernstein/Tsybakov condition holds, thus connecting PAC-Bayesian generalization and *excess risk* bounds—for the latter it has long been known that faster convergence can be obtained under Bernstein conditions. Our main technical tool is a new concentration inequality which is like Bernstein's but with X^2 taken outside its expectation.

Joint work with Peter Grünwald (CWI), Zakaria Mhammedi (Australian National University).

This work has been accepted at NeurIPS 2019, will be presented as a poster in the main conference and as a oral in the workshop "Machine Learning with guarantees", and is included in the proceedings of NeurIPS 2019.

Published: [37]

7.28. Axis 2: Attributing and Referencing (Research) Software: Best Practices and Outlook from Inria

Participant: Benjamin Guedj

Software is a fundamental pillar of modern scientiic research, not only in computer science, but actually across all elds and disciplines. However, there is a lack of adequate means to cite and reference software, for many reasons. An obvious rst 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 this paper a contribution to the ongoing eeorts 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.

Joint work with Pierre Alliez, Roberto Di Cosmo, Alain Girault, Mohand-Said Hacid, Arnaud Legrand, Nicolas Rougier (Inria).

This work has been published in the journal Computing in Science and Engineering.

Published: [14]

7.29. Axis 2: Revisiting clustering as matrix factorisation on the Stiefel manifold

Participant: Benjamin Guedj

This paper studies clustering for possibly high dimensional data (*e.g.* images, time series, gene expression data, and many other settings), and rephrase it as low rank matrix estimation in the PAC-Bayesian framework. Our approach leverages the well known Burer-Monteiro factorisation strategy from large scale optimisation, in the context of low rank estimation. Moreover, our Burer-Monteiro factors are shown to lie on a Stiefel manifold. We propose a new generalized Bayesian estimator for this problem and prove novel prediction bounds for clustering. We also devise a componentwise Langevin sampler on the Stiefel manifold to compute this estimator.

Joint work with Stéphane Chrétien (Université Lyon-2). Available as a preprint: [55]

7.30. Axis 2: A Primer on PAC-Bayesian Learning

Participant: Benjamin Guedj

This survey on PAC-Bayesian learning has been the backbone to a successful proposal for an ICML 2019 plenary tutorial.

Generalised Bayesian learning algorithms are increasingly popular in machine learning, due to their PAC generalisation properties and flexibility. The present paper aims at providing a self-contained survey on the resulting PAC-Bayes framework and some of its main theoretical and algorithmic developments.

This work has been published in the proceedings of the French Mathematical Society. Published as [66].

7.31. Axis 2: Perturbed Model Validation: A New Framework to Validate Model Relevance

Participant: Benjamin Guedj

This paper introduces Perturbed Model Validation (PMV), a new technique to validate model relevance and detect overfitting or underfitting. PMV operates by injecting noise to the training data, re-training the model against the perturbed data, then using the training accuracy decrease rate to assess model relevance. A larger decrease rate indicates better concept-hypothesis fit. We realise PMV by perturbing labels to inject noise, and evaluate PMV on four real-world datasets (breast cancer, adult, connect-4, and MNIST) and nine synthetic datasets in the classification setting. The results reveal that PMV selects models more precisely and in a more stable way than cross-validation, and effectively detects both overfitting and underfitting.

It is a joint work with Jie Zhang, Earl Barr, John Shawe-Taylor (all with UCL), and Mark Harman (UCL & Facebook). Available as a preprint: [75].

7.32. Axis 2: Decentralized learning with budgeted network load using Gaussian copulas and classifier ensembles

Participant: Benjamin Guedj

We examine a network of learners which address the same classification task but must learn from different data sets. The learners cannot share data but instead share their models. Models are shared only one time so as to preserve the network load. We introduce DELCO (standing for Decentralized Ensemble Learning with COpulas), a new approach allowing to aggregate the predictions of the classifiers trained by each learner. The proposed method aggregates the base classifiers using a probabilistic model relying on Gaussian copulas. Experiments on logistic regressor ensembles demonstrate competing accuracy and increased robustness in case of dependent classifiers. A companion python implementation is available online.

Joint work with John Klein, Olivier Colot, Mahmoud Albardan (Université de Lille).

This work has been published in the proceedings of ECML-PKDD 2019, as part (oral presentation) of the workshop Decentralized Machine Learning at the Edge. Published: [34]

7.33. Axis 2: Online k-means Clustering

Participant: Benjamin Guedj

We study the problem of online clustering where a clustering algorithm has to assign a new point that arrives to one of k clusters. The specific formulation we use is the k-means objective: At each time step the algorithm has to maintain a set of k candidate centers and the loss incurred is the squared distance between the new point and the closest center. The goal is to minimize regret with respect to the best solution to the k-means objective ($\mathcal{C}$) in hindsight. We show that provided the data lies in a bounded region, an implementation of the Multiplicative Weights Update Algorithm (*MWUA*) using a discretized grid achieves a regret bound of $\widetilde{O}(\sqrt{T})$ in expectation. We also present an online-to-offline reduction that shows that an efficient no-regret online algorithm (despite being allowed to choose a different set of candidate centres at each round) implies an offline efficient algorithm for the k-means problem. In light of this hardness, we consider the slightly weaker requirement of comparing regret with respect to $(1+\epsilon)\mathcal{C}$ and present a no-regret algorithm with runtime $O\left(T(\text{poly}(\log(T), k, d, 1/\epsilon)^{k(d+O(1))}\right)$. Our algorithm is based on maintaining an incremental coreset and an adaptive variant of the *MWUA*. We show that naïve online algorithms, such as *Follow The Leader*, fail to produce sublinear regret in the worst case. We also report preliminary experiments with synthetic and real-world data.

Joint work with Varun Kanade, Guy Rom (University of Oxford), Vincent Cohen-Addad (CNRS). Available as a preprint: [57]

7.34. Axis 2: Non-linear aggregation of filters to improve image denoising

Participant: Benjamin Guedj

We introduce a novel aggregation method to efficiently perform image denoising. Preliminary filters are aggregated in a non-linear fashion, using a new metric of pixel proximity based on how the pool of filters reaches a consensus. We provide a theoretical bound to support our aggregation scheme, its numerical performance is illustrated and we show that the aggregate significantly outperforms each of the preliminary filters.

Joint work with Juliette Rengot (Ecole des Ponts).

This work has been accepted at the Computing Conference 2020 (July 2020, London, UK) and will be included in the proceedings. Published: [33]

7.35. Axis 2: Multiple change-points detection with reproducing kernels

Participant: Alain Celisse

We tackle the change-point problem with data belonging to a general set. We build a penalty for choosing the number of change-points in the kernel-based method of Harchaoui and Cappé (2007). This penalty generalizes the one proposed by Lebarbier (2005) for a one-dimensional signal changing only through its mean. We prove a non-asymptotic oracle inequality for the proposed method, thanks to a new concentration result for some function of Hilbert-space valued random variables. Experiments on synthetic and real data illustrate the accuracy of our method, showing that it can detect changes in the whole distribution of data, even when the mean and variance are constant.

Joint work with Sylvain Arlot (Orsay) and Zaïd Harchaoui (Seattle). This work has been accepted in JMLR [15].

7.36. Axis 2: Analysis of early stopping rules based on discrepancy principle

Participant: Alain Celisse

We describe a general unified framework for analyzing the statistical performance of early stopping rules based on the minimum discrepancy principle (DP). Finite-sample bounds such as deviation or oracle inequalities are derived with high probability. Since it turns out that DP suffers some deficiencies when estimating smooth functions, refinements involving smoothing of the residuals are introduced and analyzed. Theoretical bounds established in the fixed design setting under mild assumptions such as the boundedness of the kernel. When focusing on the smoothed discrepancy principle, such bounds are even extended to the random design setting by means of a new change-of-norm argument

Joint work with Markus Reiß(Humboldt) and Martin Wahl (Humboldt). This work has been already presented several times in seminars.

7.37. Axis 3: Short-term air temperature forecasting using Nonparametric Functional Data Analysis and SARMA models

Participant: Sophie Dabo-Niang

Air temperature is a significant meteorological variable that affects social activities and economic sectors. In this paper, a non-parametric and a parametric approach are used to forecast hourly air temperature up to 24 h in advance. The former is a regression model in the Functional Data Analysis framework. The nonlinear regression operator is estimated using a kernel function. The smoothing parameter is obtained by a cross-validation procedure and used for the selection of the optimal number of closest curves. The other method applied is a Seasonal Autoregressive Moving Average (SARMA) model, the order of which is determined by the Bayesian Information Criterion. The obtained forecasts are combined using weights calculated based on the forecast errors. The results show that SARMA has a better performance for the first 6 forecasted hours, after which the Non-Parametric Functional Data Analysis (NPFDA) model provides superior results. Forecast pooling improves the accuracy of the forecasts.

It is a joint work with Stelian Curceac (Rothamsted Research, UK) Camille Ternynck (CERIM, Université de Lille) Taha B.M.J. Ouarda (INRS, Québec, Canada) Fateh Chebana (INRS, Québec, Canada). This work has been published in the journal Environmental Modelling and Software [18].

7.38. Axis 3: Mathematical Modeling and Study of Random or Deterministic Phenomena

Participant: Sophie Dabo-Niang

In order to identify mathematical modeling (including functional data analysis) and interdisciplinary research issues in evolutionary biology, epidemiology, epistemology, environmental and social sciences encountered by researchers in Mayotte, the first international conference on mathematical modeling (CIMOM'18) was held in Dembéni, Mayotte, from November 15 to 17, 2018, at the Centre Universitaire de Formation et de Recherche. The objective was to focus on mathematical research with interdisciplinarity. This contribution is a book discusses key aspects of recent developments in applied mathematical analysis and modeling. It was written after the international conference on mathematical modeling in Mayotte, where a call for chapters of the book was made. They were written in the form of journal articles, with new results extending the talks given during the conference and were reviewed by independent reviewers and book publishers It highlights a wide range of applications in the fields of biological and environmental sciences, epidemiology and social perspectives. Each chapter examines selected research problems and presents a balanced mix of theory and applications on some selected topics. Particular emphasis is placed on presenting the fundamental developments in mathematical analysis and modeling and highlighting the latest developments in different fields of probability and statistics. The chapters are presented independently and contain enough references to allow the reader to explore the various topics presented.

It is a joint work with Solym Manou-Abi and Jean-Jacques Salone (University of Mayotte, France). This book is to appear Wiley (ISTE) [21].

7.39. Axis 3: Categorical functional data analysis

Participants: Cristian Preda, Quentin Grimonprez, Vincent Vandewalle.

The research on functional data analysis is very actual. The R package "fda" is the most famous one implementing methodology for functional data. To the best of our knowledge, and quite surprisingly, there is no recent researches devoted to categorical functional data despite its ability to model real situations in different fields of applications: health and medicine (status of a patient over time), economy (status of the market), sociology (evolution of social status), and so on. We have developed the methodology to visualize, do dimension reduction and extract feature from categorical functional data. For this, the *cfda* R package has been developed.

7.40. Axis 4: Proteomic signature of early death in heart failure patients

Participants: Guillemette Marot, Vincent Vandewalle.

Heart failure (HF) remains a main cause of mortality worldwide. Risk stratification of patients with systolic chronic HF is critical to identify those who may benefit from advanced HF therapies. The aim of this study is to identify plasmatic proteins that could predict the early death (within 3 years) of HF patients with reduced ejection fraction hospitalized in CHRU de Lille. In this framework, we have performed LASSO logistic regression to perform variable selection in order to select candidates protein to predict early death in HF patients. An article has been accepted in Scientific Reports [19].

This is a joint work with Marie Cuvelliez, Florence Pinet and Christophe Bauters from INSERM.

7.41. Axis 4: Statistical analysis of high-throughput proteomic data

Participants: Guillemette Marot, Vincent Vandewalle, Wilfried Heyse.

From March until August 2019, Guillemette Marot and Vincent Vandewalle have supervised the internship of Wilfried Heyse (Master 2 Ingénieurie de Sytèmes Numériques). The purpose of this internship was to identify new circulating biomarkers of left ventricular remodelling (LVR) in patients suffering from myocardial infarction (MI). The aim is to precisely identify earlier after MI the patients at high risk of developing LVR that is quantified by imaging one year after MI. For that purpose, high throughput proteomic approach was used. This technology allows the measurement of 5000 proteins simultaneously. In parallel to these measures corresponding to the concentration of a protein in a plasma sample collected from one patient at a specific time, echocardiographic and clinical information will be collected on each of the 200 patients. Several approaches have been used to predict the LVR based on proteins measurements. In particular penalized regression such as LASSO and variable clustering. Wilfried Heyse has now started a Phd Thesis granted by INSERM and supervised by Christophe Bauters, Guillemette Marot and Vincent Vandewalle. One of the main challenge is to take into account the variations of the biomarkers according to the time (several measurement times), in order the improve the understanding of biological mechanisms involved on LVR.

This is a joint work with Florence Pinet and Christophe Bauters from INSERM.

7.42. Axis 4: Linking different kinds of Omics data through a model-based clustering approach

Participants: Guillemette Marot, Vincent Vandewalle, Wilfried Heyse.

In this work, a mixture model allowing for genes clustering using both microarray (continuous) and RNAseq (count) expression data is proposed. More generally, it answers the clustering of variables issue, when variables are of different kinds (continuous and discrete here). Variables describing the same gene are constrained to belong to the same cluster. This constraint allows us to obtain a model that links the microarray and RNAseq measurements without needing parametric constraints on the form of this link. The proposed approach has been illustrated on simulated data, as well as on real data from TCGA (The Cancer Genome Atlas). It has been presented in an international conference [49].

This is a joint work with Camille Ternynck from EA2694.

7.43. Axis 4: Real-time Audio Sources Classification

Participants: Christophe Biernacki, Maxime Baelde.

This work addresses the recurring challenge of real-time monophonic and polyphonic audio source classification. The whole power spectrum is directly involved in the proposed process, avoiding complex and hazardous traditional feature extraction. It is also a natural candidate for polyphonic events thanks to its additive property in such cases. The classification task is performed through a nonparametric kernel-based generative modeling of the power spectrum. Advantage of this model is twofold: it is almost hypothesis free and it allows to straightforwardly obtain the maximum a posteriori classification rule of online signals. Moreover it makes use of the monophonic dataset to build the polyphonic one. Then, to reach the real-time target, the complexity of the method can be tuned by using a standard hierarchical clustering preprocessing of sound models, revealing a particularly efficient computation time and classification accuracy trade-off. The proposed method reveals encouraging results both in monophonic and polyphonic classification tasks on benchmark and owned datasets, even in real-time situations. This method also has several advantages compared to the state-of-the-art methods include a reduced training time, no hyperparameters tuning, the ability to control the computation - accuracy trade-off and no training on already mixed sounds for polyphonic classification. This work is now published in an international journal [16] and Maxime Baelde defended his PhD thesis on this topic this year [11].

It is a joint work with Raphaël Greff, from the A-Volute company.

7.44. Axis 4: Matching of descriptors evolving over time

Participants: Christophe Biernacki, Anne-Lise Bedenel.

In the web domain, and in particular for insurance comparison, data constantly evolve, implying that it is difficult to directly exploit them. For example, to do a classification, performing standard learning processes require data descriptors equal for both learning and test samples. Indeed, for answering web surfer expectation, online forms whence data come from are regularly modified. So, features and data descriptors are also regularly modified. In this work, it is introduced a process to estimate and understand connections between transformed data descriptors. This estimated matching between descriptors will be a preliminary step before applying later classical learning methods. Anne-Lise Bedenel defended her PhD thesis on this topic this year [12].

It is a joint work with Laetitia Jourdan, from Université de Lille.

7.45. Axis 4: Supervised multivariate discretization and levels merging for logistic regression

Participants: Christophe Biernacki, Vincent Vandewalle, Adrien Ehrhardt.

For regulatory and interpretability reasons, the logistic regression is still widely used by financial institutions to learn the refunding probability of a loan given the applicants characteristics from historical data. Although logistic regression handles naturally both quantitative and qualitative data, three ad hoc pre-processing steps are usually performed: firstly, continuous features are discretized by assigning factor levels to predetermined intervals; secondly, qualitative features, if they take numerous values, are grouped; thirdly, interactions (products between two different features) are sparsely introduced. By reinterpreting these discretized (resp. grouped) features as latent variables and by modeling the conditional distribution of each of these latent variables given each original feature with a polytomous logistic link (resp. contingency table), a novel model-based resolution of the discretization problem is introduced. Estimation is performed via a Stochastic Expectation-Maximization (SEM) algorithm and a Gibbs sampler to find the best discretization (resp. grouping) scheme w.r.t. any classical logistic regression loss (AIC, BIC, test set AUC,...). For detecting interacting features, the same scheme is used by replacing the Gibbs sampler by a Metropolis-Hastings algorithm. The good performances of this approach are illustrated on simulated and real data from Credit Agricole Consumer Finance. Adrien Ehrhardt defended his PhD thesis on this topic this year [13]. A preprint is being finalized to be submitted to an international journal or conference [62].

This is a joint work with Philippe Heinrich from Université de Lille.

7.46. Axis 4: MASSICCC Platform for SaaS Software Availability

Participant: Christophe Biernacki.

MASSICCC is a demonstration platform giving access through a SaaS (service as a software) concept to data analysis libraries developed at Inria. It allows to obtain results either directly through a website specific display (specific and interactive visual outputs) or through an R data object download. It started in October 2015 for two years and is common to the Modal team (Inria Lille) and the Select team (Inria Saclay). In 2016, two packages have been integrated: Mixmod and MixtComp (see the specific section about MixtComp). In 2017, the BlockCluster package has been integrated and also a particular attention to provide meaningful graphical outputs (for Mixmod,MixtComp and BlockCluster) directly in the web platform itself has led to some specific developments. In 2019, MASSICCC has been presented to a workshop [39]. In addition, the Mixtcomp software is now available on the CRAN depository. Currently, a preprint for an international journal dedicated to software is also in progress.

The MASSICCC platform is available here in the web: https://massiccc.lille.inria.fr

7.47. Axis 4: Domain adaptation from a pre-trained source model

Participants: Christophe Biernacki, Pascal Germain, Luxin Zhang.

Traditional statistical learning paradigm assumes the consistency between train and test data distributions. This rarely holds in many real-life applications. The domain adaptation paradigm proposes a variety of techniques to overcome this issue. Most of the works in this area seek either for a latent space where source and target data share the same distribution, or for a transformation of the source distribution to match the target one. Both strategies require learning a model on the transformed source data. An original scenario is studied where one is given a model that has been constructed using expertise on the source data that is not accessible anymore. To use directly this model on target data, we propose to learn a transformation from the target domain to the source domain. Up to our knowledge, this is a new perspective on domain adaptation. This learning problem is introduced and formalized. We study the assumptions and the sufficient conditions mandatory to guarantee a good accuracy when using the source model directly on transformed target data. By pursuing this idea, a new domain adaptation method based on optimal transport is proposed. We experiment our method on a fraud detection problem.

Luxin Zhang begun his PhD thesis on this topic and presented this early result in an international conference [51].

It is a joint work with Yacine Kessaci, both from Worldline.

7.48. Axis 4: Reject Inference Methods in Credit Scoring: a rational review

Participants: Christophe Biernacki, Vincent Vandewalle, Adrien Ehrhardt.

The granting process of all credit institutions is based on the probability that the applicant will refund his/her loan given his/her characteristics. This probability also called score is learnt based on a dataset in which rejected applicants are de facto excluded. This implies that the population on which the score is used will be different from the learning population. Thus, this biased learning can have consequences on the scorecard's relevance. Many methods dubbed reject inference have been developed in order to try to exploit the data available from the rejected applicants to build the score. However most of these methods are considered from an empirical point of view, and there is some lack of formalization of the assumptions that are really made, and of the theoretical properties that can be expected. We propose a formalisation of these usually hidden assumptions for some of the most common reject inference methods, and we discuss the improvement that can be expected. These conclusions are illustrated on simulated data and on real data from Credit Agricole Consumer Finance (CACF), a major European loan issuer. Adrien Ehrhardt defended his PhD thesis on this topic this year [13]. A preprint is being finalized to be submitted to an international journal or conference.

This is a joint work with Philippe Heinrich from Université de Lille.

7.49. Other: Projection Under Pairwise Control

Participant: Christophe Biernacki.

Visualization of high-dimensional and possibly complex (non-continuous for instance) data onto a low-dimensional space may be difficult. Several projection methods have been already proposed for displaying such high-dimensional structures on a lower-dimensional space, but the information lost is not always easy to use. Here, a new projection paradigm is presented to describe a non-linear projection method that takes into account the projection quality of each projected point in the reduced space, this quality being directly available in the same scale as this reduced space. More specifically, this novel method allows a straightforward visualization data in R2 with a simple reading of the approximation quality, and provides then a novel variant of dimensionality reduction. This work is now under minor revision in an international journal [53].

It is a joint work with Hiba Alawieh and Nicolas Wicker, both from Université de Lille.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. COLAS company

Participant: Christophe Biernacki.

COLAS is a world leader in the construction and maintenance of transport infrastructure. This bilateral contract aims at classifying mixed data obtained with sensors coming from a study of the aging of road surfacing. The challenge is to deal with many missing (sensors failures) and correlated data (sensors proximity).

8.2. Bilateral Grants with Industry

8.2.1. EIT-Sysbooster: Nokia - Apsys/Airbus

Participant: Alain Celisse.

Nokia and Airbus are two worldwide known companies respectively working in communications and transport areas. The purpose of this contract is to perform root cause analysis to reduce (at the end) the number of failures.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. ONCOLille partnership

Participants: Sophie Dabo-Niang, Cristian Preda.

ONCOLille is a regional scientific interest group whose purpose is to develop fundamental, translational (pre-clinical) and clinical interdisciplinary cancer research, particularly in the field of resistance to therapies. Sophie Dabo-Niang is member of the executive group.

9.2. National Initiatives

9.2.1. Programme of Investments for the Future (PIA)

Bilille is a member of the PIA “Infrastructures en biologie-santé” IFB, French Institute of Bioinformatics (https://www.france-bioinformatique.fr/en). As the co-head of the platform, Guillemette Marot is thus involved in this network.

9.2.2. RHU PreciNASH

Participant: Guillemette Marot.

RHU PreciNASH

Acronym: PreciNASH

Project title: Non-alcoholic steato-hepatitis (NASH) from disease stratification to novel therapeutic approaches

Coordinator: F. Pattou

Duration: 5 years

Partners: FHU Integra and Sanofi

Abstract: PreciNASH, project coordinated by Pr. F. Pattou (UMR 859, EGID), aims at better understanding non alcoholic stratohepatitis (NASH) and improving its diagnosis and care. In this RHU, Guillemette Marot supervises a 2 years post-doc, as her team EA 2694 is a member of the FHU Integra. EA 2694 is involved in the WP1 for the development of a clinical-biological model for the prediction of NASH. Other partners of the FHU are UMR 859, UMR 1011 and UMR 8199, these last three teams being part of the labex EGID (European Genomic Institute for Diabetes). Sanofi is the main industrial partner of the RHU PreciNASH. The whole project will last 5 years (2016-2021).

9.2.3. CNRS PEPS Blanc – BayesRealForRNN project

Participants: Pascal Germain, Vera Shalaeva.

BayesRealForRNN project: PAC-Bayesian theory for recurrent neural networks: a control theoretic approach

Coordinator: Mihaly Petreczky, CNRS, UMR 9189 CRIStAL, Université de Lille

Year: 2019

Abstract: The project proposes to analyze the mathematical correctness of deep learning algorithms by combining techniques from control theory and PAC-Bayesian statistical theory. More precisely, the project proposes to concentrate on recurrent neural networks (RNNs), develop their structure theory using techniques from contol theory, and then apply this structure theory to derive PAC-Bayesian error bounds for RNNs.

9.2.4. CNRS AMIES PEPS 2 - DiagChange project

Participants: Cristian Preda, Quentin Grimonprez.

DiagChange

Coordinator: Cristian Preda, Inria MODAL

Year: 2019

Abstract: The project proposes to study the topic of change detection distribution for multivariate signal in a industrial context.The project is in collaboration with the Diagrams start-up.

9.2.5. AMIES PEPS 1 - CADIS2

Participants: Serge Iovleff, Sophie Dabo-Niang, Cristian Preda.

Partners: Société SIRS https://www.sirs-fr.com/sirs/fr/

Acronym: CADIS2

Project title: Classification Automatique D’Images Sentinel-2

Coordinator: Serge Iovleff

Year: 2019

Duration: 1 year

Abstract: In the context of several European projects, SIRS is in charge of exploring the improvements to be made to the "High Resolution Layers" as well as future prototypes such as "CORINE Land Cover +", on a European scale using the Sentinel-2 images, through the project H2020 "ECoLaSS". The CADIS2 project aims to develop, study and implement supervised classification methods to classify trees in predefined forest areas by SIRS.

9.2.6. AMIES PEPS 2 - MadiPa

Participants: Stéphane Girard, Serge Iovleff.

Partners: Société Phimeca http://phimeca.com/, Mistis team Inria Grenoble Rhône-Alpes

Acronym: MadiPa

Project title: Modèles Auto-associatifs pour la Dispersion de Polluants dans l'Atmosphère

Coordinator: Stéphane Iovleff

Duration: 18 month (start in december 2019)

Abstract: Our goal is to develop a method for predicting the dispersion of pollutants in the atmosphere from an initial emission map and meteorological data. A map of the probabilities of exceeding a critical threshold of pollutants will be estimated thanks to the construction of a metamodel: the large dimension of the problem is reduced by the use of auto-associative models, a non-linear extension of the Principal Components Analysis.

9.2.7. ANR

9.2.7.1. ANR APRIORI

Participants: Benjamin Guedj, Pascal Germain, Hemant Tyagi, Vera Shalaeva.

APRIORI 2019–2023, ANR PRC

PAC-Bayesian theory and algorithms for deep learning and representation learning.

Main coordinator of the project: Emilie Morvant, Université Jean Monnet.

Funding: 300k EUR.

2 partners - MODAL (Inria LNE), Hubert Curien Lab. (UMR CNRS 5516).

9.2.7.2. ANR BEAGLE

Participants: Benjamin Guedj, Pascal Germain.

BEAGLE 2019–2023, ANR JCJC

PAC-Bayesian theory and algorithms for agnostic learning

Main coordinator of the project: Benjamin Guedj

Funding: 180k EUR

The consortium also includes Pierre Alquier (RIKEN AIP, Japan), Peter Grünwald (CWI, The Netherlands), Rémi Bardenet (UMR CRIStAL 9189).

9.2.7.3. ANR SMILE

Participants: Christophe Biernacki, Vincent Vandewalle.

SMILE Project-2018-2022

ANR project (ANR SMILE - Statistical Modeling and Inference for unsupervised Learning at LargE-Scale)

Main coordinator of the project: Faicel Chamroukhi, LMNO, Université de Caen

4 partners - MODAL (Inria LNE), LMNO UMR CNRS 6139 (Caen), LMRS UMR CNRS 6085 (Rouen), LIS UMR CNRS 7020 (Toulon).

9.2.7.4. ANR TheraSCUD2022

Participant: Guillemette Marot.

Acronym: TheraSCUD2022

Project title: Targeting the IL-20/IL-22 balance to restore pulmonary, intestinal and metabolic homeostasis after cigarette smoking and unhealthy diet

Coordinator: P. Gosset

Duration: 3 years (2017-2020)

Partners: CIIL Institut Pasteur de Lille and UMR 1019 INRA Clermont-Ferrand

Abstract: TheraSCUD2022, project coordinated by P. Gosset (Institut Pasteur de Lille), studies inflammatory disorders associated with cigarette smoking and unhealthy diet (SCUD). Guillemette Marot is involved in this ANR project as head of bilille platform, and will supervise 1 year engineer on integration of omic data. The duration of this project is 3 years (2017-2020).

9.2.8. Working groups

Sophie Dabo-Niang belongs to the following working groups:

- STAFAV (STatistiques pour l'Afrique Francophone et Applications au Vivant)
- ERCIM Working Group on computational and Methodological Statistics, Nonparametric Statistics Team

Benjamin Guedj belongs to the following working groups (GdR) of CNRS:

- ISIS (local referee for Inria Lille - Nord Europe)
- MaDICS
- MASCOT-NUM (local referee for Inria Lille - Nord Europe).

Guillemette Marot belongs to the StatOmique working group.

9.2.9. Other initiatives

Participants: Serge Iovleff, Cristian Preda, Vincent Vandewalle.

Serge Iovleff is the head of the project CloHe granted in 2016 by the Mastodons CNRS challenge “Big data and data quality”. The project is axed on the design of classification and clustering algorithms for mixed data with missing values with applications to high spatial resolution multispectral satellite image time-series. Website. Cristian Preda and Vincent Vandewalle are also members of the CloHe project.

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

PERF-AI project (Nov 2018 - Nov 2020, involving Benjamin Guedj, Vincent Vandewalle - hired Florent Dewez, Arthur Taelpert). Two partners: Inria LNE and the company Safety Line (Paris, France).

Commercial aviation is already responsible for 3% of the total CO2 emissions, and with a constant growth rate of 5% per year, traffic will double within the next decade. With the support of new technologies such as Big Data, Artificial Intelligence, in-flight connectivity, major improvements can be introduced to optimize flight trajectories. PERF-AI focuses on the challenge of minimizing fuel consumption throughout the flight. The aim of PERF-AI is to provide a flight trajectory optimization prototype that implements new machine learning performance models.

The first step of the project that was carried out in the first year was to define, implement and test narrow system identification techniques. Several Machine Learning methods have been tried and have provided very encouraging initial results.

PERF-AI main objective is to provide a computation engine that can be used in two ways:

- support update of FMS that integrate individual aircraft performance models, that allow to perform accurate trajectory prediction;
- perform trajectory optimization on the ground using most accurate aircraft performance models.

9.3.2. Collaborations with Major European Organizations

Sophie Dabo-Niang is chair of EMS-CDC (European Mathematical society-Committe of Developing Countries).

Sophie Dabo-Niang is a member of the executive committee of CIMPA (International Centre of Pure and Applied Mathematics)

9.4. International Initiatives

9.4.1. Inria International Labs

9.4.1.1. 6PAC (IIL CWI-Inria)

Scientific leaders: Benjamin Guedj, Peter Grünwald.

Other members: Emilie Kaufmann (Inria LNE, EPI SequeL), Wouter Koolen (CWI).

Title: Making Probably Approximately Correct Learning Active, Sequential, Structure-aware, Efficient, Ideal and Safe

International Partner (Institution - Laboratory - Researcher):

CWI (Netherlands) - Machine Learning Group - Peter Grünwald (head)

Start year: 2018, renewed for 2019 and 2020

Webpage: https://bguedj.github.io/6pac/index.html

This project roots in statistical learning theory, which can be viewed as the theoretical foundations of machine learning. The most common framework is a setup in which one is given n training examples, and the goal is to build a predictor that would be efficient on new (similar) data. This efficiency should be supported by PAC (Probably Approximately Correct) guarantees, e.g. upper bounds on the excess risk of a predictor that hold with high probability. Such guarantees however often hold under stringent assumptions which are typically never met in real-life application, e.g., independent, identically distributed data. More realistic modelling of data has triggered many research efforts in several directions: first, accommodating possible data (e.g., dependent, heavy-tailed), and second, in the direction of sequential learning, in which the predictor can be built on the fly, while new data is gathered. We believe that an ever more realistic paradigm is active learning, a setup in which the learner actively requests data (possibly facing constraints, such as storage, velocity, cost, etc.) and adapts its queries to optimize its performance. The 3-years objective of 6PAC (where 6 stands for Sequential, Active, Efficient, Structured, Ideal, Safe - the six research directions we intend to contribute to) is to pave the way to new PAC generalization and sample-complexity upper and lower bounds beyond batch learning. Our ambition is to contribute to several learning setups, ranging from sequential learning (where data streams are collected) to adaptive and active learning (where data streams are requested by the learning algorithm).

9.4.2. Inria International Partners

9.4.2.1. Declared Inria International Partners

A byproduct of Benjamin Guedj's sabbatical position at University College London (UCL) since Dec 2018 is a strengthened link between UCL and Inria. DGDS has established contact with UCL President in April 2019 and a MoU has been signed between UCL and Inria in December 2019. A research group (known as Inria@UCL) has been established by Benjamin Guedj within UCL, Department for Computer Science, Centre for Artificial Intelligence. Inria@UCL initiative is expected to grow in 2020 and possibly evolve into a joint team or more. A strategic partnership between Inria and UCL will be explored in 2020.

SIMERGE

Title: Statistical Inference for the Management of Extreme Risks and Global Epidemiology

International Partner (Institution - Laboratory - Researcher):

UGB (Senegal) - LERSTAD - Abdou Ka Diongue

Serge Iovleff and Sophie Dabo-Niang are associated members of SIMERGE.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Mihai Cucuringu (University of Oxford) visited Hemant Tyagi in January 2019 for a research visit of 1 week.
- Martin Wahl (Humboldt Universität from Berlin) visited Alain Celisse in March 2019 for a research visit of 1 week and November 2019 for a research visit of 1 week.
- Apoorv Vikram Singh is currently visiting Hemant Tyagi to work on a research project which is jointly supervised by Hemant Tyagi and Mihai Cucuringu (University of Oxford). The duration of the visit is 4 months (October 1, 2019 - January 31, 2020) and is partly funded by the Alan Turing Institute, London.
- Abdou Kâ Diongue visited Serge Iovleff in June 2019 for one month.

9.5.2. Visits to International Teams

9.5.2.1. Sabbatical programme

Since Dec 2018, Benjamin Guedj is on sabbatical at University College London (UCL). He is a PI of the UCL Centre for Artificial Intelligence (UCL AI) and a visiting researcher at the Alan Turing Institute. This has led to the Inria@UCL initiative, see supra.

9.5.2.2. Research Stays Abroad

- Sophie Dabo-Niang has visited University of Kuala Lumpur, Malaysia in August 2019 and University of Mohamed V, Morroco in December 2019.
- Serge Iovleff has visited University Gaston Berger, Senegal in February 2019 and gave a course entitled “Introduction to Statistical Learning”.
- Hemant Tyagi visited Mihai Cucuringu and Benjamin Guedj at the Alan Turing Institute, UK from in October 2019.
- Alain Celisse visited Markus Reißand Martin Wahl at the Humboldt Universität, Germany in March and December 2019.
- Alain Celisse visited Benjamin Guedj at the University College London, UK in February-March and July-August 2019.
- Pascal Germain visited Benjamin Guedj at University College London, UK on several occasions totalling about 1.5 month in 2019.
- Cristian Preda visited Amarioarei Alexandru at University of Bucharest on several occasions totalling about 1 week in 2019.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. General Chair, Scientific Chair

- Hemant Tyagi and Pascal Germain are the organizers of the MODAL team scientific seminar.
- Sophie Dabo-Niang is the co-organizer of:

- Session “Recent non-parametric approaches: Applications to envi- ronmental, hydrological, oceanological and economic data analyzes” of the ISI conference (62nd World Statistics Congress), 18-23, August, 2019, Kuala Lumpur, Malaysia.
- Session “Modeling dependence through graphical models”, CMStatistics, 14-16 December 2019, London.

- Christophe Biernacki is a co-organizer (and the chair) of the special session on co-clustering called “Co-clustering: model-based or model-free approaches” at the 62nd ISI World Statistics Congress 2019 in Malaysia [45]
- Alain Celisse is the organizer and the chair of two sessions on change-points detection and early stopping rules at ERCIM 2019, London.
- Vincent Vandewalle organized a session on Model-based and multivariate functional data at CRoNoS & MDA 2019.

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

- Sophie Dabo-Niang was a member of the scientific program committee of: CRONOS-MDA2019, 14-16 April 2019, Limassol, Cyprus and African Econometric Society Conference, 11-13 July 2019, Rabat, Morocco.
- Benjamin Guedj has been a PC member for UAI 2019 and IJCAI 2019.

10.1.2.2. Reviewer

- Hemant Tyagi acted as a reviewer for journals (Foundations of Computational Mathematics, Journal of Machine Learning Research, SIAM Journal on Scientific Computing, Advances in Data Analysis and Classification) and a conference (NeurIPS 2019).
- Pascal Germain acted as a reviewer for a journal (IEEE Transactions on Pattern Analysis and Machine Intelligence) and conferences (ICML 2019, ICLR 2019, NeurIPS 2019, COLT 2019, PHM 2019).
- Benjamin Guedj served as a reviewer for journals (JMLR, Neurocomputing, Journal of international research on robotics,...) and conferences (ICML 2019, ICLR 2019, NeurIPS 2019, COLT 2019, IJCAI 2019, UAI 2019, AISTATS 2019 & 2020).
- Alain Celisse acted as a reviewer for journals (IEEE Transactions on Pattern Analysis and Machine Intelligence, Annals of Statistics, JMLR,...) and conferences (ALT 2019, COLT 2019).
- Sophie Dabo-Niang acted as a reviewer for journals Annals of Statistics, JASA, Journal of Nonparametric statistics, TEST, METRIKA.
- Christophe Biernacki acted as a reviewer for CSDA, ESWA, JCGS, JMIV.
- Serge Iovleff acted as a reviewer for CSDA and Journal of Statistics and Computing.
- Vincent Vandewalle acted as a reviewer for ADAC, Statistics in Medicine, COST.
- Cristian Preda is a reviewer for Bernoulli, MCAP, ADAC.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Christophe Biernacki is an Associate Editor of the North-Western European Journal of Mathematics (NWEJM) and for Frontiers on the topic “Computational Methods for Data Analytics”. He is also a Guest Editor for the Special Issue on Innovations in Model-Based Clustering and Classification of the journal Advances Data Analysis and Classification (ADAC).
- Cristian Preda is an associate editor for Methodological and Computing in applied probability (MCAP) and for Journal of Mathematical and Computer Science.

- Serge Iovleff is member of the Editorial Board of Astrostatistics (specialty section of Frontiers in Astronomy and Space Sciences)

10.1.4. Invited Talks

- Christophe Biernacki:
  - 3rd International Conference on Econometrics and Statistics (EcoSta 2019), Taiwan, June 25, 2019 [42]
  - 12th Scientific Meeting Classification and Data Analysis Group (CLADAG 2019), Italy, September 12, 2019 [41]
  - APSEM 2019 (Apprentissage et SEMantique) : écosystèmes pour la science ouverte et recherche par les données, Toulouse [39]
  - Séminary of Probability and Statistics of the University of Nice - Sophia Antopolis, France, November 19, 2019
- Hemant Tyagi:
  - Probability-Statistics seminar, Laboratoire de Mathematiques, Universite de Franche-Comte, Besancon, France, June 13, 2019
  - Probability-Statistics seminar, Université de Lille, France, September 18, 2019
- Pascal Germain:
  - Journées de la statistique, Nancy, France, June 6, 2019
- Benjamin Guedj:
  - Speaker (with John Shawe-Taylor) for a plenary tutorial at ICML 2019 (June 2019).
  - Several seminars in the UK (Gatsby unit, UCL CSML, UCL stats, ElementAI).
- Sophie Dabo-Niang:
  - 62nd World Statistics Congress, August, 18-23, 2019, Kuala Lumpur, Malaysia.
  - 7th Statistical Meeting of Avignon-Marseille, 14 June, 2019, Avignon, France.
  - Workshop on Non-Parametric Statistics, 18th, September, 2019, Grenoble, France.
  - CFIES 2019, 24-27 September 2019, Strasbourg, France.
  - Symposium of young Senegalese researchers in Pured and Applied Mathematics, December, 16-19, 2019, Mbour, Senegal.
- Cristian Preda:
  - Conference on scan statistics, IMS China, Dalian, 6-10 july, 2019.
  - Conference on categorical functional data at Romanian statistical society, 10-11 May 2019.
- Alain Celisse:
  - Session on Model selection at ERCIM 2018, London.
  - Several seminars in France (Nantes, Paris,...) and abroad WIAS Institute of Berlin.
- Vincent Vandewalle:
  - Session clustering categorical and mixed type data at IFCS 2019, Thessaloniki.

10.1.5. Leadership within the Scientific Community

- Benjamin Guedj and Pascal Germain are founding members of the Machine Learning and Artificial Intelligence group (MALIA) of the French statistical association (SFdS).
- Benjamin Guedj is the new French representative at the board of ECAS (since Nov 2019).
- Since 2017, Benjamin Guedj is an elected member of the board of the Statistical French Society (SFdS).

10.1.6. Scientific Expertise

Christophe Biernacki reviewed one project as an expert for the ANR and for Innoviris (Brussels).

10.1.7. Research Administration

- Sophie Dabo-Niang is the person in charge of the MeQAME axis of laboratory LEM, CNRS 9221.
- Christophe Biernacki is Scientific Head of the Inria Lille center since June 2017.
- Benjamin Guedj has been the list head for the election of the Evaluation Committee in June 2019. He has been an elected member of CE since 2017 and is a member of its executive board since Sept 2019.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

- Pascal Germain taught

  Master: Introduction aux réseaux de neurones, 15 heures, M2, Université de Lille, France
- Sophie Dabo-Niang is teaching

  Master: Spatial Statistics, 24h, M2, Université de Lille, France

  Master: Advanced Statistics, 24h, M2, Université de Lille, France

  Master: Multivariate Data Analyses, 24h, M2, Université de Lille, France

  Licence: Probability, 24h, L2, Université de Lille, France

  Licence: Multivariate Statistics, 24h, L3, Université de Lille, France
- Cristian Preda is teaching

  Polytech'Lille engineer school: Linear Models, 48h.

  Polytech'Lille engineer school: Advanced statistics, 48h.

  Polytech'Lille engineer school: Biostatistics, 10h.

  Polytech'Lille engineer school: Supervised clustering, 24h. France
- Christophe Biernacki is teaching

  New Master Data Science: Statistics, 24h, M1, Université de Lille, France
- Benjamin Guedj is teaching

  Advanced machine learning (M2, 6h), University College London, UK
- Serge Iovleff is teaching

  Licence: Analyse et méthodes numériques, 56h, Université de Lille, DUT Informatique

  Licence: R.O. et aide à la décision, 32h, Université de Lille, DUT Informatique
- Vincent Vandewalle is teaching

  Licence: Probability, 60h, Université de Lille, DUT STID

  Licence: Case study in statistics, 45h, Université de Lille, DUT STID

  Licence: R programming, 45h, Université de Lille, DUT STID

  Licence: Supervised clustering, 32h, Université de Lille, DUT STID

  Licence: Analysis, 24h, Université de Lille, DUT STID

10.2.2. Supervision

10.2.2.1. PhD defense:

- Maxime Baelde, September 20th 2019, supervised by Christophe Biernacki and Raphaël Greff on "Generative models for the classification and separation of real-time sound sources" [11].
- Anne-Lise Bedenel, April 1st 2019, supervised by Christophe Biernacki and Laetitia Jourdan on "Matching descriptors evolving over time: application to insurance comparison" [12].

- Adrien Ehrhardt, September 3rd 2019, supervised by Christophe Biernacki, Philippe Heinrich and Vincent Vandewalle on "Formalization and study of statistical problems in Credit Scoring" [13].

10.2.2.2. PhD in progress:

- Felix Biggs, Generative models and kernels, University College London, Sep 2019, Benjamin Guedj.
- Antoine Vendeville, Learning on graph to stop the propagation of fake news, University College London, Sep 2019, Benjamin Guedj.
- Luxin Zhang, Domain adaptation from a pre-trained source model – Application to fraud detection in electronic payments, February 2019, Christophe Biernacki, Pascal Germain, Yacine Kessac.
- Paul Viallard, Interpreting representation learning through PAC-Bayes theory, September 2019, Amaury Habrard, Emilie Morvant, Pascal Germain.
- Dang Khoi Pham, Planning and re-planning of nurses in an oncology department using a multi-objective and interdisciplinary approach, September 2016, Sophie Dabo-Niang.
- Solange Doumun, Performance evaluation and contribution to the development of multispectral image analysis strategies for automatic and rapid diagnosis of malaria, December 2018, Sophie Dabo-Niang.
- Alaa Ali Ayad, Statistical modeling of large spatial data and its applications in health, September 2018, Sophie Dabo-Niang.
- Wilfried Heyse, Prise en compte de la structure temporelle dans l'analyse statistique de données protéomiques à haut débit, October 2019, Christophe Bauters, Guillemette Marot and Vincent Vandewalle.
- Margot Selosse, October 2017, Christophe Biernacki and Julien Jacques.
- Filippo Antonazzo, October 2019, Christophe Biernacki and Christine Keribin.
- Yaroslav Averyanov, September 2016, Early stopping and filters estimators, Alain Celisse.

10.2.3. Juries

- Sophie Dabo-Niang acted as a reviewer and an examinator for PhD theses.
- Christophe Biernacki acted as a reviewer for the following PhD theses: Keefe Murphy December 17th 2019 (UCD Dublin), Jocelyn Chauvet April 19th 2019 (University of Montpellier).
- Christophe Biernacki acted as an examinator for the following HdR defenses: Madalina Olteanu December 10th 2019 (University Paris 1), Servane Gey February 7th 2019 (University Paris 5).
- Christophe Biernacki participated in the following juries of recruitment: MC jury université d'Avignon, Jury CR Lille, Jury DR Inria.
- Christophe Biernacki participated in the following juries of recruitment: MCF jury Université de Nice May 2019, Jury CR Paris 2019, Jury CR Inria (nationwide) 2019.
- Vincent Vandewalle participated in an MC jury Université Versailles Saint Quentin May 2019.
- Alain Celisse acted as a reviewer for the HdR defense of Zoltan Szabo in December 2019 at Ecole Polytechnique, Palaiseau.
- Cristian Preda acted as a referee for the PhD thesis of Amandine Schmutz, Université de Lyon 2, November 15, 2019.
- Cristian Preda acted as a referee for the HDR defense of Raluca Vernic, July 10, 2019, Universitatea Ovidiu, Constanta, Romania.

10.3. Popularization

10.3.1. Internal or external Inria responsibilities

Christophe Biernacki acts as a member of the working group calculation for Inria.

10.3.2. Interventions

- Pascal Germain participated in the school activity “Jouer à débattre” of the organization “L’arbre des connaissances”, as an expert in artificial intelligence (Lycée EIC de Tourcoing, February 7, 2019).
- Pascal Germain gave a vulgarization talk about neural networks and deep learning at the “Journée de l’Enseignement de l’Informatique et de l’Algorithmique” (Université de Lille, March 6, 2019).
- Christophe Biernacki participated in a round table “Deep Learning for PHM: Opportunities and Challenges”, The 10th Prognostics and System Health Management Conference Paris, France, May 2 - May 5, 2019.
- Christophe Biernacki gave a talk at the Forum Teratec in June 12nd 2019, Ecole Polytechnique, Palaiseau with Margot Correard (DiagRAMS) on Predictive maintenance solution without additional sensors [40].
- Cristian Preda was invited as a speaker at IoT Week on predictive maintenance, Méaulte, December 4, 2019.
- Cristian Preda was invited as a speaker at IT Tour by Le Monde Informatique on artificial inteligence, November 14, 2019.

11. Bibliography

Major publications by the team in recent years

[1] P. ALQUIER, B. GUEDJ.*Simpler PAC-Bayesian Bounds for Hostile Data*, in "Machine Learning", 2018 [*DOI :* 10.1007/S10994-017-5690-0], https://hal.inria.fr/hal-01385064

[2] P. BATHIA, S. IOVLEFF, G. GOVAERT.*An R Package and C++ library for Latent block models: Theory, usage and applications*, in "Journal of Statistical Software", 2016, https://hal.archives-ouvertes.fr/hal-01285610

[3] C. BIERNACKI, A. LOURME.*Unifying Data Units and Models in (Co-)Clustering*, in "Advances in Data Analysis and Classification", May 2018, vol. 12, n^{o} 41, https://hal.archives-ouvertes.fr/hal-01653881

[4] A. CELISSE.*Optimal cross-validation in density estimation with the L2-loss*, in "The Annals of Statistics", 2014, vol. 42, n^{o} 5, p. 1879–1910, https://hal.archives-ouvertes.fr/hal-00337058

[5] S. DABO-NIANG, C. TERNYNCK, A.-F. YAO.*Nonparametric prediction in the multivariate spatial context*, in "Journal of Nonparametric Statistics", 2016, vol. 28, n^{o} 2, p. 428-458 [*DOI :* 10.1080/10485252.2016.01.007], https://hal.inria.fr/hal-01425932

[6] J. DUBOIS, V. DUBOIS, H. DEHONDT, P. MAZROOEI, C. MAZUY, A. A. SÉRANDOUR, C. GHEERAERT, P. GUILLAUME, E. BAUGÉ, B. DERUDAS, N. HENNUYER, R. PAUMELLE, G. MAROT, J. S. CARROLL, M. LUPIEN, B. STAELS, P. LEFEBVRE, J. EECKHOUTE.*The logic of transcriptional regulator recruitment architecture at cis -regulatory modules controlling liver functions*, in "Genome Research", June 2017, vol. 27, n^{o} 6, p. 985–996 [*DOI :* 10.1101/GR.217075.116], https://hal.archives-ouvertes.fr/hal-01647846

[7] G. LETARTE, P. GERMAIN, B. GUEDJ, F. LAVIOLETTE.*Dichotomize and Generalize: PAC-Bayesian Binary Activated Deep Neural Networks*, in "NeurIPS 2019", Vancouver, Canada, December 2019, https://hal.inria.fr/hal-02139432

[8] M. MARBAC, C. BIERNACKI, V. VANDEWALLE.*Model-based clustering of Gaussian copulas for mixed data*, in "Communications in Statistics - Theory and Methods", December 2016, https://hal.archives-ouvertes.fr/hal-00987760

[9] C. PREDA, A. DERMOUNE.*Parametrizations, fixed and random effects*, in "Journal of Multivariate Analysis", February 2017, vol. 154, p. 162–176 [*DOI :* 10.1016/J.JMVA.2016.11.001], https://hal.archives-ouvertes.fr/hal-01655461

[10] H. TYAGI, J. VYBIRAL.*Learning general sparse additive models from point queries in high dimensions*, in "Constructive Approximation", January 2019, https://hal.inria.fr/hal-02379404

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] M. BAELDE.*Generative models for the classification and separation of real-time sound sources*, Université de Lille 1, September 2019, https://hal.archives-ouvertes.fr/tel-02399081

[12] A.-L. BEDENEL.*Matching descriptors evolving over time: application to insurance comparison*, Université de Lille I, April 2019, https://hal.archives-ouvertes.fr/tel-02399068

[13] A. EHRHARDT.*Formalization and study of statistical problems in Credit Scoring : Reject inference, discretization and pairwise interactions, logistic regression trees*, Université de Lille, September 2019, https://hal.archives-ouvertes.fr/tel-02302691

Articles in International Peer-Reviewed Journal

[14] 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

[15] S. ARLOT, A. CELISSE, Z. HARCHAOUI.*A Kernel Multiple Change-point Algorithm via Model Selection*, in "Journal of Machine Learning Research", December 2019, vol. 20, n^{o} 162, p. 1–56, https://arxiv.org/abs/1202.3878 , https://hal.archives-ouvertes.fr/hal-00671174

[16] M. BAELDE, C. BIERNACKI, R. GREFF.*Real-Time Monophonic and Polyphonic Audio Classification from Power Spectra*, in "Pattern Recognition", August 2019, vol. 92, p. 82-92 [*DOI :* 10.1016/J.PATCOG.2019.03.017], https://hal.archives-ouvertes.fr/hal-01834221

[17] M. BERNARDINI, A. BROSSA, G. CHINIGO, G. GROLEZ, G. TRIMAGLIO, L. ALLART, A. HULOT, G. MAROT, T. GENOVA, A. JOSHI, V. MATTOT, G. FROMONT, L. MUNARON, B. BUSSOLATI, N. PREVARSKAYA, A. FIORIO PLA, D. GKIKA.*Transient Receptor Potential Channel Expression Signatures in Tumor-Derived Endothelial Cells: Functional Roles in Prostate Cancer Angiogenesis*, in "Cancers", July 2019, vol. 11, n^{o} 7, 956 [*DOI :* 10.3390/CANCERS11070956], https://hal.archives-ouvertes.fr/hal-02404061

[18] S. CURCEAC, C. TERNYNCK, T. B. OUARDA, F. CHEBANA, S. DABO-NIANG.*Short-term air temperature forecasting using Nonparametric Functional Data Analysis and SARMA models*, in "Environmental Modelling and Software", January 2019, vol. 111, p. 394-408 [*DOI :* 10.1016/J.ENVSOFT.2018.09.017], https://hal.inria.fr/hal-01948928

[19] M. Cuvelliez, V. Vandewalle, M. Brunin, O. Beseme, A. Hulot, P. De Groote, P. Amouyel, C. Bauters, G. Marot, F. Pinet.*Circulating proteomic signature of early death in heart failure patients with reduced ejection fraction - Short title: Proteomic signature of early death in heart failure patients*, in "Scientific Reports", 2019, forthcoming, https://hal.inria.fr/hal-02400814

[20] M. Cuvelliez, V. Vandewalle, M. Brunin, O. Beseme, A. Hulot, P. De Groote, P. Amouyel, C. Bauters, G. Marot, F. Pinet.*Circulating proteomic signature of early death in heart failure patients with reduced ejection fraction*, in "Scientific Reports", December 2019, vol. 9, 19202 [*DOI :* 10.1038/s41598-019-55727-1], https://hal.archives-ouvertes.fr/hal-02414293

[21] S. Dabo-Niang, S. Curceac, C. Ternynck, T. B. Ouarda, F. Chebana, S. D. Niang.*Short-term air temperature forecasting using Nonparametric Functional Data Analysis and SARMA models*, in "Environmental Modelling and Software", January 2019, vol. 111, p. 394-408 [*DOI :* 10.1016/j.envsoft.2018.09.017], https://hal.inria.fr/hal-02334991

[22] S. Dabo-Niang, B. Thiam.*Kernel regression estimation with errors-in-variables for random fields*, in "Afrika Matematika", 2019 [*DOI :* 10.1007/s13370-019-00654-7], https://hal.inria.fr/hal-02334993

[23] F. Dewez, V. Montmirail.*Decrypting the Hill Cipher via a Restricted Search over the Text-Space*, in "Linköping Electronic Conference Proceedings", June 2019, https://hal.univ-cotedazur.fr/hal-02271395

[24] A. Eftekhari, J. Tanner, A. Thompson, B. Toader, H. Tyagi.*Sparse non-negative super-resolution — simplified and stabilised*, in "Applied and Computational Harmonic Analysis", August 2019 [*DOI :* 10.1016/j.acha.2019.08.004], https://hal.inria.fr/hal-02379445

[25] P. Germain, A. Habrard, F. Laviolette, E. Morvant.*PAC-Bayes and Domain Adaptation*, in "Neurocomputing", 2020, vol. 379, p. 379-397, https://arxiv.org/abs/1707.05712 [*DOI :* 10.1016/j.neucom.2019.10.105], https://hal.archives-ouvertes.fr/hal-01563152

[26] A. Goyal, E. Morvant, P. Germain, M.-R. Amini.*Multiview Boosting by Controlling the Diversity and the Accuracy of View-specific Voters*, in "Neurocomputing", 2019, https://arxiv.org/abs/1808.05784 , forthcoming [*DOI :* 10.1016/j.neucom.2019.04.072], https://hal.archives-ouvertes.fr/hal-01857463

[27] D. A. Mogilenko, J. Haas, L. L'homme, S. Fleury, S. Quemener, M. Levavasseur, C. Becquart, J. Wartelle, A. Bogomolova, L. Pineau, O. Molendi-Coste, S. Lancel, H. Dehondt, C. Gheeraert, A. Melchior, C. Dewas, A. Nikitin, S. Pic, N. Rabhi, J.-S. Annicotte, S. Oyadomari, T. Velasco-Hernandez, J. Cammenga, M. Foretz, B. Viollet, M. Vukovic, A. Villacreces, K. Kranc, P. Carmeliet, G. Marot, A. Boulter, S. J. Tavernier, L. Berod, M. P. Longhi, C. Paget, S. Janssens, D. Staumont-Sallé, E. Aksoy, B. Staels, D. Dombrowicz.*Metabolic and innate immune cues merge into a specific inflammatory response via unfolded protein-response (UPR)*, in "Cell", May 2019, vol. 177, n^{o} 5, p. 1201-1216.e19, Erratum in : Metabolic and Innate Immune Cues Merge into a Specific Inflammatory Response via the UPR. [Cell. 2019], forthcoming [*DOI :* 10.1016/j.cell.2019.03.018], https://www.hal.inserm.fr/inserm-02084447

[28] M. Selosse, J. Jacques, C. Biernacki, F. Cousson-Gélie.*Analysing a quality of life survey using a co-clustering model for ordinal data and some dynamic implications*, in "Journal of the Royal Statistical Society: Series C Applied Statistics", July 2019, https://hal.archives-ouvertes.fr/hal-01643910

[29] M. SELOSSE, J. JACQUES, C. BIERNACKI.*Model-based co-clustering for mixed type data*, in "Computational Statistics and Data Analysis", 2020, vol. 144, 106866 [*DOI :* 10.1016/J.CSDA.2019.106866], https://hal.archives-ouvertes.fr/hal-01893457

[30] H. TYAGI, J. VYBIRAL.*Learning general sparse additive models from point queries in high dimensions*, in "Constructive Approximation", January 2019, https://hal.inria.fr/hal-02379404

Invited Conferences

[31] C. BIERNACKI, G. CELEUX, J. JOSSE, F. LAPORTE.*Dealing with missing data in model-based clustering through a MNAR model*, in "CRoNos & MDA 2019 - Meeting and Workshop on Multivariate Data Analysis and Software", Limassol, Cyprus, April 2019, https://hal.inria.fr/hal-02103347

International Conferences with Proceedings

[32] M. CUCURINGU, P. DAVIES, A. GLIELMO, H. TYAGI.*SPONGE: A generalized eigenproblem for clustering signed networks*, in "AISTATS", Okinawa, Japan, April 2019, https://hal.inria.fr/hal-02379505

[33] B. GUEDJ, J. RENGOT.*Non-linear aggregation of filters to improve image denoising*, in "Computing Conference 2020", London, United Kingdom, July 2020, https://hal.inria.fr/hal-02086856

[34] J. KLEIN, M. ALBARDAN, B. GUEDJ, O. COLOT.*Decentralized learning with budgeted network load using Gaussian copulas and classifier ensembles*, in "ECML-PKDD, Decentralized Machine Learning at the Edge Workshop", Wurzburg, Germany, September 2019, https://arxiv.org/abs/1804.10028 , https://hal.archives-ouvertes.fr/hal-01779989

[35] G. LETARTE, P. GERMAIN, B. GUEDJ, F. LAVIOLETTE.*Dichotomize and Generalize: PAC-Bayesian Binary Activated Deep Neural Networks*, in "NeurIPS 2019", Vancouver, Canada, December 2019, https://hal.inria.fr/hal-02139432

[36] G. LETARTE, E. MORVANT, P. GERMAIN.*Pseudo-Bayesian Learning with Kernel Fourier Transform as Prior*, in "The 22nd International Conference on Artificial Intelligence and Statistics", Naha, Japan, Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS) 2019,, 2019, https://arxiv.org/abs/1810.12683 , https://hal.archives-ouvertes.fr/hal-01908555

[37] Z. MHAMMEDI, P. GRÜNWALD, B. GUEDJ.*PAC-Bayes Un-Expected Bernstein Inequality*, in "NeurIPS 2019", Vancouver, Canada, December 2019, https://arxiv.org/abs/1905.13367 , https://hal.inria.fr/hal-02401295

[38] V. SHALAEVA, A. FAKHRIZADEH ESFAHANI, P. GERMAIN, M. PETRECZKY.*Improved PAC-Bayesian Bounds for Linear Regression*, in "Thirty-Fourth AAAI Conference on Artificial Intelligence", New York, United States, February 2020, https://arxiv.org/abs/1912.03036 , https://hal.inria.fr/hal-02396556

Conferences without Proceedings

[39] C. BIERNACKI.*MASSICCC: A SaaS Platform for Clustering and Co-Clustering of Mixed Data*, in "APSEM 2019 ((Apprentissage et SEMantique)) : éco-systèmes pour la science ouverte et recherche par les données", Toulouse, France, October 2019, https://hal.archives-ouvertes.fr/hal-02399180

[40] C. BIERNACKI, M. CORRÉARD.*Predictive maintenance solution without additional sensors*, in "Forum TERATEC", Palaiseau, France, June 2019, https://hal.archives-ouvertes.fr/hal-02399046

[41] C. BIERNACKI, A. LOURME.*Unifying Data Units and Models in (Co-)Clustering*, in "CLADAG 2019 - 12th Scientific Meeting Classification and Data Analysis Group", Cassino, Italy, September 2019, https://hal.archives-ouvertes.fr/hal-02398982

[42] C. BIERNACKI, M. MARBAC, V. VANDEWALLE.*Gaussian Based Visualization of Gaussian and Non-Gaussian Based Clustering*, in "3rd International Conference on Econometrics and Statistics (EcoSta 2019)", Taichung, Taiwan, June 2019, https://hal.archives-ouvertes.fr/hal-02398999

[43] 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

[44] L. GAUTHERON, P. GERMAIN, A. HABRARD, G. LETARTE, E. MORVANT, M. SEBBAN, V. ZANTEDESCHI.*Revisite des "random Fourier features" basée sur l'apprentissage PAC-Bayésien via des points d'intérêts*, in "CAp 2019 - Conférence sur l'Apprentissage automatique", Toulouse, France, July 2019, https://hal.archives-ouvertes.fr/hal-02148600

[45] C. KERIBIN, C. BIERNACKI.*Co-clustering: model based or model free approaches*, in "ISI WSC 2019 - 62nd ISI World Statistics Congress", Kuala Lumpur, Malaysia, August 2019, https://hal.archives-ouvertes.fr/hal-02399031

[46] C. KERIBIN, C. BIERNACKI.*Le modèle des blocs latents, une méthode régularisée pour la classification en grande dimension*, in "JdS 2019 - 51èmes Journées de Statistique de la SFdS", Nancy, France, June 2019, https://hal.archives-ouvertes.fr/hal-02391379

[47] F. LAPORTE, C. BIERNACKI, G. CELEUX, J. JOSSE.*Modèles de classification non supervisée avec données manquantes non au hasard*, in "JdS 2019 - 51e journées de statistique de la Sfds", Nancy, France, June 2019, https://hal.archives-ouvertes.fr/hal-02398984

[48] M. MARBAC-LOURDELLE, C. BIERNACKI, V. VANDEWALLE.*Gaussian Based Visualization of Gaussian and Non-Gaussian Based Clustering*, in "SPSR 2019", Bucarest, Romania, April 2020, https://hal.archives-ouvertes.fr/hal-02400486

[49] V. VANDEWALLE, C. TERNYNCK, G. MAROT.*Linking different kinds of Omics data through a model-based clustering approach*, in "IFCS 2019", Thessalonique, Greece, August 2019, https://hal.archives-ouvertes.fr/hal-02400525

[50] P. VIALLARD, R. EMONET, P. GERMAIN, A. HABRARD, E. MORVANT.*Interpreting Neural Networks as Majority Votes through the PAC-Bayesian Theory*, in "Workshop on Machine Learning with guarantees @ NeurIPS 2019", Vancouver, Canada, 2019, https://hal.archives-ouvertes.fr/hal-02335762

[51] L. ZHANG, C. BIERNACKI, P. GERMAIN, Y. KESSACI.*Domain Adaptation from a Pre-trained Source Model: Application on fraud detection tasks*, in "12th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2019)", London, United Kingdom, December 2019, https://hal.archives-ouvertes.fr/hal-02399003

Scientific Books (or Scientific Book chapters)

[52] S. Dabo-Niang, S. MANOU-ABI, S. Jean-Jacques.*Mathematical Modeling and Study of Random or Deterministic Phenomena,*, Wiley, 2020, https://hal.inria.fr/hal-02334997

Other Publications

[53] H. Alawieh, N. Wicker, C. Biernacki.*Projection under pairwise distance controls*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01420662

[54] C. Biernacki, M. Marbac, V. Vandewalle.*Gaussian Based Visualization of Gaussian and Non-Gaussian Based Clustering*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01949155

[55] S. Chrétien, B. Guedj.*Revisiting clustering as matrix factorisation on the Stiefel manifold*, March 2019, https://arxiv.org/abs/1903.04479 - working paper or preprint, https://hal.inria.fr/hal-02064396

[56] S. Chrétien, H. Tyagi.*Multi-kernel unmixing and super-resolution using the Modified Matrix Pencil method*, November 2019, working paper or preprint, https://hal.inria.fr/hal-02379598

[57] V. Cohen-Addad, B. Guedj, V. Kanade, G. Rom.*Online k-means Clustering*, December 2019, https://arxiv.org/abs/1909.06861 - 11 pages, 1 figure, https://hal.inria.fr/hal-02401290

[58] 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

[59] 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

[60] M. Cucuringu, H. Tyagi.*Provably robust estimation of modulo 1 samples of a smooth function with applications to phase unwrapping*, November 2019, working paper or preprint, https://hal.inria.fr/hal-02379573

[61] A. D'Aspremont, M. Cucuringu, H. Tyagi.*Ranking and synchronization from pairwise measurements via SVD*, October 2019, https://arxiv.org/abs/1906.02746 - 42 pages, 9 figures, https://hal.archives-ouvertes.fr/hal-02340372

[62] A. Ehrhardt, C. Biernacki, V. Vandewalle, P. Heinrich.*Feature quantization for parsimonious and interpretable predictive models*, March 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01949135

[63] M. P. B. Gallaugher, C. Biernacki, P. D. McNicholas.*Parameter-Wise Co-Clustering for High-Dimensional Data*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01862824

[64] L. Gautheron, P. Germain, A. Habrard, E. Morvant, M. Sebban, V. Zantedeschi.*Learning Landmark-Based Ensembles with Random Fourier Features and Gradient Boosting*, June 2019, https://arxiv.org/abs/1906.06203 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02148618

[65] B. GUEDJ, B. S. DESIKAN.*Kernel-Based Ensemble Learning in Python*, December 2019, https://arxiv.org/abs/1912.08311 - 11 pages, https://hal.inria.fr/hal-02443097

[66] B. GUEDJ.*A Primer on PAC-Bayesian Learning*, May 2019, working paper or preprint, https://hal.inria.fr/hal-01983732

[67] B. GUEDJ, L. LI.*Sequential Learning of Principal Curves: Summarizing Data Streams on the Fly*, May 2019, working paper or preprint, https://hal.inria.fr/hal-01796011

[68] B. GUEDJ, L. PUJOL.*Still no free lunches: the price to pay for tighter PAC-Bayes bounds*, December 2019, https://arxiv.org/abs/1910.04460 - working paper or preprint, https://hal.inria.fr/hal-02401286

[69] F. LAPORTE, C. BIERNACKI, G. CELEUX, J. JOSSE.*Model-based clustering with missing not at random data. Missing mechanism*, July 2019, Working Group on Model-Based Clustering Summer Session, Poster, https://hal.archives-ouvertes.fr/hal-02398987

[70] G. MAZO, Y. AVERYANOV.*Constraining kernel estimators in semiparametric copula mixture models*, March 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01774629

[71] K. NOZAWA, P. GERMAIN, B. GUEDJ.*PAC-Bayesian Contrastive Unsupervised Representation Learning*, December 2019, https://arxiv.org/abs/1910.04464 - working paper or preprint, https://hal.inria.fr/hal-02401282

[72] M. SELOSSE, J. JACQUES, C. BIERNACKI.*Textual data summarization using the Self-Organized Co-Clustering model*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02115294

[73] M. SELOSSE, J. JACQUES, C. BIERNACKI.*ordinalClust: an R package for analyzing ordinal data*, December 2019, working paper or preprint, https://hal.inria.fr/hal-01678800

[74] S. N. SYLLA, S. DABO-NIANG, C. LOUCOUBAR.*Functional data analysis of parasite densities in the Senegalese villages of Dielmo and NDiop*, October 2019, working paper or preprint, https://hal.inria.fr/hal-02335001

[75] J. ZHANG, E. T. BARR, B. GUEDJ, M. HARMAN, J. SHAWE-TAYLOR.*Perturbed Model Validation: A New Framework to Validate Model Relevance*, May 2019, working paper or preprint, https://hal.inria.fr/hal-02139208

Project-Team RAPSODI

Reliable numerical approximations of dissipative systems.

IN COLLABORATION WITH: Laboratoire Paul Painlevé (LPP)

IN PARTNERSHIP WITH:
Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Numerical schemes and simulations

Table of contents

Project-Team RAPSODI

Creation of the Team: 2015 August 01, updated into Project-Team: 2017 November 01

Keywords:

Computer Science and Digital Science:

A6. - Modeling, simulation and control
A6.1. - Methods in mathematical modeling
A6.1.1. - Continuous Modeling (PDE, ODE)
A6.1.4. - Multiscale modeling
A6.1.5. - Multiphysics modeling
A6.2. - Scientific computing, Numerical Analysis & Optimization
A6.2.1. - Numerical analysis of PDE and ODE

Other Research Topics and Application Domains:

B3. - Environment and planet
B3.3. - Geosciences
B3.3.1. - Earth and subsoil
B3.4. - Risks
B3.4.2. - Industrial risks and waste
B4. - Energy
B4.2. - Nuclear Energy Production
B4.2.1. - Fission

1. Team, Visitors, External Collaborators

Research Scientists

Clément Cancès [Inria, Researcher, HDR]
Maxime Herda [Inria, Researcher]
Simon Lemaire [Inria, Researcher]

Faculty Members

Claire Chainais-Hillairet [Team leader, Université de Lille, Professor, HDR]
Benoît Merlet [Université de Lille, Professor, HDR]
Caterina Calgaro [Université de Lille, Associate Professor]
Ingrid Lacroix-Violet [Université de Lille, Associate Professor, HDR]
Thomas Rey [Université de Lille, Associate Professor]

External Collaborators

Claire Colin [Université de Lille, from Feb 2019 until Jul 2019]
Emmanuel Creusé [Université Polytechnique Hauts-de-France, Professor, HDR]

PhD Students

Sabrina Bassetto [IFPEN, PhD Student]
Claire Colin [Université de Lille, PhD Student, until Jan 2019]
Benoît Gaudeul [École Normale Supérieure de Rennes, PhD Student]
Antoine Zurek [Université de Lille, PhD Student, until Sep 2019]

Post-Doctoral Fellow

Florent Chave [Inria, Post-Doctoral Fellow]

Administrative Assistant

Amélie Supervielle [Inria]

2. Overall Objectives

2.1. Overall Objectives

Together with the diffusion of scientific computing, there has been a recent and impressive increase of the demand for numerical methods. The problems to be addressed are everyday more complex and require specific numerical algorithms. The quality of the results has to be accurately assessed, so that in-silico experiments results can be trusted. Nowadays, producing such reliable numerical results goes way beyond the abilities of isolated researchers, and must be carried out by structured teams.

The topics addressed by the RAPSODI project-team belong to the broad theme of numerical methods for the approximation of solutions of systems of partial differential equations (PDEs). Besides standard convergence properties, a good numerical method for approximating a physical problem has to satisfy at least the following three criteria:

1. preservation at the discrete level of some crucial features of the solution, such as positivity of solutions, conservation of prescribed quantities (e.g., mass, the decay of physically motivated entropies, etc.);
2. provide accurate numerical approximations at a reasonable computational cost (and ultimately maximize the accuracy at a fixed computational effort);
3. robustness with respect to physical conditions: the computational cost for a given accuracy should be essentially insensitive to change of physical parameters.

We aim to develop methods fulfilling the above quality criteria for physical models which all display a dissipative behavior, and that are motivated by industrial collaborations or multidisciplinary projects. In particular, we have identified a couple of specific situations we plan to investigate: models from corrosion science (in the framework of nuclear waste repository) [71], low-frequency electromagnetism [88], and mechanics of complex inhomogeneous fluids arising in avalanches [82] or in porous media [73].

Ideally, we should allow ourselves to design entirely new numerical methods. For some applications however (often in the context of industrial collaborations), the members of the team have to compose with existing codes. The numerical algorithms have thus to be optimized under this constraint.

2.2. Scientific Context

Some technological bottlenecks related to points (a)–(c) mentioned above are well identified. In particular, it appears that a good numerical method should handle general meshes, so that dynamic mesh adaptation strategies can be used in order to achieve (b). But it should also be of the highest possible order while remaining stable in the sense of (a), and robust in the sense of (c). There have been numerous research contributions on each point of (a)–(c) in the last decades, in particular for solving each difficulty apart, but combining them still leads to unsolved problems of crucial interest.

Let us mention for example the review paper by J. Droniou [98], where it is highlighted that all the linear methods for solving diffusion equations on general meshes suffer from the same lack of monotonicity and preserve neither the positivity of the solutions nor the decay of the entropy. Moreover, there is no complete convergence proof for the nonlinear methods exposed in [98]. The first convergence proof for a positivity preserving and entropy diminishing method designed to approximate transient dissipative equation on general meshes was proposed very recently in [83]. The idea and the techniques introduced in [83] should be extended to practical applications.

In systems of PDEs, the values of physical parameters often change the qualitative behavior of the solution. Then, one challenge in the numerical approximation of such systems is the design of methods which can be applied for a large range of parameters, as in particular in the regime of singular perturbations. Such schemes, called *asymptotic-preserving* (AP) schemes [107], are powerful tools as they permit the use of the same scheme for a given problem and for its limit with fixed discretization parameters. In many cases, the AP property of numerical schemes is just empirically established, without any rigorous proof. We aim to extend the techniques recently introduced in [79] for the drift-diffusion system, and based on the control of the numerical dissipation of entropy, to other dissipative systems in order prove the AP property of numerical schemes.

The question of the robustness of the numerical methods with respect to the physical parameters is also fundamental for fluid mixtures models. The team already developed such schemes for the variable density Navier–Stokes system [81] or [82]. We aim to propose new ones for more complex models with the same philosophy in mind. On the one hand, we will be interested in high-order schemes, which must be as simple as possible in view of 3D practical implementations. Let us stress that combining high order accuracy and stability is very challenging. On the other hand, the optimization of the computation will have to be considered, in particular with the development of some *a posteriori* error estimators. Impressive progresses have been achieved in this field [94], allowing important computational savings without compromising the accuracy of the results. Recently, we successfully applied this strategy to the Reissner-Mindlin model arising in solid mechanics [90], the dead-oil model for porous media flows [84] or the Maxwell equations in their quasi-static approximation for some eddy current problems [88] and [89]. We aim to develop new *a posteriori* estimators for other dissipative systems, like fluid mixtures models.

In a nutshell, our goal is to take advantage of and extend the most recent breakthroughs of the mathematical community to tackle in an efficient way some application-guided problems coming either from academics or from industrial partners. To this end, we shall focus on the following objectives, which are necessary for the applications we have in mind:

1. *Design and numerical analysis of structure-preserving numerical methods*.
2. *Computational optimization*.

3. Research Program

3.1. Design and analysis of structure-preserving schemes

3.1.1. Numerical analysis of nonlinear numerical methods

Up to now, the numerical methods dedicated to degenerate parabolic problems that the mathematicians are able to analyze almost all rely on the use of mathematical transformations (like e.g. the Kirchhoff's transform). It forbids the extension of the analysis to complex realistic models. The methods used in the industrial codes for solving such complex problems rely on the use of what we call NNM, i.e., on methods that preserve all the nonlinearities of the problem without reducing them thanks to artificial mathematical transforms. Our aim is to take advantage of the recent breakthrough proposed by C. Cancès & C. Guichard [83], [4] to develop efficient new numerical methods with a full numerical analysis (stability, convergence, error estimates, robustness w.r.t. physical parameters,...).

3.1.2. Design and analysis of asymptotic-preserving schemes

There has been an extensive effort in the recent years to develop numerical methods for diffusion equations that are robust with respect to heterogeneities, anisotropy, and the mesh (see for instance [98] for an extensive discussion on such methods). On the other hand, the understanding of the role of nonlinear stability properties in the asymptotic behaviors of dissipative systems increased significantly in the last decades (see for instance [85], [110]).

Recently, C. Chainais-Hillairet and co-authors [79], [86] and [87] developed a strategy based on the control of the numerical counterpart of the physical entropy to develop and analyze AP numerical methods. In particular, these methods show great promises for capturing accurately the behavior of the solutions to dissipative problems when some physical parameter is small with respect to the discretization characteristic parameters, or in the long-time asymptotic. Since it requires the use of nonlinear test functions in the analysis, strong restrictions on the physics (isotropic problems) and on the mesh (Cartesian grids, Voronoï boxes...) are required in [79], [86] and [87]. The schemes proposed in [83] and [4] allow to handle nonlinear test functions in the analysis without restrictions on the mesh and on the anisotropy of the problem. Combining the nonlinear schemes *à la* [83] with the methodology of [79], [86], [87] would provide schemes that are robust both with respect to the meshes and to the parameters. Therefore, they would be also robust under adaptive mesh refinement.

3.1.3. Design and stability analysis of numerical methods for low-Mach models

We aim at extending the range of the `NS2DDV-M` software by introducing new physical models, like for instance the low-Mach model, which gives intermediate solutions between the compressible Navier–Stokes model and the incompressible Navier–Stokes one. This model was introduced in [109] as a limiting system which describes combustion processes at low Mach number in a confined region. Within this scope, we will propose a theoretical study for proving the existence of weak solutions for a particular class of models for which the dynamic viscosity of the fluid is a specific function of the density. We will propose also the extension of a combined Finite Volume-Finite Element method, initially developed for the simulation of incompressible and variable density flows, to this class of models.

3.2. Optimizing the computational efficiency

3.2.1. High-order nonlinear numerical methods

The numerical experiments carried out in [83] show that in case of very strong anisotropy, the convergence of the proposed NNM becomes too slow (less than first order). Indeed, the method appears to strongly overestimate the dissipation. In order to make the method more competitive, it is necessary to estimate the dissipation in a more accurate way. Preliminary numerical results show that second order accuracy in space can be achieved in this way. One also aims to obtain (at least) second order accuracy in time without jeopardizing the stability. For many problems, this can be done by using so-called two-step backward differentiation formulas (BDF2) [99].

Concerning the inhomogeneous fluid models, we aim to investigate new methods for the mass equation resolution. Indeed, we aim at increasing the accuracy while maintaining some positivity-like properties and the efficiency for a wide range of physical parameters. To this end, we will consider Residual Distribution schemes, that appear as an alternative to Finite Volume methods. Residual Distribution schemes enjoy very compact stencils. Therefore, their extension from 2D to 3D yield reasonable difficulties. These methods appeared twenty years ago, but recent extensions to unsteady problems [111], [106], with high-order accuracy [66], [65], or for parabolic problems [63], [64] make them very competitive. Relying on these breakthroughs, we aim at designing new Residual Distribution schemes for fluid mixture models with high-order accuracy while preserving the positivity of the solutions.

3.2.2. A posteriori error control

The question of the *a posteriori* error estimators will also have to be addressed in this optimization context. Since the pioneering papers of Babuska and Rheinboldt more than thirty years ago [70], *a posteriori* error estimators have been widely studied. We will take advantage of the huge corresponding bibliography database in order to optimize our numerical results.

For example, we would like to generalize the results we derived for the harmonic magnetodynamic case (e.g. [88] and [89]) to the temporal magnetodynamic one, for which space/time *a posteriori* error estimators have to be developed. A space/time refinement algorithm should consequently be proposed and tested on academic as well as industrial benchmarks.

We also want to develop *a posteriori* estimators for the variable density Navier–Stokes model or some of its variants. To do so, several difficulties have to be tackled: the problem is nonlinear, unsteady, and the numerical method [81], [82] we developed combines features from Finite Elements and Finite Volumes. Fortunately, we do not start from scratch. Some recent references are devoted to the unsteady Navier–Stokes model in the Finite Element context [77], [114]. In the Finite Volume context, recent references deal with unsteady convection-diffusion equations [113], [68], [96] and [84]. We want to adapt some of these results to the variable density Navier–Stokes system, and to be able to design an efficient space-time remeshing algorithm.

3.2.3. Efficient computation of pairwise interactions in large systems of particles

Many systems are modeled as a large number of punctual individuals (N) which interact pairwise which means $N(N-1)/2$ interactions. Such systems are ubiquitous, they are found in chemistry (Van der Waals interaction between atoms), in astrophysics (gravitational interactions between stars, galaxies or galaxy clusters), in biology (flocking behavior of birds, swarming of fishes) or in the description of crowd motions. Building on the special structure of convolution-type of the interactions, the team develops computation methods based on the Non Uniform Fast Fourier Transform [102]. This reduces the $O(N^2)$ naive computational cost of the interactions to $O(N \log N)$, allowing numerical simulations involving millions of individuals.

4. Application Domains

4.1. Porous media flows

Porous media flows are of great interest in many contexts, like, e.g., oil engineering, water resource management, nuclear waste repository management, or carbon dioxyde sequestration. We refer to [73], [72] for an extensive discussion on porous media flow models.

From a mathematical point of view, the transport of complex fluids in porous media often leads to possibly degenerate parabolic conservation laws. The porous rocks can be highly heterogeneous and anisotropic. Moreover, the grids on which one intends to solve numerically the problems are prescribed by the geological data, and might be non-conformal with cells of various shapes. Therefore, the schemes used for simulating such complex flows must be particularly robust.

4.2. Corrosion and concrete carbonation

The team is interested in the theoretical and numerical analysis of mathematical models describing degradation of materials as concrete carbonation and corrosion. The study of such models is an important environmental and industrial issue. Atmospheric carbonation degrades reinforced concretes and limits the lifetime of civil engineering structures. Corrosion phenomena issues occur for instance in the reliability of nuclear power plants and the nuclear waste repository. The study of the long time evolution of these phenomena is of course fundamental in order to predict the lifetime of the structures.

From a mathematical point of view, the modeling of concrete carbonation (see [67]) as the modeling of corrosion in an underground repository (DPCM model developed by Bataillon *et al.* [71]) lead to systems of PDEs posed on moving domains. The coupling between convection-diffusion-reaction equations and moving boundary equations leads to challenging mathematical questions.

4.3. Complex fluid flows

The team is interested in some numerical methods for the simulation of systems of PDEs describing complex flows, like for instance, mixture flows, granular gases, rarefied gases, or quantum fluids.

Variable-density, low-Mach flows have been widely studied in the recent literature because of their applicability in various phenomena such as flows in high-temperature gas reactors, meteorological flows, flows with convective and/or conductive heat transfer or combustion processes. In such cases, the resolution of the full compressible Navier–Stokes system is not adapted, because of the sound waves speed. The Boussinesq incompressible model is not a better alternative for such low-speed phenomena, because the compressibility effects can not be totally cancelled due to large variations of temperature and density. Consequently, some models have been formally derived, leading to the filtering of the acoustic waves by the use of some formal asymptotic expansions and two families of methods have been developed in the literature in order to compute these flows. We are interested in particular in the so-called pressure-based methods which are more robust than density-based solvers, although their range of validity is in general more limited.

Kinetic theory of molecular gases models a gas as a system of elastically colliding spheres, conserving mechanical energy during impact. Once initialized, it takes a molecular gas not more than few collisions per particle to relax to its equilibrium state, characterized by a Maxwellian velocity distribution and a certain homogeneous density (in the absence of external forces). A granular gas is a system of dissipatively colliding, macroscopic particles (grains). This slight change in the microscopic dynamics (converting energy into heat) causes drastic changes in the behavior of the gas: granular gases are open systems, which exhibit self-organized spatio-temporal cluster formations, and have no equilibrium distribution. They can be used to model silos, avalanches, pollen or planetary rings.

The quantum models can be used to describe superfluids, quantum semiconductors, weakly interacting Bose gases or quantum trajectories of Bohmian mechanics. They have attracted considerable attention in the last decades, due in particular to the development of the nanotechnology applications. To describe quantum phenomena, there exists a large variety of models. In particular there exist three different levels of description: microscopic, mesoscopic and macroscopic. The quantum Navier–Stokes equations deal with a macroscopic description in which the quantum effects are taken into account through a third order term called the quantum Bohm potential. This Bohm potential arises from the fluid dynamical formulation of the single-state Schrödinger equation. The non-locality of quantum mechanics is approximated by the fact that the equations of state do not only depend on the particle density but also on its gradient. These equations were employed to model field emissions from metals and steady-state tunneling in metal- insulator- metal structures and to simulate ultra-small semiconductor devices.

4.4. Stratigraphy

The knowledge of the geology is a prerequisite before simulating flows within the subsoil. Numerical simulations of the geological history thanks to stratigraphy numerical codes allow to complete the knowledge of the geology where experimental data are lacking. Stratigraphic models consist in a description of the erosion and sedimentation phenomena at geological scales.

The characteristic time scales for the sediments are much larger than the characteristic time scales for the water in the river. However, the (time-averaged) water flux plays a crucial role in the evolution of the stratigraphy. Therefore, defining appropriate models that take the coupling between the rivers and the sediments into account is fundamental and challenging. Once the models are at hand, efficient numerical methods must be developed.

4.5. Low-frequency electromagnetism

Numerical simulation is nowadays an essential tool in order to design electromagnetic systems, by estimating the electromagnetic fields generated in a wide variety of devices. An important challenge for many applications is to quantify the intensity of the electric field induced in a conductor by a current generated in its neighborhood. In the low-frequency regime, we can for example quote the study of the impact on the human body of a high-tension line or, for higher frequencies, the one of a smartphone. But the ability to simulate accurately some electromagnetic fields is also very useful for non-destructive control, in the context of the maintenance of nuclear power stations for example. The development of efficient numerical tools, among which *a posteriori* error estimators, is consequently necessary to reach a high precision of calculation in order to provide estimations as reliable as possible.

5. Highlights of the Year

5.1. Highlights of the Year

In 2019, RAPSODI members have been the laureates of several calls for projects.

- T. Rey has been awarded an ERC Generator grant (116 545 euros) from I-SITE Université Lille - Nord Europe for his project MANAKINEQO (R-ERCGEN-19-007-REY). Upon the next two years, T. Rey aims at investigating mathematical properties, as well as developing efficient numerical schemes, for multiscale collisional kinetic equations of the Boltzmann type. A 20-months post-doc will be funded using this grant, as well as an international conference. Following this ERC Generator grant, T. Rey will apply for an ERC Consolidator grant.
- S. Lemaire is the PI of the ADT project ParaSkel++, which is one of the funded Actions of Technological Development of the Inria Lille - Nord Europe 2019 campaign. The aim of the project is to develop an optimized C++ platform for the arbitrary-order numerical approximation of PDEs by skeletal methods on general 2D/3D meshes, with a particular emphasis on the implementation of HPC facilities. L. Beaude has been hired as a development engineer for this project. She will start in February 2020.
- C. Cancès, C. Chainais-Hillairet and B. Merlet are involved in the H2020 project EURAD (EUropean Joint Programme on RADioactive Waste Management). The aim of their project inside EURAD is to establish an energetic formulation of the Diffusion Poisson Coupled Model leading to new large-time robust numerical methods for the simulation of the corrosion processes in an underground repository. C. Cancès is the leader of the task “Numerical methods for high-performance computing of coupled processes” within the EURAD project.

One can also mention the obtention by T. Rey of a Young Researcher PEPS grant from CNRS’s INSMI (3 500 euros, from March to November 2019). The granted project aimed at investigating high-order (in time and velocity) numerical methods for approximating the solutions to the granular gases equation.

6. New Software and Platforms

6.1. Platforms

6.1.1. Platform NS2DDV-M

NS2DDV-M [42] is a Matlab code, developed by C. Calgaro, E. Creusé, and A. Mouton (CNRS research engineer at Université de Lille), for the simulation of homogeneous and inhomogeneous fluid flows by a combined Finite Volume-Finite Element method. The code embeds parallel computation facilities, and is freely distributed, to allow for easy comparisons with concurrent codes on benchmark test-cases, and to promote new collaborations in the domain.

6.1.2. Platform KinDiff

KinDiff consists in a Python library and a Jupyter Notebook. It has been developed by M. Bessemoulin-Chatard (CNRS and Université de Nantes), M. Herda, and T. Rey. The KinDiff code approximates the solutions to the kinetic Fokker–Planck and linear BGK equations in a unidimensional periodic domain in space and a symmetric bounded unidimensional domain in velocity. Its outputs are the discretization of the distribution function in the phase space as well as the moments, along with some visualizations of these quantities. Kinetic equations display asymptotic behaviors such as return to equilibrium in large time, or convergence towards macroscopic equations in the diffusive limit. The finite volume method implemented within KinDiff reproduces these asymptotics at the discrete level. The presentation and numerical analysis of the scheme is available in the article [17]. The code is available at gitlab.com/thoma.rey/FV_HipoDiff.

7. New Results

7.1. Modeling and numerical simulation of complex fluids

In [38], N. Peton, C. Cancès *et al.* propose a new water flow driven forward stratigraphic model. Stratigraphy is a discipline of physics that aims at predicting the geological composition of the subsoil. The model enjoys the following particularities. First, the water surface flow is modelled at the continuous level, in opposition to what is currently done in this community. Second, the model incorporates a constraint on the erosion rate. A stable numerical scheme is proposed to simulate the model.

In [14], A. Ait Hammou Oulhaj and D. Maltese adapt the (positive) nonlinear Control Volume Finite Element scheme of [83] to the simulation of seawater intrusion in the subsoil nearby coastal regions. The proposed scheme is convergent even if the porous medium is anisotropic.

In [25], [41], C. Cancès *et al.* study an original model of degenerate Cahn–Hilliard type. Similarly to the classical degenerate Cahn–Hilliard model, the model can be interpreted as the gradient flow of a Ginzburg–Landau type energy, but the geometry considered here allows for more flexibility and the system thus dissipates faster than the usual degenerate Cahn–Hilliard system. Numerical evidences of this fact are given. Then, the existence of a solution to the model is established thanks to the convergence of a minimizing movement scheme.

In [19], I. Lacroix-Violet *et al.* generalize to the Navier–Stokes–Korteweg (with density-dependent viscosities satisfying the BD relation) and Euler–Korteweg systems a recent relative entropy proposed in [80]. As a concrete application, this helps justifying mathematically the convergence between global weak solutions of the quantum Navier–Stokes system and dissipative solutions of the quantum Euler system when the viscosity coefficient tends to zero. The results are based on the fact that Euler–Korteweg systems and corresponding Navier–Stokes–Korteweg systems can be reformulated through an augmented system. As a by-product of the analysis, I. Lacroix-Violet *et al.* show that this augmented formulation helps to define relative entropy estimates for the Euler–Korteweg systems in a simpler way and with less hypotheses compared to recent works [97], [100].

In [22], C. Calgaro, C. Colin, E. Creusé *et al.* investigate a specific low-Mach model for which the dynamic viscosity of the fluid is a specific function of the density. The model is reformulated in terms of the temperature and velocity, with nonlinear temperature equation, and strong solutions are considered. In addition to a local-in-time existence result for strong solutions, some convergence rates of the error between the approximation and the exact solution are obtained, following the same approach as Guillén-González *et al.* [103], [104].

In [21], C. Calgaro, C. Colin, and E. Creusé derive a combined Finite Volume-Finite Element scheme for a low-Mach model, in which a temperature field obeying an energy law is taken into account. The continuity equation is solved, whereas the state equation linking temperature, density, and thermodynamic pressure is imposed implicitly. Since the velocity field is not divergence-free, the projection method solving the momentum equation has to be adapted. This combined scheme preserves some steady-states, and ensures a discrete maximum principle on the density. Numerical results are provided and compared to other approaches using purely Finite Element schemes, on a benchmark consisting in particular in a transient injection flow [74], [101], [69], as well as in the natural convection of a flow in a cavity [108], [105], [101], [69].

In [20], C. Calgaro, C. Colin, and E. Creusé propose a combined Finite Volume-Finite Element scheme for the solution of a specific low-Mach model expressed in the velocity, pressure and temperature variables. The dynamic viscosity of the fluid is given by an explicit function of the temperature, leading to the presence of a so-called Joule term in the mass conservation equation. First, they prove a discrete maximum principle for the temperature. Second, the numerical fluxes defined for the Finite Volume computation of the temperature are efficiently derived from the discrete Finite Element velocity field obtained by the solution of the momentum equation. Several numerical tests are presented to illustrate the theoretical results and to underline the efficiency of the scheme in terms of convergence rates.

In [46], C. Calgaro and E. Creusé introduce a Finite Volume method to approximate the solution of a convection-diffusion equation involving a Joule term. They propose a way to discretize this so-called "Joule effect" term in a consistent manner with respect to the nonlinear diffusion one, in order to ensure some maximum principle properties on the solution. They investigate the numerical behavior of the scheme on two original benchmarks.

7.2. Numerical simulation in low-frequency electromagnetism

In [32], [31], E. Creusé *et al.* investigate the behavior of some Finite Element error estimators in the context of low-frequency elecromagnetism simulations, to underline the main differences in some practical situations. A more theoretical contribution is also developed to prove the equivalence of some usual discrete gauge conditions. Once again, their numerical behaviors are compared on some characteristic benchmarks.

In [58], F. Chave, S. Lemaire *et al.* introduce a three-dimensional Hybrid High-Order (HHO) method for magnetostatic problems. The proposed method is easy to implement, supports general polyhedral meshes, and allows for arbitrary orders of approximation.

7.3. Structure-preserving numerical methods

In [54], C. Cancès *et al.* propose a Finite Element scheme for the numerical approximation of degenerate parabolic problems in the form of a nonlinear anisotropic Fokker–Planck equation. The scheme is energy-stable, only involves physically motivated quantities in its definition, and is able to handle general unstructured grids. Its convergence is rigorously proven thanks to compactness arguments, under very general assumptions. Although the scheme is based on Lagrange Finite Elements of degree 1, it is locally conservative after a local post-processing giving rise to an equilibrated flux. This also allows to derive a guaranteed *a posteriori* error estimate for the approximate solution. Numerical experiments are presented in order to give evidence of a very good behavior of the proposed scheme in various situations involving strong anisotropy and drift terms.

In [55], C. Chainais-Hillairet and M. Herda apply an iterative energy method à la de Giorgi in order to establish L^∞ bounds for numerical solutions of noncoercive convection-diffusion equations with mixed Dirichlet-Neumann boundary conditions.

In [23], C. Cancès, C. Chainais-Hillairet *et al.* study a finite volume scheme for a degenerate cross-diffusion system describing the ion transport through biological membranes. The strongly coupled equations for the ion concentrations include drift terms involving the electric potential, which is coupled to the concentrations through the Poisson equation. The finite volume scheme is based on two-point flux approximations with "double" upwind mobilities. The existence of solutions to the fully discrete scheme is proven. When the particles are not distinguishable and the dynamics is driven by cross-diffusion only, it is shown that the scheme preserves the structure of the equations like nonnegativity, upper bounds, and entropy dissipation.

In [51], C. Cancès and B. Gaudeul propose a two-point flux approximation finite volume scheme for the approximation of the solutions to an entropy dissipative cross-diffusion system. The scheme is shown to preserve several key properties of the continuous system, among which positivity and decay of the entropy. Numerical experiments illustrate the behavior of the scheme.

In [48], C. Cancès, C. Chainais-Hillairet, B. Gaudeul *et al.* consider an unipolar degenerate drift-diffusion system arising in the modeling of organic semiconductors. They design four different finite volume schemes based on four different formulations of the fluxes. They provide a stability analysis and existence results for the four schemes; the convergence is established for two of them.

In [24], C. Cancès *et al.* compare energy-stable finite volume schemes for multiphase flows in porous media with schemes based on the Wasserstein gradient flow structure of the equations, that has recently been highlighted in [3]. The model is approximated by means of the minimizing movement (or JKO) scheme, that C. Cancès *et al.* solve thanks to the ALG2-JKO scheme proposed in [76].

In [50], C. Cancès *et al.* propose a variational finite volume scheme for the computation of Wasserstein gradient flows. The discrete solution is the minimizer of a discrete action, keeping track at the discrete level of the optimal character of the gradient flow. The spatial discretization relies on upstream mobility fluxes, while an implicit linearization of the Wasserstein distance is used in order to reduce the computational cost by avoiding an inner time-stepping as in the related contributions of the literature.

In [61], T. Rey *et al.* present a new finite volume method for computing numerical approximations of a system of nonlocal transport equations modeling interacting species. In this work, the nonlocal continuity equations are treated as conservative transport equations with a nonlocal, nonlinear, rough velocity field. Some properties of the method are analyzed, and numerical simulations are performed.

In [15], I. Lacroix-Violet *et al.* are interested in the numerical integration in time of nonlinear Schrödinger equations using different methods preserving the energy or a discrete analog of it. In particular, they give a rigorous proof of the order of the relaxation method (presented in [78] for cubic nonlinearities) and they propose a generalized version that allows to deal with general power law nonlinearities. Numerical simulations for different physical models show the efficiency of these methods.

7.4. Cost reduction for numerical methods

In [36], S. Lemaire builds a bridge between the Hybrid High-Order [93] and Virtual Element [75] methods, which are the two main new-generation approaches to the arbitrary-order approximation of PDEs on meshes with general, polytopal cells. The Virtual Element method writes in functional terms and is naturally conforming; at the opposite, the Hybrid High-Order method writes in algebraic terms and is naturally nonconforming. It has been remarked a few years ago that the Hybrid High-Order method can be viewed as a nonconforming version of the Virtual Element method. Here, S. Lemaire ends up unifying the Hybrid High-Order and Virtual Element approaches by showing that the Virtual Element method can be reformulated as a (newborn) conforming Hybrid High-Order method. This parallel has interesting consequences as it sheds new light on the *a priori* analysis of Virtual Element methods, and on the differences between the conforming and nonconforming cases.

In [30], [40], S. Lemaire *et al.* design and analyze (in the periodic setting) arbitrary-order nonconforming multiscale methods for highly oscillatory elliptic problems, which are applicable on coarse grids that may feature general polytopal cells. The construction of these methods is based on the Hybrid High-Order framework [93]. As standard with such multiscale approaches, the general workflow of the method splits into an offline, massively parallelizable stage where all fine-scale computations are performed, and the online, fully coarse-scale stage.

In [52], C. Cancès and D. Maltese propose a reduced model for the migration of hydrocarbons in heterogeneous porous media. Their model keeps track of the time variable. This allows to compute steady-states that cannot be reached by the commonly used ray-tracing and invasion-percolation algorithms. An efficient finite volume scheme allowing for very large time steps is then proposed.

In [57], F. Chave proposes a new definition of the normal fracture diffusion-dispersion coefficient for a reduced model of passive transport in fractured porous media, and numerically studies the impact on the discrete solution on a few test-cases.

In [37], T. Rey *et al.* present high-order, fully explicit time integrators for nonlinear collisional kinetic equations, including the full Boltzmann equation. The methods, called projective integration, first take a few small steps with a simple, explicit method (forward Euler) to damp out the stiff components of the solution. Then, the time derivative is estimated and used in a Runge–Kutta method of arbitrary order. The procedure can be recursively repeated on a hierarchy of projective levels to construct telescopic projective integration methods. The method is illustrated with numerical results in one and two space dimensions.

In [60], I. Lacroix-Violet *et al.* introduce a new class of numerical methods for the time integration of evolution equations set as Cauchy problems of ODEs or PDEs. The systematic design of these methods mixes the Runge–Kutta collocation formalism with collocation techniques, in such a way that the methods are linearly implicit and have high order. The fact that these methods are implicit allows to avoid CFL conditions when the large systems to integrate come from the space discretization of evolution PDEs. Moreover, these methods are expected to be efficient since they only require to solve one linear system of equations at each time step, and efficient techniques from the literature can be used to do so.

7.5. Asymptotic analysis

In [18], C. Cancès *et al.* derive the porous medium equation as the hydrodynamic limit of an interacting particle system which belongs to the family of exclusion processes, with nearest neighbor exchanges. The particles follow a degenerate dynamics, in the sense that the jump rates can vanish for certain configurations, and there exist blocked configurations that cannot evolve. Our approach, which is based on the relative entropy method, is tailored to deal with vanishing initial densities.

In [13], A. Ait Hammou Oulhaj, C. Cancès, C. Chainais-Hillairet *et al.* study the large-time behavior of the solutions to a two-phase extension of the porous media equation, which models the seawater intrusion problem. Their goal is to identify the self-similar solutions that correspond to steady-states of a rescaled version of the problem. They fully characterize the unique steady-states that are identified as minimizers of a convex energy and shown to be radially symmetric. Moreover, they prove the convergence of the solution to the time-dependent model towards the unique stationary state as time goes to infinity. They also provide numerical illustrations of the stationary states and exhibit numerical convergence rates.

In [16], C. Chainais-Hillairet *et al.* propose a new proof of existence of a solution to the scheme introduced in [1] for drift-diffusion systems, which does not require any assumption on the time step. The result relies on the application of a topological degree argument which is based on the positivity and on uniform-in-time upper bounds of the approximate densities. They also establish uniform-in-time lower bounds satisfied by the approximate densities. These uniform-in-time upper and lower bounds ensure the exponential decay of the scheme towards the thermal equilibrium as shown in [1].

In [26], C. Chainais-Hillairet and M. Herda study the large-time behavior of the solutions to Finite Volume discretizations of convection-diffusion equations or systems endowed with non-homogeneous Dirichlet and Neumann type boundary conditions. Their results concern various linear and nonlinear models such as Fokker–Planck equations, porous media equations, or drift-diffusion systems for semiconductors. For all of these models, some relative entropy principle is satisfied and implies exponential decay to the stationary state. They show that in the framework of Finite Volume schemes on orthogonal meshes, a large class of two-point monotone fluxes preserve this exponential decay of the discrete solution to the discrete steady-state of the scheme.

In [49], C. Cancès, C. Chainais-Hillairet, M. Herda *et al.* analyze the large-time behavior of a family of nonlinear finite volume schemes for anisotropic convection-diffusion equations set in a bounded bidimensional domain and endowed with either Dirichlet and/or no-flux boundary conditions. They show that the solutions to the two-point flux approximation (TPFA) and discrete duality finite volume (DDFV) schemes under consideration converge exponentially fast toward their steady-state. The analysis relies on discrete entropy estimates and discrete functional inequalities. As a by-product of their analysis, they establish new discrete Poincaré–Wirtinger, Beckner and logarithmic Sobolev inequalities. Their theoretical results are illustrated by numerical simulations.

In [56], C. Chainais-Hillairet *et al.* introduce a nonlinear DDFV scheme for an anisotropic linear convection-diffusion equation with mixed boundary conditions and establish the exponential decay of the scheme towards its steady-state.

In [39], A. Zurek studies the large-time regime of the moving interface appearing in a concrete carbonation model. He proves that the approximate free boundary, given by an implicit-in-time Finite Volume scheme, propagates in time following a $\sqrt{t}$-law. This result is illustrated by numerical experiments.

In [17], M. Herda, T. Rey *et al.* are interested in the asymptotic analysis of a Finite Volume scheme for one-dimensional linear kinetic equations, with either Fokker–Planck or linearized BGK collision operator. Thanks to appropriate uniform estimates, they establish that the proposed scheme is asymptotic-preserving in the diffusive limit. Moreover, they adapt to the discrete framework the hypocoercivity method proposed in [95] to prove the exponential return to equilibrium of the approximate solution. They obtain decay estimates that are uniform in the diffusive limit. Finally, they present an efficient implementation of the proposed numerical schemes, and perform numerous numerical simulations assessing their accuracy and efficiency in capturing the correct asymptotic behaviors of the models.

In [44], M. Herda *et al.* are interested in the large-time behavior of linear kinetic equations with heavy-tailed local equilibria. Their main contribution concerns the kinetic Lévy–Fokker–Planck equation, for which they adapt hypocoercivity techniques in order to show that solutions converge exponentially fast to the global equilibrium. Compared to the classical kinetic Fokker–Planck equation, the issues here concern the lack of symmetry of the non-local Lévy–Fokker–Planck operator and the understanding of its regularization properties. As a complementary related result, they also treat the case of the heavy-tailed BGK equation.

In [35], M. Herda *et al.* consider various sets of Vlasov–Fokker–Planck equations modeling the dynamics of charged particles in a plasma under the effect of a strong magnetic field. For each of them, in a regime where the strength of the magnetic field is effectively stronger than that of collisions, they first formally derive asymptotically reduced models. In this regime, strong anisotropic phenomena occur; while equilibrium along magnetic field lines is asymptotically reached, the asymptotic models capture a nontrivial dynamics in the perpendicular directions. They do check that in any case the obtained asymptotic model defines a well-posed dynamical system and when self-consistent electric fields are neglected they provide a rigorous mathematical justification of the formally derived systems. In this last step they provide a complete control on solutions by developing anisotropic hypocoercive estimates.

In [45], T. Rey *et al.* propose a new mathematical model intended to describe dynamically the evolution of knowledge in structured societies of interacting individuals. This process, termed cumulative culture, has been extensively studied by evolutionary anthropologists, both theoretically and experimentally. Some of the mathematical properties of the new model are analyzed, and exponential convergence towards a global equilibrium is shown for a simplified model. A numerical method is finally proposed to simulate the complete model.

In [43], following the ideas of V. V. Zhikov and A. L. Pyatnitskii, and more precisely the stochastic two-scale convergence, B. Merlet *et al.* establish a homogenization theorem in a stochastic setting for two nonlinear equations: the equation of harmonic maps into the sphere and the Landau–Lifshitz equation. Homogenization results for nonlinear problems are known to be difficult. In this particular case the equations have strong nonlinear features, in particular, in general their solutions are not unique. Here the authors take advantage of the different equivalent definitions of weak solutions to the nonlinear problem to apply typical linear homogenization recipes.

7.6. Applied calculus of variations

In [34], B. Merlet *et al.* study a variational problem which models the behavior of topological singularities on the surface of a biological membrane in P_β-phase (see [112]). The problem combines features of the Ginzburg–Landau model in 2D and of the Mumford–Shah functional. As in the classical Ginzburg–Landau theory, a prescribed number of point vortices appear in the moderate energy regime; the model allows for discontinuities, and the energy penalizes their length. The novel phenomenon here is that the vortices have a fractional degree $1/m$ with m prescribed. Those vortices must be connected by line discontinuities to form clusters of total integer degrees. The vortices and line discontinuities are therefore coupled through a topological constraint. As in the Ginzburg–Landau model, the energy is parameterized by a small length scale $\varepsilon > 0$. B. Merlet *et al.* perform a complete Γ-convergence analysis of the model as $\varepsilon \downarrow 0$ in the moderate energy regime. Then, they study the structure of minimizers of the limit problem. In particular, the line discontinuities of a minimizer solve a variant of the Steiner problem.

In [27], B. Merlet *et al.* consider the branched transportation problem in 2D associated with a cost per unit length of the form $1 + \beta\theta$ where θ denotes the amount of transported mass and $\beta > 0$ is a fixed parameter (notice that the limit case $\beta = 0$ corresponds to the classical Steiner problem). Motivated by the numerical approximation of this problem, they introduce a family of functionals $(\{\mathcal{F}_\varepsilon\}_{\varepsilon>0})$ which approximate the above branched transport energy. They justify rigorously the approximation by establishing the equicoercivity and the Γ-convergence of $\{\mathcal{F}_\varepsilon\}$ as $\varepsilon \downarrow 0$. The functionals are modeled on the Ambrosio–Tortorelli functional and are easy to optimize in practice. Numerical evidences of the efficiency of the method are presented.

In [28], B. Merlet *et al.* establish new results on the approximation of k-dimensional surfaces (k-rectifiable currents) by polyhedral surfaces with convergence in h-mass and with preservation of the boundary (the approximating polyhedral surface has the same boundary as the limit). This approximation result is required in the convergence study of [29].

In [29], B. Merlet *et al.* consider a generalization of branched transportation in arbitrary dimension and codimension: minimize the h-mass of some oriented k-dimensional branched surface in $\mathbf{R}^n$ with some prescribed boundary. Attached to the surface is a multiplicity $m(x)$ which is not necessarily an integer and is a conserved quantity (Kirchhoff current law is satisfied at branched points). The h-mass is defined as the integral of a cost $h(|m(x)|)$ over the branched surface. As usual in branched transportation, the cost function is a lower-semicontious, sublinear increasing function with $h(0) = 0$ (for instance $h(m) = \sqrt{1 + am^2}$ if $m \neq 0$ and $h(0) = 0$). For numerical purpose, it is convenient to approximate the measure defined by the k-dimensional surfaces by smooth functions in $\mathbf{R}^n$. In this spirit, B. Merlet *et al.* propose phase field approximations of the branched surfaces and of their energy in the spirit of the Ambrosio–Tortorelli functional. The convergence of these approximations towards the original k-dimensional branched transportation problem is established in the sense of Γ-convergence. Next, considering the cost $h(m) = \sqrt{1 + am^2}$ and sending a to 0, a phase field approximation of the Plateau problem is obtained. Numerical experiments show the efficiency of the method. These numerical results are exceptional as they are obtained without any guess on the topology of the minimizing k-surface (as opposed to methods based on parameterizations of the k-surface).

In [33], [62], B. Merlet *et al.* study a family of functionals penalizing oblique oscillations. These functionals naturally appear in some variational problems related to pattern formation and are somewhat reminiscent of those introduced by Bourgain, Brezis and Mironescu to characterize Sobolev functions. More precisely, for a function u defined on a tensor product $\Omega_1 \times \Omega_2$, the family of functionals $\{E_\varepsilon(u)\}_{\varepsilon>0}$ that we consider vanishes if u is of the form $u(x_1)$ or $u(x_2)$. We prove the converse property and related quantitative results. In particular, we describe the fine properties of functions with $\sup_\varepsilon E_\varepsilon(u) < \infty$ by showing that roughly, such u is piecewise of the form $u(x_1)$ or $u(x_2)$ on domains separated by lines where the energy concentrates. It turns out that this problem naturally leads to the study of various differential inclusions and has connections with branched transportation models.

7.7. Approximation theory

In [59], M. Herda *et al.* propose an iterative algorithm for the numerical computation of sums of squares of polynomials approximating given data at prescribed interpolation points. The method is based on the definition of a convex functional G arising from the dualization of a quadratic regression over the Cholesky factors of the sum of squares decomposition. In order to justify the construction, the domain of G, the boundary of the domain and the behavior at infinity are analyzed in details. When the data interpolate a positive univariate polynomial, we show that in the context of the Lukacs sum of squares representation, G is coercive and strictly convex which yields a unique critical point and a corresponding decomposition in sum of squares. For multivariate polynomials which admit a decomposition in sum of squares and up to a small perturbation of size ε, G^ε is always coercive and so its minimum yields an approximate decomposition in sum of squares. Various unconstrained descent algorithms are proposed to minimize G. Numerical examples are provided, for univariate and bivariate polynomials.

In [47], M. Herda *et al.* investigate the numerical approximation of bounded functions by polynomials satisfying the same bounds. The contribution makes use of the recent algebraic characterization found in [91] and [92] where an interpretation of monovariate polynomials with two bounds is provided in terms of a quaternion algebra and the Euler four-squares formulas. Thanks to this structure, the authors generate a new nonlinear projection algorithm onto the set of polynomials with two bounds. The numerical analysis of the method provides theoretical error estimates showing stability and continuity of the projection. Some numerical tests illustrate this novel algorithm for constrained polynomial approximation.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

A new contractual collaboration between C. Cancès and IFPEN corresponding to the supervision of the PhD thesis of S. Bassetto started in January 2019. This contract is part of the Inria-IFPEN framework agreement.

8.2. Bilateral Grants with Industry

C. Bataillon (CEA) and L. Trenty (ANDRA) are involved in the `EURAD` project on corrosion modeling together with C. Cancès, C. Chainais-Hillairet, and B. Merlet. More details in Section 9.3.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. ERC Generator

T. Rey has been awarded an ERC Generator grant (116 545 euros) from I-SITE Université Lille - Nord Europe for his project `MANAKINEQO` (R-ERCGEN-19-007-REY). In the next two years, T. Rey aims at investigating mathematical properties, as well as developing efficient numerical schemes, for multiscale collisional kinetic equations of the Boltzmann type. A 20-months post-doc will be funded using this grant, as well as an international conference. Following this ERC Generator grant, T. Rey will apply for an ERC Consolidator grant.

9.1.2. Actions of Technological Development (ADT)

S. Lemaire is the PI of the ADT project `ParaSkel++`, which is one of the funded ADT of the Inria Lille - Nord Europe 2019 campaign. The aim of the project is to develop an optimized C++ platform for the arbitrary-order numerical approximation of PDEs by skeletal methods on general 2D/3D meshes, with a particular emphasis on the implementation of HPC facilities. L. Beaude has been hired as a development engineer for this project. She will start in February 2020.

In the same vein, T. Rey is part of the ADT project `SIMPAPH` led by the MEPHYSTO-POST team, that has as well been funded as a result of the Inria Lille - Nord Europe 2019 campaign. The aim is to develop robust numerical methods to solve large systems of stochastic differential equations describing (among others) particles in an optic fiber, schools of fish, or microscopic particles. The expected code will attempt to solve these multiscale problems using different approaches, and to be versatile enough to act as an industrial benchmark. A. Roget has been hired as a development engineer for this project.

9.2. National Initiatives

9.2.1. ANR

C. Chainais-Hillairet has been a member of the ANR `MOONRISE` project. The `MOONRISE` project aimed at exploring modeling, mathematical, and numerical issues originating from the presence of high oscillations in nonlinear PDEs mainly from the physics of nanotechnologies and from the physics of plasmas.

Title: `MOdels`, `Oscillations`, and `NumeRIcal SchEmes`

Type: Fondements du numérique (DS0705) - 2014

ANR reference: ANR-14-CE23-0007

Coordinator: F. Méhats (Université de Rennes 1)

Duration: October 2014 - June 2019

C. Chainais-Hillairet and T. Rey are members of the ANR MOHYCON project. The MOHYCON project is related to the analysis and simulation of multiscale models of semiconductors. As almost all current electronic technology involves the use of semiconductors, there is a strong interest for modeling and simulating the behavior of such devices, which was recently reinforced by the development of organic semiconductors used for example in solar panels or in mobile phones and television screens (among others).

Title: Multiscale MOdels and HYbrid numerical methods for semiCONductors

Type: Société de l'information et de la communication (DS07) - 2017

ANR reference: ANR-17-CE40-0027

Coordinator: M. Bessemoulin-Chatard (CNRS and Université de Nantes)

Duration: January 2018 - December 2020

C. Cancès is a member of the ANR COMODO project. The COMODO project focuses on the mathematical and numerical study of cross-diffusion systems in moving domains. The targeted application is the simulation of the building of solar plants by the vapour deposition process.

Title: CrOss-diffusion equations in MOving DOmains

Type: Modèles numériques, simulation, applications (CE46) - 2019

ANR reference: ANR-19-CE46-0002

Coordinator: V. Ehrlacher (École des Ponts ParisTech and Inria Paris)

Duration: January 2020 - December 2023

M. Herda is a member of the ANR JCJC MICMOV project. The MICMOV project aims at gathering PDE analysts, probability theorists, and theoretical physicists to work on the derivation of macroscopic properties of physical systems from their microscopic description. The rigorous microscopic description of moving interfaces, the understanding of macroscopic nonlocal effects, and the mathematical apprehension of the underlying atomic mechanisms, are particularly important matters of this project.

Title: MICroscopic description of MOVing interfaces

Type: Mathématiques (CE40) - 2019

Coordinator: M. Simon (Inria Lille - Nord Europe)

9.2.2. LabEx CEMPI

Title: Centre Européen pour les Mathématiques, la Physique et leurs Interactions

Coordinator: S. De Bièvre (LPP, Université de Lille)

Duration: January 2012 - December 2019, extended in 2019

Partners: Laboratoire Paul Painlevé (LPP) and Laser Physics department (PhLAM), Université de Lille

The "Laboratoire d'Excellence" Centre Européen pour les Mathématiques, la Physique et leurs Interactions (CEMPI), a project of the Laboratoire de mathématiques Paul Painlevé (LPP) and the laboratoire de Physique des Lasers, Atomes et Molécules (PhLAM), was created in the context of the "Programme d'Investissements d'Avenir" in February 2012.

The association Painlevé-PhLAM creates in Lille a research unit for fundamental and applied research and for training and technological development that covers a wide spectrum of knowledge stretching from pure and applied mathematics to experimental and applied physics.

One of the three focus areas of CEMPI research is the interface between mathematics and physics. This focus area encompasses three themes. The first is concerned with key problems of a mathematical, physical and technological nature coming from the study of complex behavior in cold atoms physics and nonlinear optics, in particular fiber optics. The two other themes deal with fields of mathematics such as algebraic geometry, modular forms, operator algebras, harmonic analysis and quantum groups that have promising interactions with several branches of theoretical physics.

9.2.3. PEPS

T. Rey has been the laureate in 2019 of a Young Researcher PEPS grant from CNRS's INSMI (3 500 euros, from March to November 2019). The granted project aimed at investigating high-order (in time and velocity) numerical methods for approximating the solutions to the granular gases equation.

9.3. European Initiatives

C. Cancès, C. Chainais-Hillairet and B. Merlet are involved in the H2020 project EURAD (EUropean Joint Programme on RADioactive Waste Management). The aim of their project inside EURAD is to establish an energetic formulation of the Diffusion Poisson Coupled Model leading to new large-time robust numerical methods for the simulation of the corrosion processes in an underground repository.

C. Cancès is the leader of the task "Numerical methods for high-performance computing of coupled processes" within the EURAD project.

9.4. International Initiatives

C. Cancès is a member of the Indo-French Center for Applied Mathematics (IFCAM) project "Conservation laws: BV^s, control, interfaces" (PIs: S. Ghoshal, TIFR Centre For Applicable Mathematics, India and S. Junca, Université de Nice).

9.5. International Research Visitors

9.5.1. Visits of International Scientists

In January-February, E. Daus (TU Vienna, Austria) visited C. Cancès and C. Chainais-Hillairet during two weeks.

In 2019, RAPSODI members also invited several researchers for short visits (a week or less) in Lille.

- R. Bailo (Imperial College London, UK) came in November to work with T. Rey (funded by his Young Researcher PEPS grant).
- R. V. Sabariego (KU Leuven, Electrical Engineering ESAT/Electa, EnergyVille, Belgium) came in May to work with E. Creusé.
- E. Bretin and S. Masnou (Institut Camille Jordan, Lyon) and M. Goldman (CNRS and LJLL/Université Paris Diderot) came in June and December to work with B. Merlet.
- A. Trescases (CNRS and Institut de Mathématiques de Toulouse) came in September to work with M. Herda.
- C. Bataillon (CEA), V. Ehrlacher (École des Ponts ParisTech and Inria Paris) and C. Perrin (CNRS and Université d'Aix-Marseille) came to work with C. Cancès.
- M. Cassier (CNRS and Institut Fresnel, Marseille) came in February-March to work with S. Lemaire.

On a slightly different note, from March to October, G. Robillard has been an AIRLab resident (Artiste en Immersion Recherche dans un Laboratoire) in order to work with C. Calgaro and E. Creusé (with a support from the Communauté d'Universités et d'Établissements Lille Nord-de-France).

9.5.2. Visits to International Teams

B. Gaudeul spent two weeks in WIAS Berlin, Germany in November in order to work with J. Fuhrmann on the extension of the results obtained in [48] to Nernst–Planck–Poisson systems with ion size and solvation effects.

M. Herda spent one week at Imperial College London, UK in February to work with P. Degond on a Fokker–Planck approach to the study of robustness in gene expression.

T. Rey visited 3 times J. A. Carrillo and J. Hu at Imperial College London, UK between February and March, for 3 days long stays funded by his Young Researcher PEPS grant, to work (in particular) on the development of a new high-order numerical method for solving the granular gases equation.

C. Cancès and B. Merlet spent one week at the University of Lisbon, Portugal in December to work with L. Monsaingeon.

C. Cancès spent one week at Université de Tours to work with B. Andreianov.

B. Merlet visited E. Bretin and S. Masnou at Institut Camille Jordan in Lyon in February, and visited several times M. Goldman at LJLL/Université Paris Diderot in March, May, October, and November.

F. Chave and S. Lemaire spent 3 days at Université de Montpellier in September to work with D. A. Di Pietro on arbitrary-order polytopal methods for electromagnetism.

9.5.3. Research Stays Abroad

M. Herda was in residence at the Hausdorff Research Institute for Mathematics (University of Bonn, Germany) from May 19 to July 7 in the framework of the Junior Trimester Program in Kinetic Theory, that gave young mathematicians the opportunity to carry out collaborative research in kinetic theory. M. Herda was part of a project in collaboration with N. Ayi (Sorbonne Université), M. Breden (École Polytechnique), J. Guerand (University of Cambridge, UK), H. Hivert (Centrale Lyon), and I. Tristani (ENS Paris) on the study of a fractional kinetic Fokker–Planck equation. This collaboration has already led to the article [44], and a second article is in preparation. A collaboration was also initiated with M. Breden and A. Trescases (CNRS and Institut de Mathématiques de Toulouse) on the derivation of cross-diffusion systems from kinetic models.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

M. Herda co-organized the yearly Journée de la Fédération de Recherche Mathématique du Nord-Pas-de-Calais which took place on October 15 in Lille. This one-day conference brought together all the maths research laboratories from the Hauts-de-France region and gave the opportunity to new permanent members to present their research.

E. Creusé was part of the organizing committee of the first Rencontres Mathématiques Valenciennoises, that were held on March 13. This meeting was organized for maths students at Université Polytechnique Hauts-de-France as well as for high-school students of Lycée Wallon in Valenciennes. Several researchers presented some applications of mathematics in industry as well as in academic research. This meeting gathered around 500 participants.

T. Rey organized the yearly ANEDP team day at Laboratoire Paul Painlevé on December 17.

C. Cancès and T. Rey co-organized mini-symposia on, respectively, (i) cross-diffusion systems and (ii) novel time discretization methods and moment models for kinetic equations at ICIAM 2019 in Valencia. S. Lemaire co-organized a mini-symposium on polytopal methods at MAFELAP 2019 in Brunel University, London.

10.1.2. Journal

10.1.2.1. Member of the Editorial Boards

C. Chainais-Hillairet is a member of the editorial board of the North-Western European Journal of Mathematics.

10.1.2.2. Reviewer - Reviewing Activities

RAPSODI team members are regular reviewers for all the main international journals in numerical analysis/scientific computing and PDEs.

10.1.3. Invited Talks

C. Calgaro was one of the speakers of the seventh Amarena Days held in Amiens.

C. Cancès was invited to give a talk in Paris during the closure workshop of the MANON common laboratory (CEA Saclay / LJLL Sorbonne Université). He was also invited to give a talk in Kalvåg (Norway) during the workshop Analysis and Computation of Coupled Systems. He finally gave seminars in Compiègne and at the University of Lisbon.

C. Chainais-Hillairet was an invited plenary speaker at the POEMs conference held at CIRM in Marseille, and for the Journées Corrosion MaNu 2019 held in Orsay. She also gave a talk at ENUMATH 2019 that was held in Egmond aan Zee. She finally gave seminars in Nice, Orsay and Paris (Laboratoire Jacques-Louis Lions), as well as a talk for the ANEDP team day.

F. Chave was an invited speaker in ENUMATH 2019 that was held in Egmond aan Zee. He also presented a poster at the POEMs conference held at CIRM in Marseille. He finally gave a seminar for the ANEDP team at Laboratoire Paul Painlevé.

B. Gaudeul was an invited speaker of the Journées Corrosion MaNu 2019 held in Orsay. He also presented a poster at the conference Numerical Methods for Multiscale Models arising in Physics and Biology held in Nantes. He finally gave a seminar at WIAS Berlin, as well as a talk for the PhD students seminar in Lille.

M. Herda was an invited speaker in the workshop Analytical and Computational Problems for Mixtures and Plasma Dynamics which took place at the Hausdorff Research Institute for Mathematics in Bonn in the framework of the Junior Trimester Program in Kinetic Theory. He was also one of the speakers of the mini-symposium “Novel time discretization methods and moment models for kinetic equations” co-organized by T. Rey at ICIAM 2019 in Valencia. Finally, he gave a talk in the Applied PDEs seminar of Imperial College London.

I. Lacroix-Violet was an invited speaker in the workshop Women in PDEs that was held in Vienna. She also gave several seminars in Université Paris-Est Créteil, Nancy, Dijon, and Lyon, as well as a talk for the ANEDP team day.

S. Lemaire was an invited speaker at the GAMM 2019 conference held in Vienna, and for the POEMs conference held at CIRM in Marseille. He also gave a talk for the yearly Journée de la Fédération de Recherche Mathématique du Nord-Pas-de-Calais held in Lille, and a seminar at ENSTA ParisTech.

B. Merlet was an invited speaker of the fourth Journées Optimisation de Formes et Applications held in Palaiseau. He also gave two seminars in École Polytechnique and at the University of Lisbon, as well as a talk for the ANEDP team day.

T. Rey was an invited plenary speaker in the Summer School Trails in kinetic theory: foundational aspects and numerical methods from the Junior Trimester Program in Kinetic Theory of the Hausdorff Research Institute for Mathematics, held in Bonn. He was also one of the speakers of the conference Numerical Methods for Multiscale Models arising in Physics and Biology held in Nantes. He finally gave a colloquium in the mathematics department of Université Paris-Est Créteil, and a seminar at ENS Ker Lann in Rennes.

A. Zurek was an invited speaker of the Journée EDP held in Calais. He was also a contributed speaker in the conference Numerical Methods for Multiscale Models arising in Physics and Biology held in Nantes. He finally gave a talk during the thirteenth Journée des Doctorants en Mathématiques du Nord-Pas-de-Calais in Lens.

10.1.4. Leadership within the Scientific Community

C. Cancès is the head of the French Research Group GdR MaNu on Mathematics for Nuclear Energy (funded by CNRS’s INSMI).

He is also the leader of the task “Numerical methods for high-performance computing of coupled processes” within the EURAD project (H2020, EJP COFUND); cf. Section 9.3.

10.1.5. Research Administration

C. Calgaro is a member of the Conseil de Département de Mathématiques. She has been a member of the CNU 26 until September.

C. Cancès is a member of the Scientific Board of the Inria Lille - Nord Europe research center.

Until September, C. Chainais-Hillairet has been a member of the Conseil de Département de Mathématiques. Since September, she is a member of the Conseil de la Faculté des Sciences et Technologies of Université de Lille. She also joined in January the CNRS evaluation committee of the Laboratoire de Mathématiques de Reims.

B. Gaudeul is the delegate of the PhD students at the Commission Mixte in Mathematics of Université de Lille.

I. Lacroix-Violet is an elected member of the Conseil du Laboratoire Paul Painlevé. She is also a member of the Comité d'Éthique de la Recherche.

S. Lemaire is a member of the Commission de Développement Technologique of the Inria Lille - Nord Europe research center.

B. Merlet is an elected member of the Conseil du Laboratoire Paul Painlevé.

T. Rey is also an elected member of the Conseil du Laboratoire Paul Painlevé. T. Rey and M. Herda (since September) are in charge of the organization of the weekly seminar of the ANEDP team at the Laboratoire Paul Painlevé. T. Rey is a member of the team of Opération Postes.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

RAPSODI team members are strongly involved in teaching at Université de Lille. C. Calgaro is in charge of the Master of Mathematical Engineering. I. Lacroix-Violet is responsible of the IESP2A's third year at Polytech Lille. B. Merlet is in charge of the Master 2 of Scientific Computing. E. Creusé is the director of the Mathematics Department of the Université Polytechnique Hauts-de-France.

In 2019, C. Cancès gave practicals (16h) of scientific computing at the École Centrale de Lille. He also gave lectures (32h) on fundamental notions in mathematics in the framework of the newly created Master on Data Sciences of the Université de Lille and the École Centrale de Lille. F. Chave gave practicals to first-year students at Polytech Lille. M. Herda gave lectures (26h) on mathematical methods for signal processing in the Master of Numerical and Statistical Engineering of Université de Lille. S. Lemaire gave lectures (32h) on mathematical tools for simulation in the Master 2 of Scientific Computing at Université de Lille.

C. Chainais-Hillairet gave lectures about "Finite volume methods, dissipative problems, and long-time behavior" in 2 Summer Schools, at Le Lioran (GdR `EGRIN`) in June, and at CIRM in Marseille (CEMRACS 2019) in July.

10.2.2. Supervision

Post-doc in progress: F. Chave, on "High-order polytopal discretization methods for electromagnetism", since 12/01/2018; supervisors: S. Lemaire and E. Creusé.

PhD in progress: A. Nahas, on "Vortices in Bose–Einstein condensates", since 09/01/2019; advisors: I. Lacroix-Violet and G. Dujardin (Inria Lille - Nord Europe). The thesis is half funded by the LabEx `CEMPI`.

PhD in progress: S. Bassetto, on "Towards a more robust and accurate treatment of capillary effects in multiphase flow simulations in porous media", since 01/01/2019; advisors: C. Cancès, G. Enchéry and Q.-H. Tran (IFPEN).

PhD in progress: B. Gaudeul, on the "Numerical approximation of cross-diffusion systems arising in physics and biology", since 09/01/2018; advisors: C. Chainais-Hillairet and C. Cancès.

PhD: A. Zurek defended his PhD thesis [11] entitled “Problèmes à interface mobile pour la dégradation de matériaux et la croissance de biofilms : analyse numérique et modélisation” on September 26; advisors: C. Chainais-Hillairet and B. Merlet.

PhD: C. Colin defended her PhD thesis [12] entitled “Analyse et simulation numérique par méthode combinée Volumes Finis - Éléments Finis de modèles de type faible Mach” on May 10; advisors: E. Creusé and C. Calgaro.

M2 Modélisation - Optimisation - Sécurité group research in progress (Université Polytechnique Hauts-de-France): L. Romanowicz and A. Violet, on the “Numerical study of the rise of a suspension bridge”, 2019-2020; advisor: E. Creusé.

M2 internship: L. Maisonneuve, on the “Mathematical modeling of a knowledge model in structured populations”, from April to September; advisors: T. Rey, S. Billiard and M. Derex.

M2 internship: A. Nahas, on the “Numerical study of the local minimizers of a Bose–Einstein energy functional in 2D”, from January to July; advisors: I. Lacroix-Violet and G. Dujardin (Inria Lille - Nord Europe).

M2 internship: N. Aghouzzaf, on “Keller–Segel equations for chemotaxis”, from December 2018 to July; advisor: T. Rey.

M1 internship: Q. Fourche, on “Spectral moment methods for linear kinetic equations”, from June to July; advisor: T. Rey.

M1 internship: M. Jonval, on “Nonlinear numerical schemes for drift-diffusion equations”, from June to July; advisor: C. Chainais-Hillairet.

M1 CHPS group research (Université de Lille): M. Ghestin and A. Rotolo, on “Systems of interacting particles and their mean-field limits”, from February to May; advisor: T. Rey.

M1 MAS group research (Université de Lille): C. Germain and A. Jossien, on “Deterministic predictive models of the average annual temperature in Morocco in 2019”, 2019; advisor: C. Calgaro.

M1 Modélisation - Optimisation - Sécurité group research (Université Polytechnique Hauts-de-France): I. Kahiyehmoumin, A. Ndao and M. Fall, on a “Domain decomposition method for the finite difference approximation of a diffusion problem”, 2019; advisor: E. Creusé.

L3 internship: G. Helbecque, on the “Numerical simulation of charged particles in a strong magnetic field”, from April to May; advisor: M. Herda.

10.2.3. Juries

C. Calgaro was a jury member for the PhD defense of M. Tsegmid (Université du Littoral-Côte-d’Opale). She is also a member of the Jury de l’Agrégation de Mathématiques, which is a national hiring committee for the highest level of high-school teachers.

C. Cancès reported on T. Beltzung’s (CEA / UVSQ), J. W. Both’s (University of Bergen, Norway), and D. Shylaja’s (Monash University / ITT Bombay, Australia / India) PhD theses. He is also a member of the Jury de l’Agrégation de Mathématiques.

C. Chainais-Hillairet reported on A. Gerstenmayer’s (T.U. Vienna) PhD thesis. She was also a jury member for the PhD defenses of L. Boittin (Paris), C. Colin (Lille), R. Chalayer (Clermont-Ferrand), P. Mennuni (Lille), and M. Id Moulay (Pau). She finally was part of the selection committee for an associate professor (MCF) position at Orsay.

E. Creusé reported on P. Daniel’s PhD thesis, defended on March 22 at Sorbonne Université, and entitled “Éléments finis hp-adaptatifs avec contraction d’erreur garantie et solveurs multi-niveaux inexacts”. He was also a jury member for the PhD defenses of J. Tomezyk (Université Polytechnique Hauts-de-France) and G. Jeanmasson (Université de Bordeaux). He finally was part of the selection committee for an associate professor (MCF) position at Université Polytechnique Hauts-de-France.

B. Gaudeul was a jury member for the Tournoi Français des Jeunes Mathématiciennes et Mathématiciens.

I. Lacroix-Violet was a jury member for the PhD defense of I. Bensouilah (Lille), and for the habilitation of D. Sanchez (Toulouse).

B. Merlet reported on M. Bonafini's PhD thesis, defended on April 11 at the University of Trento (Italy), and entitled "Variational and convex approximations of 1-dimensional optimal networks and hyperbolic obstacle problems".

10.3. Popularization

C. Calgaro is in charge of the communication of the Laboratoire Paul Painlevé. She organized in 2019 various events to promote mathematics among young people:

- "Les mathématiciens de Lille fêtent les 80 ans du CNRS" (about 600 participants from high-school or Classes Préparatoires);
- a conference of E. Ghys for the presence of the Jury de l'Agrégation de Mathématiques in Lille (about 200 participants);
- les "Mathématiques itinérantes", a series of conferences in high-schools.

E. Creusé gave two popularization talks at Université Polytechnique Hauts-de-France:

- "Le calcul scientifique pour la simulation et le contrôle des écoulements" for the first Rencontres Mathématiques Valenciennoises in March;
- "Comment prévoir la population de poissons dans l'océan ou le déplacement d'oiseaux migrateurs dans le ciel ? Les maths nous répondent !" for the Semaine des Mathématiques 2019 in March.

M. Herda gave a talk at Inria Lille - Nord Europe for the internal scientific popularization event "30 minutes of Science" in January.

T. Rey was one of the four speakers of the event "Les mathématiciens de Lille fêtent les 80 ans du CNRS".

11. Bibliography

Major publications by the team in recent years

[1] M. Bessemoulin-Chatard, C. Chainais-Hillairet.*Exponential decay of a finite volume scheme to the thermal equilibrium for drift–diffusion systems*, in "Journal of Numerical Mathematics", 2017, vol. 25, n^{o} 3, p. 147-168 [*DOI :* 10.1515/JNMA-2016-0007], https://hal.archives-ouvertes.fr/hal-01250709

[2] C. Calgaro, E. Creusé, T. Goudon, S. Krell.*Simulations of non homogeneous viscous flows with incompressibility constraints*, in "Mathematics and Computers in Simulation", 2017, vol. 137, p. 201-225, https://hal.archives-ouvertes.fr/hal-01246070

[3] C. Cancès, T. Gallouët, L. Monsaingeon.*Incompressible immiscible multiphase flows in porous media: a variational approach*, in "Analysis & PDE", 2017, vol. 10, n^{o} 8, p. 1845–1876 [*DOI :* 10.2140/APDE.2017.10.1845], https://hal.archives-ouvertes.fr/hal-01345438

[4] C. Cancès, C. Guichard.*Numerical analysis of a robust free energy diminishing Finite Volume scheme for parabolic equations with gradient structure*, in "Foundations of Computational Mathematics", 2017, vol. 17, n^{o} 6, p. 1525-1584, https://hal.archives-ouvertes.fr/hal-01119735

[5] C. CHAINAIS-HILLAIRET, B. MERLET, A. VASSEUR.*Positive Lower Bound for the Numerical Solution of a Convection-Diffusion Equation*, in "FVCA8 2017 - International Conference on Finite Volumes for Complex Applications VIII", Lille, France, Springer, June 2017, p. 331-339 [*DOI :* 10.1007/978-3-319-57397-7_26], https://hal.archives-ouvertes.fr/hal-01596076

[6] D. A. DI PIETRO, A. ERN, S. LEMAIRE.*An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators*, in "Computational Methods in Applied Mathematics", June 2014, vol. 14, n^o 4, p. 461-472 [*DOI :* 10.1515/CMAM-2014-0018], https://hal.archives-ouvertes.fr/hal-00978198

[7] G. DIMARCO, R. LOUBÈRE, J. NARSKI, T. REY.*An efficient numerical method for solving the Boltzmann equation in multidimensions*, in "Journal of Computational Physics", 2018, vol. 353, p. 46-81 [*DOI :* 10.1016/J.JCP.2017.10.010], https://hal.archives-ouvertes.fr/hal-01357112

[8] F. FILBET, M. HERDA.*A finite volume scheme for boundary-driven convection-diffusion equations with relative entropy structure*, in "Numerische Mathematik", 2017, vol. 137, n^o 3, p. 535-577, https://hal.archives-ouvertes.fr/hal-01326029

[9] I. LACROIX-VIOLET, A. VASSEUR.*Global weak solutions to the compressible quantum Navier–Stokes equation and its semi-classical limit*, in "Journal de Mathématiques Pures et Appliquées", 2018, vol. 114, p. 191-210, https://hal.archives-ouvertes.fr/hal-01347943

[10] B. MERLET.*A highly anisotropic nonlinear elasticity model for vesicles I. Eulerian formulation, rigidity estimates and vanishing energy limit*, in "Arch. Ration. Mech. Anal.", 2015, vol. 217, n^o 2, p. 651–680 [*DOI :* 10.1007/S00205-014-0839-5], https://hal.archives-ouvertes.fr/hal-00848547

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] A. ZUREK.*Free boundary problems for the degradation of materials and biofilms growth: numerical analysis and modelisation*, Université de Lille, September 2019, https://tel.archives-ouvertes.fr/tel-02397231

[12] C. COLIN.*Analysis and numerical simulation by a combined Finite Volumes - Finite Elements method of low Mach type models*, Université de Lille / Laboratoire Paul Painlevé, May 2019, https://hal.archives-ouvertes.fr/tel-02406716

Articles in International Peer-Reviewed Journal

[13] A. AIT HAMMOU OULHAJ, C. CANCÈS, C. CHAINAIS-HILLAIRET, P. LAURENÇOT.*Large time behavior of a two phase extension of the porous medium equation*, in "Interfaces and Free Boundaries", 2019, vol. 21, p. 199-229, https://arxiv.org/abs/1803.10476 [*DOI :* 10.4171/IFB/421], https://hal.archives-ouvertes.fr/hal-01752759

[14] A. AIT HAMMOU OULHAJ, D. MALTESE.*Convergence of a positive nonlinear control volume finite element scheme for an anisotropic seawater intrusion model with sharp interfaces*, in "Numerical Methods for Partial Differential Equations", 2020, vol. 36, n^o 1, p. 133-153 [*DOI :* 10.1002/NUM.22422], https://hal.archives-ouvertes.fr/hal-01906872

[15] C. BESSE, S. DESCOMBES, G. DUJARDIN, I. LACROIX-VIOLET.*Energy preserving methods for nonlinear Schrödinger equations*, in "IMA Journal of Numerical Analysis", 2019, forthcoming [*DOI :* 10.1016/J.APNUM.2019.11.008], https://hal.archives-ouvertes.fr/hal-01951527

[16] M. BESSEMOULIN-CHATARD, C. CHAINAIS-HILLAIRET.*Uniform-in-time Bounds for approximate Solutions of the drift-diffusion System*, in "Numerische Mathematik", 2019, vol. 141, n$^{\mathrm{o}}$ 4, p. 881-916 [*DOI :* 10.1007/S00211-018-01019-1], https://hal.archives-ouvertes.fr/hal-01659418

[17] M. BESSEMOULIN-CHATARD, M. HERDA, T. REY.*Hypocoercivity and diffusion limit of a finite volume scheme for linear kinetic equations*, in "Mathematics of Computation", September 2019, 39 pages [*DOI :* 10.1090/MCOM/3490], https://hal.archives-ouvertes.fr/hal-01957832

[18] O. BLONDEL, C. CANCÈS, M. SASADA, M. SIMON.*Convergence of a Degenerate Microscopic Dynamics to the Porous Medium Equation*, in "Annales de l'Institut Fourier", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-01710628

[19] D. BRESCH, M. GISCLON, I. LACROIX-VIOLET.*On Navier-Stokes-Korteweg and Euler-Korteweg Systems: Application to Quantum Fluids Models*, in "Archive for Rational Mechanics and Analysis", 2019, vol. 233, n$^{\mathrm{o}}$ 3, p. 975-1025, https://arxiv.org/abs/1703.09460 [*DOI :* 10.1007/S00205-019-01373-W], https://hal.archives-ouvertes.fr/hal-01496960

[20] C. CALGARO, C. COLIN, E. CREUSÉ.*A combined finite volume - finite element scheme for a low-Mach system involving a Joule term*, in "AIMS Mathematics", 2019, vol. 5, n$^{\mathrm{o}}$ 1, p. 311-331, forthcoming, https://hal.archives-ouvertes.fr/hal-02398893

[21] C. CALGARO, C. COLIN, E. CREUSÉ.*A combined Finite Volumes -Finite Elements method for a low-Mach model*, in "International Journal for Numerical Methods in Fluids", 2019, vol. 90, n$^{\mathrm{o}}$ 1, p. 1-21 [*DOI :* 10.1002/FLD.4706], https://hal.archives-ouvertes.fr/hal-01574894

[22] C. CALGARO, C. COLIN, E. CREUSÉ, E. ZAHROUNI.*Approximation by an iterative method of a low Mach model with temperature dependent viscosity*, in "Mathematical Methods in the Applied Sciences", 2019, vol. 42, n$^{\mathrm{o}}$ 1, p. 250-271 [*DOI :* 10.1002/MMA.5342], https://hal.archives-ouvertes.fr/hal-01801242

[23] C. CANCÈS, C. CHAINAIS-HILLAIRET, A. GERSTENMAYER, A. JÜNGEL.*Convergence of a Finite-Volume Scheme for a Degenerate Cross-Diffusion Model for Ion Transport*, in "Numerical Methods for Partial Differential Equations", 2019, vol. 35, n$^{\mathrm{o}}$ 2, p. 545-575, https://arxiv.org/abs/1801.09408 [*DOI :* 10.1002/NUM.22313], https://hal.archives-ouvertes.fr/hal-01695129

[24] C. CANCÈS, T. GALLOUËT, M. LABORDE, L. MONSAINGEON.*Simulation of multiphase porous media flows with minimizing movement and finite volume schemes*, in "European Journal of Applied Mathematics", 2019, vol. 30, n$^{\mathrm{o}}$ 6, p. 1123-1152 [*DOI :* 10.1017/S0956792518000633], https://hal.archives-ouvertes.fr/hal-01700952

[25] C. CANCÈS, D. MATTHES, F. NABET.*A two-phase two-fluxes degenerate Cahn-Hilliard model as constrained Wasserstein gradient flow*, in "Archive for Rational Mechanics and Analysis", 2019, vol. 233, n$^{\mathrm{o}}$ 2, p. 837–866 [*DOI :* 10.1007/S00205-019-01369-6], https://hal.archives-ouvertes.fr/hal-01665338

[26] C. CHAINAIS-HILLAIRET, M. HERDA.*Large-time behaviour of a family of finite volume schemes for boundary-driven convection-diffusion equations*, in "IMA Journal of Numerical Analysis", November 2019,

https://arxiv.org/abs/1810.01087 , forthcoming [*DOI :* 10.1093/IMANUM/DRZ037], https://hal.archives-ouvertes.fr/hal-01885015

[27] A. CHAMBOLLE, L. A. D. FERRARI, B. MERLET.*A phase-field approximation of the Steiner problem in dimension two*, in "Advances in Calculus of Variation", 2019, vol. 12, n^{o} 2, p. 157–179, https://arxiv.org/abs/1609.00519v1 - 27 pages, 8 figures [*DOI :* 10.1515/ACV-2016-0034], https://hal.archives-ouvertes.fr/hal-01359483

[28] A. CHAMBOLLE, L. A. D. FERRARI, B. MERLET.*Strong approximation in h-mass of rectifiable currents under homological constraint*, in "Advances in Calculus of Variation", 2019, https://arxiv.org/abs/1806.05046 , forthcoming [*DOI :* 10.1515/ACV-2018-0079], https://hal.archives-ouvertes.fr/hal-01813234

[29] A. CHAMBOLLE, L. A. D. FERRARI, B. MERLET.*Variational approximation of size-mass energies for k-dimensional currents*, in "ESAIM: Control, Optimisation and Calculus of Variations", 2019, vol. 25 (2019), n^{o} 43, 39, https://arxiv.org/abs/1710.08808 , forthcoming, https://hal.archives-ouvertes.fr/hal-01622540

[30] M. CICUTTIN, A. ERN, S. LEMAIRE.*A Hybrid High-Order method for highly oscillatory elliptic problems*, in "Computational Methods in Applied Mathematics", 2019, vol. 19, n^{o} 4, p. 723-748 [*DOI :* 10.1515/CMAM-2018-0013], https://hal.archives-ouvertes.fr/hal-01467434

[31] E. CREUSÉ, P. DULAR, S. NICAISE.*About the gauge conditions arising in Finite Element magnetostatic problems*, in "Computers and Mathematics with Applications", 2019, vol. 77, n^{o} 6, p. 1563-1582, https://hal.archives-ouvertes.fr/hal-01955649

[32] E. CREUSÉ, Y. LE MENACH, S. NICAISE, F. PIRIOU, R. TITTARELLI.*Two Guaranteed Equilibrated Error Estimators for Harmonic Formulations in Eddy Current Problems*, in "Computers and Mathematics with Applications", 2019, vol. 77, n^{o} 6, p. 1549-1562, https://hal.archives-ouvertes.fr/hal-01955692

[33] M. GOLDMAN, B. MERLET.*Recent results on non-convex functionals penalizing oblique oscillations*, in "Rendiconti del Seminario Matematico", 2019, https://hal.archives-ouvertes.fr/hal-02382214

[34] M. GOLDMAN, B. MERLET, V. MILLOT.*A Ginzburg-Landau model with topologically induced free discontinuities*, in "Annales de l'Institut Fourier", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-01643795

[35] M. HERDA, L. M. RODRIGUES.*Anisotropic Boltzmann-Gibbs dynamics of strongly magnetized Vlasov-Fokker-Planck equations*, in "Kinetic and Related Models ", 2019, vol. 12, n^{o} 3, p. 593-636, https://arxiv.org/abs/1610.05138 [*DOI :* 10.3934/KRM.2019024], https://hal.archives-ouvertes.fr/hal-01382854

[36] S. LEMAIRE.*Bridging the Hybrid High-Order and Virtual Element methods*, in "IMA Journal of Numerical Analysis", 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-01902962

[37] W. MELIS, T. REY, G. SAMAEY.*Projective and telescopic projective integration for the nonlinear BGK and Boltzmann equations*, in "SMAI Journal of Computational Mathematics", 2019, vol. 5, p. 53-88, https://arxiv.org/abs/1712.06362 [*DOI :* 10.5802/SMAI-JCM.43], https://hal.archives-ouvertes.fr/hal-01666346

[38] N. PETON, C. CANCÈS, D. GRANJEON, Q.-H. TRAN, S. WOLF.*Numerical scheme for a water flow-driven forward stratigraphic model*, in "Computational Geosciences", 2019, forthcoming [*DOI :* 10.1007/S10596-019-09893-W], https://hal.archives-ouvertes.fr/hal-01870347

[39] A. ZUREK.*Numerical approximation of a concrete carbonation model: study of the $\sqrt{t}$-law of propagation*, in "Numerical Methods for Partial Differential Equations", May 2019, vol. 35, n° 5, p. 1801-1820 [*DOI :* 10.1002/NUM.22377], https://hal.archives-ouvertes.fr/hal-01839277

International Conferences with Proceedings

[40] M. CICUTTIN, A. ERN, S. LEMAIRE.*On the implementation of a multiscale Hybrid High-Order method*, in "ENUMATH 2017", Bergen, Norway, I. BERRE, K. KUMAR, J. M. NORDBOTTEN, I. S. POP, F. A. RADU (editors), Numerical Mathematics and Advanced Applications - ENUMATH 2017, Springer, Cham, 2019, vol. 126, p. 509-517 [*DOI :* 10.1007/978-3-319-96415-7_46], https://hal.archives-ouvertes.fr/hal-01661925

[41] D. MATTHES, C. CANCÈS, F. NABET.*A degenerate Cahn-Hilliard model as constrained Wasserstein gradient flow*, in "GAMM annual meeting", Vienna, Austria, International Association for Applied Mathematics and Mechanics, November 2019, vol. 19, n° 1 [*DOI :* 10.1002/PAMM.201900158], https://hal.archives-ouvertes.fr/hal-02377146

Software

[42] A. MOUTON, C. CALGARO, E. CREUSÉ.*NS2DDV - Navier-Stokes 2D à Densité Variable*, September 2019 [SWH-ID : swh:1:dir:a126af9f1534e0b9f3431531a5f4751ad9b7b2fe], Software, https://hal.archives-ouvertes.fr/hal-02137040

Other Publications

[43] F. ALOUGES, A. DE BOUARD, B. MERLET, L. NICOLAS.*Stochastic homogenization of the Landau-Lifshitz-Gilbert equation*, February 2019, https://arxiv.org/abs/1902.05743 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02020241

[44] N. AYI, M. HERDA, H. HIVERT, I. TRISTANI.*A note on hypocoercivity for kinetic equations with heavy-tailed equilibrium*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02389146

[45] S. BILLIARD, M. DEREX, L. MAISONNEUVE, T. REY.*Convergence of knowledge in a cultural evolution model with population structure, random social learning and credibility biases*, November 2019, 25 pages, https://hal.archives-ouvertes.fr/hal-02357188

[46] C. CALGARO, E. CREUSÉ.*A finite volume method for a convection- diffusion equation involving a Joule term*, 2019, working paper or preprint, https://hal.inria.fr/hal-02432936

[47] M. CAMPOS PINTO, F. CHARLES, B. DESPRÉS, M. HERDA.*A projection algorithm on the set of polynomials with two bounds*, May 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02128851

[48] C. CANCÈS, C. CHAINAIS-HILLAIRET, J. FUHRMANN, B. GAUDEUL.*A numerical analysis focused comparison of several Finite Volume schemes for an Unipolar Degenerated Drift-Diffusion Model*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02194604

[49] C. CANCÈS, C. CHAINAIS-HILLAIRET, M. HERDA, S. KRELL.*Large time behavior of nonlinear finite volume schemes for convection-diffusion equations*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02360155

[50] C. Cancès, T. O. Gallouët, G. Todeschi. *A variational finite volume scheme for Wasserstein gradient flows*, July 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02189050

[51] C. Cancès, B. Gaudeul. *Entropy diminishing finite volume approximation of a cross-diffusion system*, 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02418908

[52] C. Cancès, D. Maltese. *A gravity current model with capillary trapping for oil migration in multilayer geological basins*, August 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02272965

[53] C. Cancès, F. Nabet. *Energy stable discretization for two-phase porous media flows*, January 2020, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02442233

[54] C. Cancès, F. Nabet, M. Vohralík. *Convergence and a posteriori error analysis for energy-stable finite element approximations of degenerate parabolic equations*, January 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01894884

[55] C. Chainais-Hillairet, M. Herda. *L^∞ bounds for numerical solutions of noncoercive convection-diffusion equations*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02404546

[56] C. Chainais-Hillairet, S. Krell. *Exponential decay to equilibrium of nonlinear DDFV schemes for convection-diffusion equations*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02408212

[57] F. Chave. *Numerical study of the fracture diffusion-dispersion coefficient for passive transport in fractured porous media*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02412691

[58] F. Chave, D. A. Di Pietro, S. Lemaire. *A three-dimensional Hybrid High-Order method for magnetostatics*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02407175

[59] B. Després, M. Herda. *Computation of sum of squares polynomials from data points*, August 2019, https://arxiv.org/abs/1812.02444 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01946539

[60] G. Dujardin, I. Lacroix-Violet. *High order linearly implicit methods for evolution equations: How to solve an ODE by inverting only linear systems*, November 2019, https://arxiv.org/abs/1911.06016 - working paper or preprint, https://hal.inria.fr/hal-02361814

[61] A. El Keurti, T. Rey. *Finite Volume Method for a System of Continuity Equations Driven by Nonlocal Interactions*, December 2019, https://arxiv.org/abs/1912.06423 - 8 pages, https://hal.archives-ouvertes.fr/hal-02408246

[62] M. Goldman, B. Merlet. *Non-convex functionals penalizing simultaneous oscillations along independent directions: rigidity estimates*, May 2019, https://arxiv.org/abs/1905.07905 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02132896

References in notes

[63] R. ABGRALL.*A review of residual distribution schemes for hyperbolic and parabolic problems: the July 2010 state of the art*, in "Commun. Comput. Phys.", 2012, vol. 11, n° 4, p. 1043–1080, http://dx.doi.org/10.4208/cicp.270710.130711s

[64] R. ABGRALL, G. BAURIN, A. KRUST, D. DE SANTIS, M. RICCHIUTO.*Numerical approximation of parabolic problems by residual distribution schemes*, in "Internat. J. Numer. Methods Fluids", 2013, vol. 71, n° 9, p. 1191–1206, http://dx.doi.org/10.1002/fld.3710

[65] R. ABGRALL, A. LARAT, M. RICCHIUTO.*Construction of very high order residual distribution schemes for steady inviscid flow problems on hybrid unstructured meshes*, in "J. Comput. Phys.", 2011, vol. 230, n° 11, p. 4103–4136, http://dx.doi.org/10.1016/j.jcp.2010.07.035

[66] R. ABGRALL, A. LARAT, M. RICCHIUTO, C. TAVÉ.*A simple construction of very high order non-oscillatory compact schemes on unstructured meshes*, in "Comput. & Fluids", 2009, vol. 38, n° 7, p. 1314–1323, http://dx.doi.org/10.1016/j.compfluid.2008.01.031

[67] T. AIKI, A. MUNTEAN.*A free-boundary problem for concrete carbonation: front nucleation and rigorous justification of the $\sqrt{t}$-law of propagation*, in "Interfaces Free Bound.", 2013, vol. 15, n° 2, p. 167–180, http://dx.doi.org/10.4171/IFB/299

[68] B. AMAZIANE, A. BERGAM, M. EL OSSMANI, Z. MGHAZLI.*A posteriori estimators for vertex centred finite volume discretization of a convection-diffusion-reaction equation arising in flow in porous media*, in "Internat. J. Numer. Methods Fluids", 2009, vol. 59, n° 3, p. 259–284, http://dx.doi.org/10.1002/fld.1456

[69] M. AVILA, J. PRINCIPE, R. CODINA.*A finite element dynamical nonlinear subscale approximation for the low Mach number flow equations*, in "J. Comput. Phys.", 2011, vol. 230, n° 22, p. 7988–8009, http://dx.doi.org/10.1016/j.jcp.2011.06.032

[70] I. BABUŠKA, W. C. RHEINBOLDT.*Error estimates for adaptive finite element computations*, in "SIAM J. Numer. Anal.", 1978, vol. 15, n° 4, p. 736–754

[71] C. BATAILLON, F. BOUCHON, C. CHAINAIS-HILLAIRET, C. DESGRANGES, E. HOARAU, F. MARTIN, S. PERRIN, M. TUPIN, J. TALANDIER.*Corrosion modelling of iron based alloy in nuclear waste repository*, in "Electrochim. Acta", 2010, vol. 55, n° 15, p. 4451–4467

[72] J. BEAR, Y. BACHMAT.*Introduction to modeling of transport phenomena in porous media*, Springer, 1990, vol. 4

[73] J. BEAR.*Dynamic of Fluids in Porous Media*, American Elsevier, New York, 1972

[74] A. BECCANTINI, E. STUDER, S. GOUNAND, J.-P. MAGNAUD, T. KLOCZKO, C. CORRE, S. KUDRIAKOV.*Numerical simulations of a transient injection flow at low Mach number regime*, in "Internat. J. Numer. Methods Engrg.", 2008, vol. 76, n° 5, p. 662–696, http://dx.doi.org/10.1002/nme.2331

[75] L. BEIRÃO DA VEIGA, F. BREZZI, A. CANGIANI, G. MANZINI, L. D. MARINI, A. RUSSO.*Basic principles of virtual element methods*, in "Math. Models Methods Appl. Sci. (M3AS)", 2013, vol. 23, n° 1, p. 199–214

[76] J.-D. BENAMOU, G. CARLIER, M. LABORDE.*An augmented Lagrangian approach to Wasserstein gradient flows and applications*, in "Gradient flows: from theory to application", ESAIM Proc. Surveys, EDP Sci., Les Ulis, 2016, vol. 54, p. 1–17, https://doi.org/10.1051/proc/201654001

[77] S. BERRONE, V. GARBERO, M. MARRO.*Numerical simulation of low-Reynolds number flows past rectangular cylinders based on adaptive finite element and finite volume methods*, in "Comput. & Fluids", 2011, vol. 40, p. 92–112, http://dx.doi.org/10.1016/j.compfluid.2010.08.014

[78] C. BESSE.*Analyse numérique des systèmes de Davey–Stewartson*, Université Bordeaux 1, 1998

[79] M. BESSEMOULIN-CHATARD, C. CHAINAIS-HILLAIRET, M.-H. VIGNAL.*Study of a fully implicit scheme for the drift-diffusion system. Asymptotic behavior in the quasi-neutral limit*, in "SIAM, J. Numer. Anal.", 2014, vol. 52, n^{o} 4, https://epubs.siam.org/doi/abs/10.1137/130913432

[80] D. BRESCH, P. NOBLE, J.-P. VILA.*Relative entropy for compressible Navier-Stokes equations with density dependent viscosities and various applications*, in "LMLFN 2015—low velocity flows—application to low Mach and low Froude regimes", ESAIM Proc. Surveys, EDP Sci., Les Ulis, 2017, vol. 58, p. 40–57

[81] C. CALGARO, E. CREUSÉ, T. GOUDON.*An hybrid finite volume-finite element method for variable density incompressible flows*, in "J. Comput. Phys.", 2008, vol. 227, n^{o} 9, p. 4671–4696

[82] C. CALGARO, E. CREUSÉ, T. GOUDON.*Modeling and simulation of mixture flows: application to powder-snow avalanches*, in "Comput. & Fluids", 2015, vol. 107, p. 100–122, http://dx.doi.org/10.1016/j.compfluid.2014.10.008

[83] C. CANCÈS, C. GUICHARD.*Convergence of a nonlinear entropy diminishing Control Volume Finite Element scheme for solving anisotropic degenerate parabolic equations*, in "Mathematics of Computation", 2016, vol. 85, n^{o} 298, p. 549-580, https://hal.archives-ouvertes.fr/hal-00955091

[84] C. CANCÈS, I. S. POP, M. VOHRALÍK.*An a posteriori error estimate for vertex-centered finite volume discretizations of immiscible incompressible two-phase flow*, in "Math. Comp.", 2014, vol. 83, n^{o} 285, p. 153–188, http://dx.doi.org/10.1090/S0025-5718-2013-02723-8

[85] J. A. CARRILLO, A. JÜNGEL, P. A. MARKOWICH, G. TOSCANI, A. UNTERREITER.*Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities*, in "Monatsh. Math.", 2001, vol. 133, n^{o} 1, p. 1–82, http://dx.doi.org/10.1007/s006050170032

[86] C. CHAINAIS-HILLAIRET.*Entropy method and asymptotic behaviours of finite volume schemes*, in "Finite volumes for complex applications. VII. Methods and theoretical aspects", Springer Proc. Math. Stat., Springer, Cham, 2014, vol. 77, p. 17–35

[87] C. CHAINAIS-HILLAIRET, A. JÜNGEL, S. SCHUCHNIGG.*Entropy-dissipative discretization of nonlinear diffusion equations and discrete Beckner inequalities*, in "Modelisation Mathématique et Analyse Numérique", 2016, vol. 50, n^{o} 1, p. 135-162, https://hal.archives-ouvertes.fr/hal-00924282

[88] E. CREUSÉ, S. NICAISE, Z. TANG, Y. LE MENACH, N. NEMITZ, F. PIRIOU.*Residual-based a posteriori estimators for the $\mathbf{A}-\phi$ magnetodynamic harmonic formulation of the Maxwell system*, in "Math. Models Methods Appl. Sci.", 2012, vol. 22, n^{o} 5, p. 1150028-30, http://dx.doi.org/10.1142/S021820251150028X

[89] E. CREUSÉ, S. NICAISE, Z. TANG, Y. LE MENACH, N. NEMITZ, F. PIRIOU.*Residual-based a posteriori estimators for the $\mathbf{T}/\Omega$ magnetodynamic harmonic formulation of the Maxwell system*, in "Int. J. Numer. Anal. Model.", 2013, vol. 10, n$^{\mathrm{o}}$ 2, p. 411–429

[90] E. CREUSÉ, S. NICAISE, E. VERHILLE.*Robust equilibrated a posteriori error estimators for the Reissner-Mindlin system*, in "Calcolo", 2011, vol. 48, n$^{\mathrm{o}}$ 4, p. 307–335, http://dx.doi.org/10.1007/s10092-011-0042-0

[91] B. DESPRÉS.*Polynomials with bounds and numerical approximation*, in "Numerical Algorithms", 2017, vol. 76, n$^{\mathrm{o}}$ 3, p. 829–859, https://doi.org/10.1007/s11075-017-0286-0

[92] B. DESPRÉS, M. HERDA.*Correction to: Polynomials with bounds and numerical approximation*, in "Numerical Algorithms", 2018, vol. 77, n$^{\mathrm{o}}$ 1, p. 309–311, https://doi.org/10.1007/s11075-017-0441-7

[93] D. A. DI PIETRO, A. ERN, S. LEMAIRE.*An arbitrary-order and compact-stencil discretization of diffusion on general meshes based on local reconstruction operators*, in "Comput. Methods Appl. Math.", 2014, vol. 14, n$^{\mathrm{o}}$ 4, p. 461–472, https://doi.org/10.1515/cmam-2014-0018

[94] D. A. DI PIETRO, M. VOHRALÍK.*A Review of Recent Advances in Discretization Methods, a Posteriori Error Analysis, and Adaptive Algorithms for Numerical Modeling in Geosciences*, in "Oil & Gas Science and Technology-Rev. IFP", June 2014, p. 1-29, (online first)

[95] J. DOLBEAULT, C. MOUHOT, C. SCHMEISER.*Hypocoercivity for linear kinetic equations conserving mass*, in "Trans. Amer. Math. Soc.", 2015, vol. 367, n$^{\mathrm{o}}$ 6, p. 3807–3828, https://doi.org/10.1090/S0002-9947-2015-06012-7

[96] V. DOLEJŠÍ, A. ERN, M. VOHRALÍK.*A framework for robust a posteriori error control in unsteady nonlinear advection-diffusion problems*, in "SIAM J. Numer. Anal.", 2013, vol. 51, n$^{\mathrm{o}}$ 2, p. 773–793, http://dx.doi.org/10.1137/110859282

[97] D. DONATELLI, E. FEIREISL, P. MARCATI.*Well/ill posedness for the Euler–Korteweg–Poisson system and related problems*, in "Comm. Partial Differential Equations", 2015, vol. 40, p. 1314-1335

[98] J. DRONIOU.*Finite volume schemes for diffusion equations: introduction to and review of modern methods*, in "Math. Models Methods Appl. Sci.", 2014, vol. 24, n$^{\mathrm{o}}$ 8, p. 1575-1620

[99] E. EMMRICH.*Two-step BDF time discretisation of nonlinear evolution problems governed by monotone operators with strongly continuous perturbations*, in "Comput. Methods Appl. Math.", 2009, vol. 9, n$^{\mathrm{o}}$ 1, p. 37–62

[100] J. GIESSELMANN, C. LATTANZIO, A.-E. TZAVARAS.*Relative energy for the Korteweg theory and related Hamiltonian flows in gas dynamics*, in "Arch. Rational Mech. Analysis", 2017, vol. 223, p. 1427-1484

[101] V. GRAVEMEIER, W. A. WALL.*Residual-based variational multiscale methods for laminar, transitional and turbulent variable-density flow at low Mach number*, in "Internat. J. Numer. Methods Fluids", 2011, vol. 65, n$^{\mathrm{o}}$ 10, p. 1260–1278, http://dx.doi.org/10.1002/fld.2242

[102] L. GREENGARD, J.-Y. LEE.*Accelerating the nonuniform fast Fourier transform*, in "SIAM Rev.", 2004, vol. 46, n$^{\mathrm{o}}$ 3, p. 443–454, http://dx.doi.org/10.1137/S003614450343200X

[103] F. GUILLÉN-GONZÁLEZ, P. DAMÁZIO, M. A. ROJAS-MEDAR.*Approximation by an iterative method for regular solutions for incompressible fluids with mass diffusion*, in "J. Math. Anal. Appl.", 2007, vol. 326, n^{o} 1, p. 468–487, http://dx.doi.org/10.1016/j.jmaa.2006.03.009

[104] F. GUILLÉN-GONZÁLEZ, M. SY.*Iterative method for mass diffusion model with density dependent viscosity*, in "Discrete Contin. Dyn. Syst. Ser. B", 2008, vol. 10, n^{o} 4, p. 823–841, http://dx.doi.org/10.3934/dcdsb.2008.10.823

[105] V. HEUVELINE.*On higher-order mixed FEM for low Mach number flows: application to a natural convection benchmark problem*, in "Internat. J. Numer. Methods Fluids", 2003, vol. 41, n^{o} 12, p. 1339–1356, http://dx.doi.org/10.1002/fld.454

[106] M. E. HUBBARD, M. RICCHIUTO.*Discontinuous upwind residual distribution: a route to unconditional positivity and high order accuracy*, in "Comput. & Fluids", 2011, vol. 46, p. 263–269, http://dx.doi.org/10.1016/j.compfluid.2010.12.023

[107] S. JIN.*Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations*, in "SIAM, J. Sci. Comput.", 1999, vol. 21, p. 441-454

[108] P. LE QUÉRÉ, R. MASSON, P. PERROT.*A Chebyshev collocation algorithm for 2D non-Boussinesq convection*, in "Journal of Computational Physics", 1992, vol. 103, p. 320-335

[109] A. MAJDA, J. SETHIAN.*The derivation and numerical solution of the equations for zero Mach number combustion*, in "Combustion Science and Technology", 1985, vol. 42, p. 185–205

[110] F. OTTO.*The geometry of dissipative evolution equations: the porous medium equation*, in "Comm. Partial Differential Equations", 2001, vol. 26, n^{o} 1-2, p. 101–174

[111] M. RICCHIUTO, R. ABGRALL.*Explicit Runge-Kutta residual distribution schemes for time dependent problems: second order case*, in "J. Comput. Phys.", 2010, vol. 229, n^{o} 16, p. 5653–5691, http://dx.doi.org/10.1016/j.jcp.2010.04.002

[112] D. RUPPEL, E. SACKMANN.*On defects in different phases of two-dimensional lipid bilayers*, in "J. Phys. France", 1983, vol. 44, n^{o} 9, p. 1025-1034 [*DOI :* 10.1051/JPHYS:019830044090102500], https://jphys.journaldephysique.org/articles/jphys/abs/1983/09/jphys_1983__44_9_1025_0/jphys_1983__44_9_1025_0.html

[113] M. VOHRALÍK.*Residual flux-based a posteriori error estimates for finite volume and related locally conservative methods*, in "Numer. Math.", 2008, vol. 111, n^{o} 1, p. 121–158, http://dx.doi.org/10.1007/s00211-008-0168-4

[114] J. DE FRUTOS, B. GARCÍA-ARCHILLA, J. NOVO.*A posteriori error estimations for mixed finite-element approximations to the Navier-Stokes equations*, in "J. Comput. Appl. Math.", 2011, vol. 236, n^{o} 6, p. 1103–1122, http://dx.doi.org/10.1016/j.cam.2011.07.033

Project-Team RMOD

Analyses and Languages Constructs for Object-Oriented Application Evolution

IN COLLABORATION WITH: Centre de Recherche en Informatique, Signal et Automatique de Lille

IN PARTNERSHIP WITH:
Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Distributed programming and Software engineering

Table of contents

Project-Team RMOD

Creation of the Project-Team: 2009 July 01

Keywords:

Computer Science and Digital Science:

A1.3.3. - Blockchain
A2. - Software
A2.1. - Programming Languages
A2.1.3. - Object-oriented programming
A2.1.8. - Aspect-oriented programming
A2.1.10. - Domain-specific languages
A2.1.12. - Dynamic languages
A2.3.1. - Embedded systems
A2.5. - Software engineering
A2.5.1. - Software Architecture & Design
A2.5.3. - Empirical Software Engineering
A2.5.4. - Software Maintenance & Evolution
A2.6. - Infrastructure software
A2.6.3. - Virtual machines

Other Research Topics and Application Domains:

B2. - Health
B2.7. - Medical devices
B5. - Industry of the future
B5.9. - Industrial maintenance
B6.5. - Information systems
B7. - Transport and logistics

1. Team, Visitors, External Collaborators

Research Scientists

Stéphane Ducasse [Team leader, Inria, Senior Researcher, HDR]
Steven Costiou [Inria, Researcher, from Oct 2019]
Marcus Denker [Inria, Researcher]

Faculty Members

Nicolas Anquetil [Université de Lille, Associate Professor, HDR]
Vincent Aranega [Université de Lille, Associate Professor]
Anne Etien [Université de Lille, Associate Professor, HDR]
Damien Pollet [Université de Lille, Associate Professor]

External Collaborators

Luc Fabresse [IMT Lille-Douai]
Pavel Krivanek [Inria, until Jun 2019]
Guillermo Polito [CNRS]
Esteban Lorenzano [InriaSoft]
Olivier Auverlot [Université de Lille]

Technical Staff

Andy Amoordon [Inria, Engineer, from Apr 2019 until Jul 2019]
Santiago Bragagnolo [Inria, Engineer, from Apr 2019]
Cyril Ferlicot-Delbecque [Inria, Engineer]
Guillaume Larchevêque [Inria, Engineer, until Oct 2019]
Allex Oliveira [Inria, Engineer]
Pablo Tesone [Inria, Engineer, from Apr 2019]
Christophe Demarey [Inria, Engineer, 60%]

PhD Students

Lionel Akue [Inria, PhD Student, until Apr 2019]
Julien Delplanque [Université de Lille, PhD Student]
Thomas Dupriez [Université de Lille, PhD Student]
Carolina Hernández Phillips [Inria, PhD Student]
Mahugnon Honoré Houekpetodji [Inria, PhD Student]
Jason Lecerf [CEA, PhD Student, until Nov 2019]
Matteo Marra [VUB Brussels, Belgium, PhD Student]
Pierre Misse Chanabier [Inria, PhD Student, from Oct 2019]
Théo Rogliano [Inria, PhD Student, from Oct 2019]
Pablo Tesone [École des Mines de Douai, PhD Student, until Mar 2019]
Benoît Verhaeghe [Berger-Levrault, PhD Student]
Oleksandr Zaitsev [Inria, PhD Student, from Mar 2019]

Post-Doctoral Fellow

Steven Costiou [Inria, Post-Doctoral Fellow, until Sep 2019]

Visiting Scientists

Abdelhakim Bouremel [University of Skikda, Algeria, Oct 2019]
Mohamad Chakroun [Independent, Mar 2019]
Victor Martín Dias [University of Chile, Chile, until Sep 2019]
Christopher Fuhrman [École de technologie supérieure de Montréal, Canada, until Sep 2019]
Yann-Gaël Guéhéneuc [Concordia University, Canada, from Apr 2019 until May 2019]
Sebastijan Kaplar [University of Novi Sad, Serbia, Aug 2019]
Manuel Leuenberger [University of Bern, Switzerland, from Sep 2019 until Nov 2019]
Milton Mamani Torres [Object Profile SpA, Chile, Aug 2019]
Nina Medic [University of Novi Sad, Serbia, from Jun 2019 until Jul 2019]
Hayatou Oumarou [University of Maroua, Cameroun from Sep 2019 until Oct 2019]
Giuseppe Antonio Pierro [University of Cagliari, Italy, until July 2019]
Gordana Rakic [University of Novi Sad, Serbia, from Nov 2019]
Moussa Saker [University Badji Mokhtar-Annaba, Algeria, from Nov 2019]

Administrative Assistant

Julie Jonas [Inria]

2. Overall Objectives

2.1. Introduction

Keywords: Software evolution, Maintenance, Program visualization, Program analyses, Meta modelling, Software metrics, Quality models, Object-oriented programming, Reflective programming, Traits, Dynamically typed languages, Dynamic software update, Pharo, Moose.

RMoD's general vision is defined in two objectives: remodularization and modularity constructs. These two views are the two faces of a same coin: maintenance could be eased with better engineering and analysis tools and programming language constructs could let programmers define more modular applications.

2.2. Reengineering and remodularization

While applications must evolve to meet new requirements, few approaches analyze the implications of their original structure (modules, packages, classes) and their transformation to support their evolution. Our research focuses on the *remodularization* of object-oriented applications. Automated approaches including clustering algorithms are not satisfactory because they often ignore user inputs. Our vision is that we need better approaches to support the transformation of existing software. The reengineering challenge tackled by RMoD is formulated as follows:

How to help remodularize existing software applications?

We are developing analyses and algorithms to remodularize object-oriented applications. This is why we started studying and building tools to support the *understanding of applications* at the level of packages and modules. This allows us to understand the results of the *analyses* that we are building.

We seek to create tools to help developers perform large refactoring. How can they keep track of changes in various locations in a system while ensuring *integrity of current and new code* by *uniformally applying new design choices*.

2.3. Constructs for modular and isolating programming languages

Dynamically-typed programming languages such as JavaScript are getting new attention as illustrated by the large investment of Google in the development of the Chrome V8 JavaScript engine and the development of a new dynamic language DART. This new trend is correlated to the increased adoption of dynamic programming languages for web-application development, as illustrated by Ruby on Rails, PHP and JavaScript. With web applications, users expect applications to be always available and getting updated on the fly. This continuous evolution of application is a real challenge [40]. Hot software evolution often requires *reflective* behavior and features. For instance in CLOS and Smalltalk each class modification automatically migrates existing instances on the fly.

At the same time, there is a need for *software isolation*, i.e., applications should reliably run co-located with other applications in the same virtual machine with neither confidential information leaks nor vulnerabilities. Indeed, often for economical reasons, web servers run multiple applications on the same virtual machine. Users need confined applications. It is important that (1) an application does not access information of other applications running on the same virtual machine and (2) an application authorized to manipulate data cannot pass such authorization or information to other parts of the application that should not get access to it.

Static analysis tools have always been confronted to reflection [37]. Without a full treatment of reflection, static analysis tools are both incomplete and unsound. Incomplete because some parts of the program may not be included in the application call graph, and unsound because the static analysis does not take into account reflective features [46]. In reflective languages such as F-Script, Ruby, Python, Lua, JavaScript, Smalltalk and Java (to a certain extent), it is possible to nearly change any aspect of an application: change objects, change classes dynamically, migrate instances, and even load untrusted code.

Reflection and isolation concerns are a priori antagonistic, pulling language design in two opposite directions. Isolation, on the one hand, pulls towards more static elements and types (e.g., ownership types). Reflection, on the other hand, pulls towards fully dynamic behavior. This tension is what makes this a real challenge: As experts in reflective programming, dynamic languages and modular systems, we believe that by working on this important tension we can make a breakthrough and propose innovative solutions in resolving or mitigating this tension. With this endeavor, we believe that we are working on a key challenge that can have an impact on future programming languages. The language construct challenge tackled by RMoD is formulated as follows:

What are the language modularity constructs to support isolation?

In parallel we are continuing our research effort on traits [0] by assessing trait scalability and reuse on a large case study and developing a pure trait-based language. In addition, we dedicate efforts to remodularizing a

[0] Traits are groups of methods that can be composed orthogonally to simple inheritance. Contrary to mixin, the class has the control of the composition and conflict management.

meta-level architecture in the context of the design of an isolating dynamic language. Indeed at the extreme, modules and structural control of reflective features are the first steps towards flexible, dynamic, yet isolating, languages. As a result, we expect to demonstrate that having adequate composable units and scoping units will help the evolution and recomposition of an application.

3. Research Program

3.1. Software Reengineering

Strong coupling among the parts of an application severely hampers its evolution. Therefore, it is crucial to answer the following questions: How to support the substitution of certain parts while limiting the impact on others? How to identify reusable parts? How to modularize an object-oriented application?

Having good classes does not imply a good application layering, absence of cycles between packages and reuse of well-identified parts. Which notion of cohesion makes sense in presence of late-binding and programming frameworks? Indeed, frameworks define a context that can be extended by subclassing or composition: in this case, packages can have a low cohesion without being a problem for evolution. How to obtain algorithms that can be used on real cases? Which criteria should be selected for a given remodularization?

To help us answer these questions, we work on enriching Moose, our reengineering environment, with a new set of analyses [31], [30]. We decompose our approach in three main and potentially overlapping steps:

1. Tools for understanding applications,
2. Remodularization analyses,
3. Software Quality.

3.1.1. *Tools for understanding applications*

Context and Problems. We are studying the problems raised by the understanding of applications at a larger level of granularity such as packages or modules. We want to develop a set of conceptual tools to support this understanding.

Some approaches based on Formal Concept Analysis (FCA) [59] show that such an analysis can be used to identify modules. However the presented examples are too small and not representative of real code.

Research Agenda.

FCA provides an important approach in software reengineering for software understanding, design anomalies detection and correction, but it suffers from two problems: (i) it produces lattices that must be interpreted by the user according to his/her understanding of the technique and different elements of the graph; and, (ii) the lattice can rapidly become so big that one is overwhelmed by the mass of information and possibilities [20]. We look for solutions to help people putting FCA to real use.

3.1.2. *Remodularization analyses*

Context and Problems. It is a well-known practice to layer applications with bottom layers being more stable than top layers [47]. Until now, few works have attempted to identify layers in practice: Mudpie [61] is a first cut at identifying cycles between packages as well as package groups potentially representing layers. DSM (dependency structure matrix) [60], [55] seems to be adapted for such a task but there is no serious empirical experience that validates this claim. From the side of remodularization algorithms, many were defined for procedural languages [43]. However, object-oriented programming languages bring some specific problems linked with late-binding and the fact that a package does not have to be systematically cohesive since it can be an extension of another one [62], [34].

As we are designing and evaluating algorithms and analyses to remodularize applications, we also need a way to understand and assess the results we are obtaining.

Research Agenda. We work on the following items:

Layer identification. We propose an approach to identify layers based on a semi-automatic classification of package and class interrelationships that they contain. However, taking into account the wish or knowledge of the designer or maintainer should be supported.

Cohesion Metric Assessment. We are building a validation framework for cohesion/coupling metrics to determine whether they actually measure what they promise to. We are also compiling a number of traditional metrics for cohesion and coupling quality metrics to evaluate their relevance in a software quality setting.

3.1.3. Software Quality

Research Agenda. Since software quality is fuzzy by definition and a lot of parameters should be taken into account we consider that defining precisely a unique notion of software quality is definitively a Grail in the realm of software engineering. The question is still relevant and important. We work on the two following items:

Quality models. We studied existing quality models and the different options to combine indicators — often, software quality models happily combine metrics, but at the price of losing the explicit relationships between the indicator contributions. There is a need to combine the results of one metric over all the software components of a system, and there is also the need to combine different metric results for any software component. Different combination methods are possible that can give very different results. It is therefore important to understand the characteristics of each method.

Bug prevention. Another aspect of software quality is validating or monitoring the source code to avoid the emergence of well known sources of errors and bugs. We work on how to best identify such common errors, by trying to identify earlier markers of possible errors, or by helping identifying common errors that programmers did in the past.

3.2. Language Constructs for Modular Design

While the previous axis focuses on how to help remodularizing existing software, this second research axis aims at providing new language constructs to build more flexible and recomposable software. We will build on our work on traits [57], [32] and classboxes [21] but also start to work on new areas such as isolation in dynamic languages. We will work on the following points: (1) Traits and (2) Modularization as a support for isolation.

3.2.1. Traits-based program reuse

Context and Problems. Inheritance is well-known and accepted as a mechanism for reuse in object-oriented languages. Unfortunately, due to the coarse granularity of inheritance, it may be difficult to decompose an application into an optimal class hierarchy that maximizes software reuse. Existing schemes based on single inheritance, multiple inheritance, or mixins, all pose numerous problems for reuse.

To overcome these problems, we designed a new composition mechanism called Traits [57], [32]. Traits are pure units of behavior that can be composed to form classes or other traits. The trait composition mechanism is an alternative to multiple or mixin inheritance in which the composer has full control over the trait composition. The result enables more reuse than single inheritance without introducing the drawbacks of multiple or mixin inheritance. Several extensions of the model have been proposed [29], [51], [22], [33] and several type systems were defined [35], [58], [52], [45].

Traits are reusable building blocks that can be explicitly composed to share methods across unrelated class hierarchies. In their original form, traits do not contain state and cannot express visibility control for methods. Two extensions, stateful traits and freezable traits, have been proposed to overcome these limitations. However, these extensions are complex both to use for software developers and to implement for language designers.

Research Agenda: Towards a pure trait language. We plan distinct actions: (1) a large application of traits, (2) assessment of the existing trait models and (3) bootstrapping a pure trait language.

- To evaluate the expressiveness of traits, some hierarchies were refactored, showing code reuse [24]. However, such large refactorings, while valuable, may not exhibit all possible composition problems, since the hierarchies were previously expressed using single inheritance and following certain patterns. We want to redesign from scratch the collection library of Smalltalk (or part of it). Such a redesign should on the one hand demonstrate the added value of traits on a real large and redesigned library and on the other hand foster new ideas for the bootstrapping of a pure trait-based language.

  In particular we want to reconsider the different models proposed (stateless [32], stateful [23], and freezable [33]) and their operators. We will compare these models by (1) implementing a trait-based collection hierarchy, (2) analyzing several existing applications that exhibit the need for traits. Traits may be flattened [50]. This is a fundamental property that confers to traits their simplicity and expressiveness over Eiffel's multiple inheritance. Keeping these aspects is one of our priority in forthcoming enhancements of traits.
- Alternative trait models. This work revisits the problem of adding state and visibility control to traits. Rather than extending the original trait model with additional operations, we use a fundamentally different approach by allowing traits to be lexically nested within other modules. This enables traits to express (shared) state and visibility control by hiding variables or methods in their lexical scope. Although the traits' "flattening property" no longer holds when they can be lexically nested, the combination of traits with lexical nesting results in a simple and more expressive trait model. We formally specify the operational semantics of this combination. Lexically nested traits are fully implemented in AmbientTalk, where they are used among others in the development of a Morphic-like UI framework.
- We want to evaluate how inheritance can be replaced by traits to form a new object model. For this purpose we will design a minimal reflective kernel, inspired first from ObjVlisp [28] then from Smalltalk [38].

3.2.2. Reconciling Dynamic Languages and Isolation

Context and Problems. More and more applications require dynamic behavior such as modification of their own execution (often implemented using reflective features [42]). For example, F-script allows one to script Cocoa Mac-OS X applications and Lua is used in Adobe Photoshop. Now in addition more and more applications are updated on the fly, potentially loading untrusted or broken code, which may be problematic for the system if the application is not properly isolated. Bytecode checking and static code analysis are used to enable isolation, but such approaches do not really work in presence of dynamic languages and reflective features. Therefore there is a tension between the need for flexibility and isolation.

Research Agenda: Isolation in dynamic and reflective languages. To solve this tension, we will work on *Sure*, a language where isolation is provided by construction: as an example, if the language does not offer field access and its reflective facilities are controlled, then the possibility to access and modify private data is controlled. In this context, layering and modularizing the meta-level [25], as well as controlling the access to reflective features [26], [27] are important challenges. We plan to:

- Study the isolation abstractions available in erights (http://www.erights.org) [49], [48], and Java's class loader strategies [44], [39].
- Categorize the different reflective features of languages such as CLOS [41], Python and Smalltalk [53] and identify suitable isolation mechanisms and infrastructure [36].
- Assess different isolation models (access rights, capabilities [54],...) and identify the ones adapted to our context as well as different access and right propagation.
- Define a language based on
  - the decomposition and restructuring of the reflective features [25],

– the use of encapsulation policies as a basis to restrict the interfaces of the controlled objects [56],
– the definition of method modifiers to support controlling encapsulation in the context of dynamic languages.

An open question is whether, instead of providing restricted interfaces, we could use traits to grant additional behavior to specific instances: without trait application, the instances would only exhibit default public behavior, but with additional traits applied, the instances would get extra behavior. We will develop *Sure*, a modular extension of the reflective kernel of Smalltalk (since it is one of the languages offering the largest set of reflective features such as pointer swapping, class changing, class definition,...) [53].

4. Application Domains

4.1. Programming Languages and Tools

Many of the results of RMoD are improving programming languages or development tools for such languages. As such the application domain of these results is as varied as the use of programming languages in general. Pharo, the language that RMoD develops, is used for a very broad range of applications. From pure research experiments to real world industrial use (the Pharo Consortium has more than 25 company members).

Examples are web applications, server backends for mobile applications or even graphical tools and embedded applications

4.2. Software Reengineering

Moose is a language-independent environment for reverse and re-engineering complex software systems. Moose provides a set of services including a common meta-model, metrics evaluation and visualization. As such Moose is used for analyzing software systems to support understanding and continuous development as well as software quality analysis.

5. Highlights of the Year

5.1. Highlights of the Year

- Steven Costiou was hired as CR.
- We released Pharo 7. More information at http://pharo.org.
- The paper *Rotten Green Tests* has been accepted at ICSE.
  https://hal.inria.fr/hal-02002346

5.1.1. Awards

- Best paper award: *SATToSE 2019 Migrating GWT to angular 6 using MDE.*
  https://hal.inria.fr/hal-02304301
- 2nd place best paper award IWST 2019: *Illicium: a modular transpilation toolchain from Pharo to C.*
  https://hal.archives-ouvertes.fr/hal-02297860
- 3rd place best paper award IWST 2019: *GildaVM: a Non-Blocking I/O Architecture for the Cog VM.*
  https://hal.archives-ouvertes.fr/view/index/docid/2379275

6. New Software and Platforms

6.1. Moose

Moose: Software and Data Analysis Platform

KEYWORDS: Software engineering - Meta model - Software visualisation

FUNCTIONAL DESCRIPTION: Moose is an extensive platform for software and data analysis. It offers multiple services ranging from importing and parsing data, to modeling, to measuring, querying, mining, and to building interactive and visual analysis tools. The development of Moose has been evaluated to 200 man/year.

Mots-cles : MetaModeling, Program Visualization, Software metrics, Code Duplication, Software analyses, Parsers

- Participants: Anne Etien, Nicolas Anquetil, Olivier Auverlot, Stéphane Ducasse, Julien Delplanque, Guillaume Larcheveque, Cyril Ferlicot-Delbecque and Pavel Krivanek
- Partners: Université de Berne - Sensus - Synectique - Pleiad - USI - Vrije Universiteit Brussel
- Contact: Stéphane Ducasse
- URL: http://www.moosetechnology.org

6.2. Pharo

KEYWORDS: Live programmation objet - Reflective system - Web Application

FUNCTIONAL DESCRIPTION: Pharo is a pure object reflective and dynamic language inspired by Smalltalk. In addition, Pharo comes with a full advanced programming environment developed under the MIT License. It provides a platform for innovative development both in industry and research. By providing a stable and small core system, excellent developer tools, and maintained releases, Pharo's goal is to be a platform to build and deploy mission critical applications, while at the same time continue to evolve. Pharo 60 got 100 contributors world-wide. It is used by around 30 universities, 15 research groups and around 40 companies.

- Participants: Christophe Demarey, Clement Bera, Damien Pollet, Esteban Lorenzano, Marcus Denker, Stéphane Ducasse and Guillermo Polito
- Partners: BetaNine - Reveal - Inceptive - Netstyle - Feenk - ObjectProfile - GemTalk Systems - Greyc Université de Caen - Basse-Normandie - Université de Berne - Yesplan - RMod - Pleiad - Sensus - Université de Bretagne Occidentale - École des Mines de Douai - ENSTA - Uqbar foundation Argentina - LAM Research - ZWEIDENKER - LifeWare - JPMorgan Chase - KnowRoaming - ENIT - Spesenfuchs - FINWorks - Esug - FAST - Ingenieubüro Schmidt - Projector Software - HRWorks - Inspired.org - Palantir Solutions - High Octane - Soops - Osoco - Ta Mère SCRL - University of Yaounde 1 - Software Quality Laboratory, University of Novi Sad - Software Institute Università della Svizzera italiana - Universdad Nacional de Quilmes - UMMISCO IRD - Université technique de Prague
- Contact: Marcus Denker
- URL: http://www.pharo.org

6.3. Pillar

KEYWORDS: HTML - LaTeX - HTML5

FUNCTIONAL DESCRIPTION: Pillar is a markup syntax and associated tools to write and generate documentation and books. Pillar is currently used to write several books and other documentation. It is used in the tools developed by Feenk.com.

- Partner: Feenk
- Contact: Stéphane Ducasse
- URL: https://github.com/Pillar-markup/pillar

7. New Results

7.1. Dynamic Languages: Virtual Machines

Illicium A modular transpilation toolchain from Pharo to C. The Pharo programming language runs on the OpenSmalltalk-VM. This Virtual Machine (VM) is mainly written in Slang, a subset of the Smalltalk language dedicated to VM development. Slang is transpiled to C using the Slang-to-C transpiler. The generated C is then compiled to produce the VM executable binary code. Slang is a powerful dialect for generating C because it benefits from the tools of the Smalltalk environment, including a simulator that runs and debugs the VM. However, the Slang-to-C transpiler is often too permissive. For example, the Slang-to-C transpiler generates invalid C code from some Smalltalk concepts it does not support. This makes the Slang code hard to debug as the errors are caught very late during the development process, which is worsen by the loss of the mapping between the generated C code and Slang. The Slang-to-C transpiler is also hard to extend or adapt to modify part of the translation process. We present Illicium, a new modular transpilation toolchain based on a subset of Pharo targeting C through AST transformations. This toolchain translates the Pharo AST into a C AST to generate C code. Using ASTs as source and target artifacts enables analysis, modification and validation at different levels during the translation process. The main translator is split into smaller and replaceable translators to increase modularity. Illicium also allows the possibility to introduce new translators and to chain them together, increasing reusability. To evaluate our approach, we show with a use case how to extend the transpilation process with a translation that requires changes not considered in the original C AST. [7]

GildaVM: a Non-Blocking I/O Architecture for the Cog VM. The OpenSmalltalk virtual machine (VM) was historically designed as a single-threaded VM. All VM code including the Smalltalk interpreter, the garbage collector and the just-in-time compiler run in the same single native thread. While this VM provides concurrency through green threads, it cannot take advantage of multi-core processors. This architecture performs really well in practice until the VM accesses external resources such as e.g., FFI callouts, which block the single VM thread and prevent green threads to benefit from the processor. We present GildaVM, a multi-threaded VM architecture where one thread at a time executes the VM while allowing non-blocking I/O in parallel. The ownership of the VM is orchestrated by a Global Interpreter Lock (GIL) as in the standard implementations of Python and Ruby. However, within a single VM thread concurrency is still possible through green threads. We present a prototype implementation of this architecture running on top of the Stack flavour of the OpenSmalltalk VM. We finally evaluate several aspects of this architecture like FFI and thread-switch overhead. While current benchmarks show good results for long FFI calls, short FFI calls require more research to minimize the overhead of thread-switch. [9]

7.2. Dynamic Languages: Language Constructs for Modular Design

Magic Literals in Pharo Literals are constant values (numbers, strings, etc.) used in the source code. Magic literals are the ones used without a clear explanation of their meaning. Presence of such literals harms source code readability, decreases its modularity, and encourages code duplication. Identifying magic literals is not straightforward. A literal can be considered self-explanatory in one context and magic in another. We need a heuristic to help developers spot magic literals. We study and characterize the literals in Pharo. We implemented a heuristic to detect magic literals and integrated it as a code critic rule for System Browser and Critics Browser in Pharo 7. We run our heuristic on 112,500 Pharo methods which reported 23,292 magic literals spread across 8,986 methods. We manually validated our approach on a random subset of 100 methods and found that 62% of the reported literals in those methods are indeed magic. [3]

Towards easy program migration using language virtualization Migrating programs between language versions is a daunting task. A developer writes a program in a particular version of a language and cannot foresee future language changes. In this article, we explore a solution to gradual program migration based on virtualization at the programming language level. Our language virtualization approach adds a backwards-compatibility layer on top of a recent language version, allowing developers to load and run old programs on the more recent infrastructure. Developers are then able to migrate the program to the new language version or are able to run it as it is. Our virtualization technique is based on a dynamic module implementation and code intercession techniques. Migrated and non-migrated parts co-exist in the meantime allowing an incremental migration procedure. We validate it by migrating legacy Pharo programs, MuTalk and Fuel. [10]

7.3. Dynamic Languages: Debugging

Sindarin: A Versatile Scripting API for the Pharo Debugger Debugging is one of the most important and time consuming activities in software maintenance, yet mainstream debuggers are not well-adapted to several debugging scenarios. This has led to the research of new techniques covering specific families of complex bugs. Notably, recent research proposes to empower developers with scripting DSLs, plugin-based and moldable debuggers. However, these solutions are tailored to specific use-cases, or too costly for one-time-use scenarios. We argue that exposing a debugging scripting interface in mainstream debuggers helps in solving many challenging debugging scenarios. For this purpose, we present Sindarin, a scripting API that eases the expression and automation of different strategies developers pursue during their debugging sessions. Sindarin provides a GDB-like API, augmented with AST-bytecode-source code mappings and object-centric capabilities. To demonstrate the versatility of Sindarin, we reproduce several advanced breakpoints and non-trivial debugging mechanisms from the literature. [4]

Challenges in Debugging Bootstraps of Reflective Kernels The current explosion of embedded systems (i.e., IoT, Edge Computing) implies the need for generating tailored and customized software for these systems. Instead of using specific runtimes (e.g., MicroPython, eLua, mRuby), we advocate that bootstrapping specific language kernels is a promising higher-level approach because the process takes advantage of the generated language abstractions, easing the task for a language developer. Nevertheless, bootstrapping language kernels is still challenging because current debugging tools are not suitable for fixing the possible failures that occur during the process. We take the Pharo bootstrap process as an example to analyse the different challenges a language developer faces. We propose a taxonomy of failures appearing during bootstrap and their causes. Based on this analysis, we identify future research directions: (1) prevention measures based on the reification of implicit virtual machine contracts, and (2) hybrid debugging tools that unify the debugging of high-level code from the bootstrapped language with low-level code from the virtual machine. [6]

7.4. Software Reengineering

Decomposing God Classes at Siemens A group of developers at Siemens Digital Industry Division approached our team to help them restructure a large legacy system. Several problems were identified, including the presence of God classes (big classes with thousands of lines of code and hundred of methods). They had tried different approaches considering the dependencies between the classes, but none were satisfactory. Through interaction during the last three years with a lead software architect of the project, we designed a software visualization tool and an accompanying process that allows her to propose a decomposition of a God Class in a matter of one or two hours even without prior knowledge of the class (although actually implementing the decomposition in the source code could take a week of work). We present the process that was formalized to decompose God Classes and the tool that was designed. We give details on the system itself and some of the classes that were decomposed. The presented process and visualisations have been successfully used for the last three years on a real industrial system at Siemens. [1]

Rotten Green Tests Unit tests are a tenant of agile programming methodologies, and are widely used to improve code quality and prevent code regression. A green (passing) test is usually taken as a robust sign that the code under test is valid. However, some green tests contain assertions that are never executed. We call such tests Rotten Green Tests. Rotten Green Tests represent a case worse than a broken test: they report that the code under test is valid, but in fact do not test that validity. We describe an approach to identify rotten green tests by combining simple static and dynamic call-site analyses. Our approach takes into account test helper methods, inherited helpers, and trait compositions, and has been implemented in a tool called DrTest. DrTest reports no false negatives, yet it still reports some false positives due to conditional use or multiple test contexts. Using DrTest we conducted an empirical evaluation of 19,905 real test cases in mature projects of the Pharo ecosystem. The results of the evaluation show that the tool is effective; it detected 294 tests as rotten-green tests that contain assertions that are not executed. Some rotten tests have been "sleeping" in Pharo for at least 5 years. [2]

Migrating GWT to Angular 6 using MDE In the context of a collaboration with Berger-Levrault, a major IT company, we are working on the migration of a GWT application to Angular. We focus on the GUI aspect of this migration which, even if both are web frameworks, is made difficult because they use different programming languages (Java for one, Typescript for the other) and different organization schemas (e.g. different XML files). Moreover, the new application must mimic closely the visual aspect of the old one so that the users of the application are not disturbed. We propose an approach in three steps that uses a meta-model to represent the GUI at a high abstraction level. We evaluated this approach on an application comprising 470 Java (GWT) classes representing 56 screens. We are able to model all the web pages of the application and 93% of the wid-gets they contain, and we successfully migrated (i.e., the result is visually equal to the original) 26 out of 39 pages (66%). We give examples of the migrated pages, both successful and not. [14] [11] [12]

Empirical Study of Programming to an Interface A popular recommendation to programmers in object-oriented software is to *program to an interface, not an implementation* (PTI). Expected benefits include increased simplicity from abstraction, decreased dependency on implementations, and higher flexibility. Yet, interfaces must be immutable, excessive class hierarchies can be a form of complexity, and *speculative generality* is a known code smell. To advance the empirical knowledge of PTI, we conducted an empirical investigation that involves 126 Java projects on GitHub, aiming to measuring the decreased dependency benefits (in terms of cochange). [13]

Exposing Test Analysis Results with DrTests Tests are getting the cornerstone of continuous development process and software evolution. Tests are the new gold. To improve test quality, a plethora of analyses is proposed such as test smells, mutation testing, test coverage. The problem is that each analysis often needs a particular way to expose its results to the developer. There is a need for an architecture supporting test running and analysis in a modular and extensible way. We present an extensible plugin-based architecture to run and report test results. DrTests is a new test browser that implements such plugin-based architecture. DrTests supports the execution of rotten tests, comments to tests, coverage and profiling tests. [5]

7.5. Blockchain Software Engineering

SmartAnvil: Open-Source Tool Suite for Smart Contract Analysis Smart contracts are new computational units with special properties: they act as classes with aspectual concerns; their memory structure is more complex than mere objects; they are obscure in the sense that once deployed it is difficult to access their internal state; they reside in an append-only chain. There is a need to support the building of new generation tools to help developers. Such support should tackle several important aspects: (1) the static structure of the contract, (2) the object nature of published contracts, and (3) the overall data chain composed of blocks and transactions. We present SmartAnvil an open platform to build software analysis tools around smart contracts. We illustrate the general components and we focus on three important aspects: support for static analysis of Solidity smart contracts, deployed smart contract binary analysis through inspection, and blockchain navigation and querying. SmartAnvil is open-source and supports a bridge to the Moose data and software analysis platform. [18]

The Influence Factors on Ethereum Transaction Fees In Ethereum blockchain, the user needs to set a Gas price to get a transaction processed and approved by Miners. To have the transaction executed, the Gas price has to be greater than or equal to the lowest Ethereum transaction fees. We present a set of data sampled every 15 seconds, from December 1st, 2018 to December 15, 2018, coming from different blockchain web APIs. The aim is to investigate whether and to what extent different variables - such as the number of pending transactions, the value of the USD/Ether pair, average electricity prices around the world, and the number of miners - influence the Ethereum transaction fees. This study is relevant from an economic perspective because more and more companies in different economic fields are adopting Ethereum blockchain. From historical data analysis, we found that only some of these variables do have an influence. For example, the number of pending transactions and the number of miners have a major influence on Ethereum transaction fees when compared to the other variables. [8]

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Pharo Consortium

Participants: Esteban Lorenzano, Marcus Denker, Stéphane Ducasse
From 2012, ongoing.

The Pharo Consortium was founded in 2012 and is growing constantly. By the end 2019, it has 25 company members, 19 academic partners. Inria supports the consortium with one full time engineer starting in 2011. In 2018, the Pharo Consortium joined InriaSoft.

More at http://consortium.pharo.org.

8.2. Bilateral Grants with Industry

8.2.1. Berger-Levrault: Remodularization of Architecture

Participants: Nicolas Anquetil, Santiago Bragagnolo, Stéphane Ducasse, Anne Etien, Benoît Verhaeghe
From 2017, ongoing.

We started a new collaboration with the software editor Berger-Levrault about software architecture remodularization. The collaboration started with an end study project exploring the architecture used in the company in order to later migrate from GWT to Angular since GWT will not be backward supported anymore in the next versions. A PhD CIFRE thesis started in 2019: Benoît Verhaeghe, *Support à l'automatisation de la migration d'interface d'applications Web : le cas de GWT vers Angular*.

8.2.2. Arolla: Machine Learning-Based Recommenders to Support Software Evolution

Participants: Nicolas Anquetil, Stéphane Ducasse, Anne Etien, Oleksandr Zaitsev

We started a new collaboration with the council company, Arolla, about software evolution. Arolla has daily problems with identifying architecture, design, and deviations from those artifacts. The goal of the thesis is to experiment which learning techniques can help with semi-automatically extracting design and architectural aspects and their violations. A PhD CIFRE has started in 2019: Oleksandr Zaitsev, *Machine Learning-Based Tools to Support Software Evolution*.

A second CIFRE around legacy system cartography will start in 2020.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. CAR IMT Douai

Participants: Pablo Tesone, Guillermo Polito, Marcus Denker, Stéphane Ducasse with: L. Fabresse and N. Bouraqadi (IMT Douai)
From 2009, ongoing.

We have signed a convention with the CAR team led by Noury Bouraqadi of IMT Douai. In this context we co-supervised three PhD students (Mariano Martinez-Peck, Nick Papoylias and Guillermo Polito). The team is also an important contributor and supporting organization of the Pharo project.

Pablo Tesone did a PhD co-supervided by RMOD and Pr. L. Fabresse and N. Bouraqadi (finished in 2018). Currently, three PhD Students are co-supervised:

- PhD in progress: Théo Rogliano, *On multiple language kernel*, started Oct 2019, Stéphane Ducasse, Luc Fabresse
- PhD in progress: Pierre Misse-Chanabier, *Modular, green, versatile Virtual Machines*, started Oct 2019, Stéphane Ducasse, Noury Bouraqadi
- PhD in progress: Carolina Hernández, *Tools for MicroKernels* Guillermo Polito and Luc Fabresse

We are collaborating in the Context of CPER Data since 2018.

9.1.2. CPER DATA

Participants: Marcus Denker, Stéphane Ducasse, Allex Oliveira with: L. Fabresse and N. Bouraqadi (IMT Douai)
From 2018, ongoing.

Funding to work one year on the PharoThings Platform. We are creating content for a website and a Demo in callaboration with IMT Douai.

9.2. National Initiatives

9.2.1. CEA List

Participants: Jason Lecerf, Stéphane Ducasse with Thierry Goubier (CEA List)
From 2016, PhD finished 2019.

Jason Lecerf started a shared PhD Oct 2016 and finished November 2019: *Designing Language-Agnostic Code Transformation Engines*.

9.3. European Initiatives

9.3.1. Collaborations in European Programs, Except FP7 & H2020

University of Novi Sad, Serbia
Participants: Stéphane Ducasse, Anne Etien, Nicolas Anquetil, Vincent Aranega

A collaboration with the University of Novi Sad, Serbia, started in 2018 with the university joining the Pharo Consortium as an academic member.

We have handed in a bilateral project (Campus France) between Novi Sad and RMOD: *An innovative visual environment in service of developer experience*. We expect results by the end of this year.

A Master thesis has been cosupervised. Nina Medic: *Graph library with layout algorithms in Pharo*.

Visitors:

- Sebastijan Kaplar [University of Novi Sad, Serbia, Aug 2019]
- Gordana Rakic [University of Novi Sad, Serbia, from Nov 2019]
- Nina Medic [University of Novi Sad, Serbia, from Jun 2019 until Jul 2019]

University of Prague, Czech Republic
Participants: Stéphane Ducasse.

From 2015, ongoing.

We are working with Dr. Robert Pergl from the University of Prague. Stéphane Ducasse gave a lecture at the University of Prague in 2018, the next lecture is planned for 2020.

University of Cagliari, Italy
Participants: Stéphane Ducasse

We are working on sofware engineering problems in the context of blockchain based sofware.

Visitor: Giuseppe Antonio Pierro [University of Cagliari, until July 2019].

University of Bern, Switzerland
Participants: Stéphane Ducasse, Marcus Denker

We are working on dynamic software update to, for example, automatically transform users of deprecated code.

Visitor: Manuel Leuenberger [University of Bern, from Sep 2019 until Nov 2019]

Siemens AG, Germany
Participants: Stéphane Ducasse, Anne Etien, Nicolas Anquetil

The Siemens Digital Industry Division approached our team to help them restructure a large legacy systems. The joined work resulted in a publication in 2019: *Decomposing God Classes at Siemens* [1].

9.4. International Initiatives

9.4.1. Inria International Labs

Discussions with Inria Chile have started about organizing Pharo lectures in Chile. A first visit to Inria Chile in fall 2019 did not happen due to the political situation in Chile.

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

VUB Brussels, Belgium
Participants: Guillermo Polito, Stéphane Ducasse, Marcus Denker.

Collaboration with SOFT started 2016, from 2020 Inria Lille North-European associated team funding with SOFT/VUB for 2 years.

Student: Matteo Marra, collaboration with Eliza Gonzalez Boix.

Marcus Denker gave a lecture at VUB in October 2019.

9.4.3. Inria International Partners

9.4.3.1. Informal International Partners

Uqbar Argentina

Participants: Pablo Tesone, Esteban Lorenzano, Guillermo Polito, Stéphane Ducasse.

From 2015, ongoing.

We are working with the Uqbar team from different Argentinian universities. We hired three of the people: Nicolas Passerini(engineer), Esteban Lorenzano (engineer) and Pablo Tesone (PhD).

Pharo in Research:

Participants: Pablo Tesone, Esteban Lorenzano, Guillermo Polito, Marcus Denker, Stéphane Ducasse.
From 2009, ongoing.

We are building an ecosystem around Pharo with international research groups, universities and companies. Several research groups (such as Software Composition Group – Bern, and Pleaid – Santiago) are using Pharo. Many universities are teaching OOP using Pharo and its books. Several companies worldwide are deploying business solutions using Pharo.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Abdelhakim Bouremel [University of Skikda, Algeria, Oct 2019]
- Mohamad Chakroun [Mar 2019]
- Victor Martín Dias [University of Chile, Chile, until Sep 2019]
- Christopher Fuhrman [École de technologie supérieure de Montréal, Canada, until Sep 2019]
- Yann-Gaël Guéhéneuc [Concordia University, Canada, from Apr 2019 until May 2019]
- Sebastijan Kaplar [University of Novi Sad, Serbia, Aug 2019]
- Manuel Leuenberger [University of Bern, Switzerland, from Sep 2019 until Nov 2019]
- Milton Mamani Torres [Object Profile SpA, Chile, Aug 2019]
- Nina Medic [University of Novi Sad, Serbia, from Jun 2019 until Jul 2019]
- Hayatou Oumarou [University of Maroua, Cameroun, from Sep 2019 until Oct 2019]
- Giuseppe Antonio Pierro [University of Cagliari, Italy, until July 2019]
- Gordana Rakic [University of Novi Sad, Serbia, from Nov 2019]
- Moussa Saker [University Badji Mokhtar-Annaba, Algeria, from Dec 2019]

9.5.1.1. Internships

- Dayne Lorena Guerra Calle [Inria, from Feb 2019 until Aug 2019]
- Chia Yu Li [Inria, until Jul 2019]
- Iona Thomas [Centrale Lille, from Jul 2019 until Aug 2019]
- Oleksandr Zaitsev [Inria, until Feb 2019]

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

10.1.1.1. Member of the Organizing Committees

- Marcus Denker and Stéphane Ducasse are in the board of ESUG and organized ESUG 2019, the yearly Smalltalk conference that brings together research and industry http://www.esug.org/

10.1.2. Scientific Events: Selection

10.1.2.1. Chair of Conference Program Committees

- Anne Etien has been PC chair of IWST since 2015.
- Anne Etien has been PC chair of Sattose 2019.
- Santiago Bragagonolo: IWBOSE 2020 CFP Manager, Chair (workshop on blockchain software engineering).
- Stéphane Ducasse: PC-chair International Conference on Software Reuse 2020.

10.1.2.2. Member of the Conference Program Committees

- Marcus Denker has been a member of ICOOLPS 2019.
- Steven Costiou has been member of IWST 2019.
- Anne Etien has been member of BENEVOL 2019, IWor 2019, VISSOFT 2019, ICPC 2019, CIEL 2019, GDRGPL 2019, Modelwards Program Committees.
- Guillermo Polito has been PC of IWST 2019

- Guillermo Polito has been PC of Meta Workshop 2019.
- Vincent Aranega was in the PC of IWST 2019.

10.1.2.3. Reviewing Activities

- Vincent Aranega reviewed for MODELSWARD19.
- Damien Pollet reviewed for ICSOFT.

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

- Anne Etien has been a member of the editorial board for the Special issue on Dynamic Languages of *Science of Computer Programming* journal.
- Anne Etien has been editor for *CEUR* journal vol 2510.

10.1.3.2. Reviewing Activities

- Anne Etien was reviewer for *Journal of Systems and Software*.
- Steven Costiou was reviewer for *Science of Computer Programming*.
- Guillermo Polito was reviewer for *Journal IEEE Transactions on Industrial Informatics*.
- Damien Pollet was reviewer for *Science of Computer Programming*.
- Gordana Rakic was reviewer for the 3rd call for *The Programming Journal* (and Programming 2020 conference).
- Gordana Rakic was reviewer for the *Computing and Informatics* journal.

10.1.4. Invited Talks

Steven Costiou: Keynote ICCR2019 6th International Conference on Cloud and Robotics (Remote and live debugging of IOT applications: tools for researchers and developers.)

10.1.5. Scientific Expertise

- Anne Etien: expert for the French government on Research Tax Credit.
- Anne Etien: expert for the ANRT to evaluate a CIFRE file.
- Steven Costiou: program committee of the GDR-GPL 2020 challenges.

10.1.6. Research Administration

- Steven Costiou: representative of the post-doc researchers of the CRIStAL lab for the 2019 HCERES evaluation.
- Anne Etien is animator of the Software Engineering thematic Group of the CRIStAL lab since 2019.
- Anne Etien is elected at the committee of Inria Lille - Nord Europe center since 2016.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master: Marcus Denker, 2 hours, Advanced Reflection. MetaLinks, VUB Brussels, Belgium.

Master: Steven Costiou, Fondamentaux du debugging, 20h, M2, Université de Lille

Master: Steven Costiou, Fondamentaux du debugging, 12h eTD, M2, Université de Bretagne Occidentale

Master: Steven Costiou, Introduction au C, 10h, M1 + M2, Polytech-Lille

Master: Steven Costiou, IOT, 13h, M2, Polytech-Lille

Master: Steven Costiou, IOT, 8h, M2, Université de Picardie Jules Verne

Licence: Steven Costiou, Systèmes d'information, 36h, L3, Université de Bretagne Occidentale

Master: Christophe Demarey, Intégration continue, 16 EdTD, M2, Université de Lille, France

Master: Bragagnolo Santiago, Robotique avec ROS, 36h TD, M2, ISEN

Master: Bragagnolo Santiago, Introduction a la blockchain, 6h TD, M2, Université de Lille

Licence: Bragagnolo Santiago, Programmation par objet, 26.3h TD, Polytech-Lille, France

Licence: Anne Etien, Bases de données, 30h, L3, Polytech-Lille, France

Licence: Anne Etien, Programmation par objet, 40h, L3, Polytech-Lille, France

Licence: Anne Etien, Bases de données, 30h, M1, Polytech-Lille, France

Master: Anne Etien, Test et Maintenance, 10h, M1, Polytech-Lille, France

Master: Anne Etien, Test et Maintenance, 14h, M2, Polytech-Lille, France

Master: Anne Etien, Système d'information objet, 10h, M1, Polytech Lille, France

Master: Anne Etien, Bases de données Avancées, 17h, M1, Polytech-Lille, France

Licence: Vincent Aranega, Programmation Python, 42h, niveau (L1), Université de Lille, France

Master: Vincent Aranega, Génie Logiciel, 42h, niveau (M1), Université de Lille, France

Master: Vincent Aranega, Paradigme de Programmation par la Pratique, 40h, niveau (M1), Université de Lille, France

Master: Vincent Aranega, Paradigme de Programmation par la Pratique, 15h, niveau (M1), Université de Lille, France

Licence: Vincent Aranega, Programmation C, 42h, niveau (L2), Université de Lille, France

Licence: Vincent Aranega, Programmation avancée, 32h, niveau (L3), Polytech-Lille, France

Licence: Nicolas Anquetil, Principes des Système d'Exploitation, 30h, L2, IUT-A, Université de Lille, France

Licence: Nicolas Anquetil, Conception OO Avancée, 34h, L2, IUT-A, Université de Lille, France

Licence: Nicolas Anquetil, Modélisation Mathématique, 36h, L2, UT-A, Université de Lille, France

Licence: Nicolas Anquetil, Production d'Application, 30h, L2, IUT-A, Université de Lille, France

Licence: Nicolas Anquetil, Programation Mobile, 30h, L2, IUT-A, Université de Lille, France

Licence: Nicolas Anquetil, Projets Agiles, 12h, L2, IUT-A, Université de Lille, France

Licence: Nicolas Anquetil, Graphes & Languages, 64h, L2, IUT-A, Université de Lille, France

Licence: Nicolas Anquetil, Suivi de Stages, 15h, L2, IUT-A, Université de Lille, France

Licence: Nicolas Anquetil, Interfaces Hommes-Machines, 32h, L2, IUT-A, Université de Lille, France

Licence: Benoît Verhaeghe, Structure de données, 22 EdTD, L3, Polytech Lille, France

Licence: Benoît Verhaeghe, Système d'exploitation, 20 EdTD, L2, Université de Montpellier, France

Licence: Benoît Verhaeghe, Fondamentaux Architecture et Systèmes d'Exploitation, 15 EdTD, L3, Polytech Montpellier, France

Master: Benoît Verhaeghe, Architecture Logicielles, 25 EdTD, M1, Polytech Lille, France

Master: Benoît Verhaeghe, Évolution et restructuration, 9 EdTD, M2, Université de Montpellier, France

License: Damien Pollet, OpenDevs, 35h, L3, Université de Lille, France

License: Damien Pollet, Programmation objet en Java, 140h, L3, IMT Lille-Douai, France

License: Damien Pollet, Systèmes numériques, 18h, L3, IMT Lille-Douai, France

Master: Damien Pollet, Technologies des systèmes d'informations, 28h, M1, IMT Lille-Douai, France

Master: Damien Pollet, Ingénierie du logiciel, 7h, M2, IMT Lille-Douai, France

Master: Damien Pollet, Algorithmes pour les réseaux, 22h, M2, IMT Lille-Douai, France

Licence: Julien Delplanque, Structure de données, 18 EdTD, L3, Polytech Lille, France

Licence: Julien Delplanque, Base de données relationnelles, 16 EdTD, L3, Polytech Lille, France

Master: Julien Delplanque, Base de données, 12 EdTD, M1, Polytech Lille, France

Master: Julien Delplanque, Informatique Industrielle 2, 12 EdTD, M1, Polytech Lille, France

Master: Stéphane Ducasse. Pharo, 20h, University of Prague, Czech Republic

Master: Stéphane Ducasse, Advanced Design, 16h, M1, Université Brest, France

Licence: Stéphane Ducasse, Interpreters, 24h, L3, Université de Lille, France

Master: Stéphane Ducasse, Pharo, 2h, M1, Polytech Lille, France

Master: Stéphane Ducasse, Advanced Design, 2h, M1, ENS Paris Sud, France

Master: Stéphane Ducasse, Advanced Design, 20h, M1, Université de Picardie, France

Master: Stéphane Ducasse, Advanced Design, 20h, M2, Tunis ENSI, Tunisia

Master: Stéphane Ducasse, Advanced Design, 9h, M1, Université de Brest, France

Master: Stéphane Ducasse, Modelisation, 18h, M1, Paris Pantheon Sorbonne, France

Licence: Carolina Hernández Phillips, Algorithmique et Programmation, 12 EdTD, L3, IMT Lille Douai, France

E-learning

Pharo Mooc, 7 weeks, Licences and Master students

Pedagogical resources: *TinyBlog: Develop your First Web App with Pharo* [16] and *Pharo with Style* [17]

10.2.2. Supervision

PhD: Jason Lecerf, *Designing Language-Agnostic Code Transformation Engines*, 26 Nov 2019, CEA, Thierry Goubier, Stéphane Ducasse

PhD in progress: Oleksandr Zaitsev, *Machine Learning-Based Tools to Support Software Evolution*, started Jul 2019, Stéphane Ducasse, Nicolas Anquetil

PhD in progress: Benoît Verhaeghe, *Support à l'automatisation de la migration d'interface d'applications Web : le cas de GWT vers Angular*, started Jan 2019, Anne Etien, Nicolas Anquetil

PhD in progress: Théo Rogliano, *On multiple language kernel*, started Oct 2019, Stéphane Ducasse, Luc Fabresse

PhD in progress: Pierre Misse-Chanabier, *Modular, green, versatile Virtual Machines*, started Oct 2019, Stéphane Ducasse, Noury Bouraqadi

PhD in progress: Julien Delplanque, *Software Engineering Techniques Applied to Databases*, started Oct 2017, Anne Etien, Nicolas Anquetil

PhD in progress: Thomas Dupriez, *New Generation Debugger and Application Monitoring*, started Oct 2018, Stéphane Ducasse, Steven Costiou, Guillermo Polito

PhD in progress: Mahugnon Honoré Houekpetodji, *Multi-Facet Actionable for Information System Rejuvenation*, SPI Lille, France, Stéphane Ducasse, Nicolas Anquetil, Nicolas Dias, Jérome Sudich

PhD in progress: Carolina Hernández, *Tools for MicroKernels* Guillermo Polito and Luc Fabresse

10.2.3. Juries

Dynamic program analysis for suggesting test improvements to developers, Oscar Luis Vera Perez, KTH Royal Institute of Technology, (Sweden), 17/12/2019.

Adaptation non-anticipée de comportement : application au déverminage de programmes en cours d'exécution, Steven Costiou, Université de Bretagne Occidentale, Brest (France), 26/11/2018.

Modélisation et évaluation de la sécurité des parcours d'authentification Youssou Ndaye, Université Rennes 1, France, 10/12/19

Challenges in the collaborative evolution of a proof language and its ecosystem Théo Zimmermann, Université de Paris, France, 12/12/19

ARIANE: Automated Re-documentation to Improve software Architecture uNderstanding and Evolution, Alexandre Le Borgne, IMT Mines Alès, France, 12/12/19.

Contribution à la conception d'un Système de Recommandation dédié à la réutilisation de composants logiciels, Brice Evrard Tarehy, Université de Fianarantsoa, Madagascar, Jan. 2020.

10.3. Popularization

10.3.1. Articles and contents

- Olivier Auverlot, Stéphane Ducasse, Luc Fabresse. *TinyBlog: Créer votre première application web avec Pharo* [15].
- Olivier Auverlot, Stéphane Ducasse, Luc Fabresse. *TinyBlog: Develop your First Web App with Pharo* [16].
- Stéphane Ducasse. *Pharo with Style* [17].

10.3.2. Education

- A MOOC for Pharo is online (Stéphane Ducasse).
  http://mooc.pharo.org
- The Pharo Consortium was part of GSOC (Google Summer of Code) again in 2019. Oleksandr Zaitsev was co-organizer. RMoD mentored some of the Students.

10.3.3. Interventions

- Multiple public Pharo Sprints in Lille.
  https://association.pharo.org/events
- RMOD co-organized and participated at ESUG 2019.
  http://www.esug.org/wiki/pier/Conferences/2019

10.3.4. Internal action

- 30 Minutes de Sciences, Inria Lille - Nord Europe. 25 octobre 2019: *Hard bugs and how to track them with Unanticipated Debugging*, Steven Costiou.

11. Bibliography

Publications of the year

International Conferences with Proceedings

[1] N. Anquetil, A. Etien, G. Andreo, S. Ducasse.*Decomposing God Classes at Siemens*, in "International Conference on Software Maintenance and Evolution (ICSME)", Cleveland, United States, October 2019, https://hal.inria.fr/hal-02395836

[2] J. Delplanque, S. Ducasse, G. Polito, A. P. Black, A. Etien.*Rotten Green Tests*, in "ICSE 2019 - International Conference on Software Engineering", Montréal, Canada, May 2019, https://hal.inria.fr/hal-02002346

[3] J. Delplanque, S. Ducasse, O. Zaitsev.*Magic Literals in Pharo*, in "IWST19 - International Workshop on Smalltalk Technologies Cologne", Köln, Germany, August 2019, https://hal.inria.fr/hal-02266137

[4] T. Dupriez, G. Polito, S. Costiou, V. Aranega, S. Ducasse.*Sindarin: A Versatile Scripting API for the Pharo Debugger*, in "Proceedings of the 15th ACM SIGPLAN International Symposium on Dynamic Languages", Athens, Greece, October 2019, https://arxiv.org/abs/1909.03658 [*DOI :* 10.1145/3359619.3359745], https://hal.archives-ouvertes.fr/hal-02280915

[5] D. Guerra Calle, J. Delplanque, S. Ducasse.*Exposing Test Analysis Results with DrTests*, in "International Workshop on Smalltalk Technologies", Cologne, Germany, August 2019, https://hal.inria.fr/hal-02404040

[6] C. Hernández Phillips, G. Polito, L. Fabresse, S. Ducasse, N. Bouraqadi, P. Tesone.*Challenges in Debugging Bootstraps of Reflective Kernels*, in "IWST19 - International workshop on Smalltalk Technologies", Cologne, Germany, August 2019, https://hal.archives-ouvertes.fr/hal-02297710

[7] P. Misse-Chanabier, V. Aranega, G. Polito, S. Ducasse.*Illicium A modular transpilation toolchain from Pharo to C*, in "IWST19 — International Workshop on Smalltalk Technologies", Köln, Germany, August 2019, https://hal.archives-ouvertes.fr/hal-02297860

[8] G. A. Pierro, H. S. C. Rocha.*The Influence Factors on Ethereum Transaction Fees*, in "2019 IEEE/ACM 2nd International Workshop on Emerging Trends in Software Engineering for Blockchain (WETSEB)", Montreal, Canada, IEEE, May 2019, p. 24-31 [*DOI :* 10.1109/WETSEB.2019.00010], https://hal.inria.fr/hal-02403098

[9] G. Polito, P. Tesone, E. Miranda, D. Simmons.*GildaVM: a Non-Blocking I/O Architecture for the Cog VM*, in "International Workshop on Smalltalk Technologies", Cologne, Germany, August 2019, https://hal.archives-ouvertes.fr/hal-02379275

[10] T. Rogliano, G. Polito, P. Tesone.*Towards easy program migration using language virtualization*, in "IWST19 - International Workshop on Smalltalk Technologies", Cologne, Germany, August 2019, https://hal.archives-ouvertes.fr/hal-02297756

[11] B. Verhaeghe, N. Anquetil, S. Ducasse, A. Seriai, L. Deruelle, M. Derras.*Migrating GWT to Angular 6 using MDE*, in "SATToSE 2019 - 12th Seminar on Advanced Techniques & Tools for Software Evolution", Bolzano, Italy, July 2019, https://hal.inria.fr/hal-02304301

[12] B. Verhaeghe, A. Etien, N. Anquetil, A. Seriai, L. Deruelle, S. Ducasse, M. Derras.*GUI Migration using MDE from GWT to Angular 6: An Industrial Case*, in "SANER 2019 - 26th edition of the IEEE International Conference on Software Analysis, Evolution and Reengineering", Hangzhou, China, February 2019, https://hal.archives-ouvertes.fr/hal-02019015

[13] B. Verhaeghe, C. Fuhrman, L. Guerrouj, N. Anquetil, S. Ducasse.*Empirical Study of Programming to an Interface*, in "Automated Software Engineering (ASE 2019)", San Diego, United States, November 2019, https://hal.inria.fr/hal-02353681

Conferences without Proceedings

[14] B. VERHAEGHE, A. ETIEN, S. DUCASSE, A. SERIAI, L. DERUELLE, M. DERRAS.*Migration de GWT vers Angular 6 en utilisant l'IDM*, in "CIEL 2019 - 8ème Conférence en Ingénierie du Logiciel", Toulouse, France, June 2019, https://hal.inria.fr/hal-02304296

Scientific Books (or Scientific Book chapters)

[15] O. AUVERLOT, S. DUCASSE, L. FABRESSE.*TinyBlog: Créer votre Première Application Web avec Pharo*, Square Bracket Associates, 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-02297691

[16] O. AUVERLOT, S. DUCASSE, L. FABRESSE.*TinyBlog: Develop your First Web App with Pharo*, Square Bracket Associates, April 2019, forthcoming, https://hal.archives-ouvertes.fr/hal-02297688

[17] S. DUCASSE.*Pharo with Style*, Square Bracket Associates, April 2019, https://hal.archives-ouvertes.fr/hal-02299550

[18] S. DUCASSE, H. S. C. ROCHA, S. BRAGAGNOLO, M. DENKER, C. FRANCOMME.*SmartAnvil: Open-Source Tool Suite for Smart Contract Analysis*, in "Blockchain and Web 3.0: Social, economic, and technological challenges", Routledge, February 2019, https://hal.inria.fr/hal-01940287

Research Reports

[19] M. DENKER, N. ANQUETIL, S. DUCASSE, A. ETIEN, D. POLLET.*Project-Team RMoD 2018 Activity Report*, Inria, February 2019, https://hal.inria.fr/hal-02006630

References in notes

[20] N. ANQUETIL.*A Comparison of Graphs of Concept for Reverse Engineering*, in "Proceedings of the 8th International Workshop on Program Comprehension", Washington, DC, USA, IWPC'00, IEEE Computer Society, 2000, p. 231–240, http://rmod.lille.inria.fr/archives/papers/Anqu00b-ICSM-GraphsConcepts.pdf

[21] A. BERGEL, S. DUCASSE, O. NIERSTRASZ.*Classbox/J: Controlling the Scope of Change in Java*, in "Proceedings of 20th International Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA'05)", New York, NY, USA, ACM Press, 2005, p. 177–189 [*DOI :* 10.1145/1094811.1094826], http://scg.unibe.ch/archive/papers/Berg05bclassboxjOOPSLA.pdf

[22] A. BERGEL, S. DUCASSE, O. NIERSTRASZ, R. WUYTS.*Stateful Traits*, in "Advances in Smalltalk — Proceedings of 14th International Smalltalk Conference (ISC 2006)", LNCS, Springer, August 2007, vol. 4406, p. 66–90, http://dx.doi.org/10.1007/978-3-540-71836-9_3

[23] A. BERGEL, S. DUCASSE, O. NIERSTRASZ, R. WUYTS.*Stateful Traits and their Formalization*, in "Journal of Computer Languages, Systems and Structures", 2008, vol. 34, n^{o} 2-3, p. 83–108, http://dx.doi.org/10.1016/j.cl.2007.05.003

[24] A. P. BLACK, N. SCHÄRLI, S. DUCASSE.*Applying Traits to the Smalltalk Collection Hierarchy*, in "Proceedings of 17th International Conference on Object-Oriented Programming Systems, Languages and Applications (OOPSLA'03)", October 2003, vol. 38, p. 47–64 [*DOI :* 10.1145/949305.949311], http://scg.unibe.ch/archive/papers/Blac03aTraitsHierarchy.pdf

[25] G. BRACHA, D. UNGAR.*Mirrors: design principles for meta-level facilities of object-oriented programming languages*, in "Proceedings of the International Conference on Object-Oriented Programming, Systems,

Languages, and Applications (OOPSLA'04), ACM SIGPLAN Notices", New York, NY, USA, ACM Press, 2004, p. 331–344, http://bracha.org/mirrors.pdf

[26] D. CAROMEL, J. VAYSSIÈRE.*Reflections on MOPs, Components, and Java Security*, in "ECOOP '01: Proceedings of the 15th European Conference on Object-Oriented Programming", Springer-Verlag, 2001, p. 256–274

[27] D. CAROMEL, J. VAYSSIÈRE.*A security framework for reflective Java applications*, in "Software: Practice and Experience", 2003, vol. 33, n^{o} 9, p. 821–846, http://dx.doi.org/10.1002/spe.528

[28] P. COINTE.*Metaclasses are First Class: the ObjVlisp Model*, in "Proceedings OOPSLA '87, ACM SIGPLAN Notices", December 1987, vol. 22, p. 156–167

[29] S. DENIER.*Traits Programming with AspectJ*, in "Actes de la Première Journée Francophone sur le Développement du Logiciel par Aspects (JFDLPA'04)", Paris, France, P. COINTE (editor), September 2004, p. 62–78

[30] S. DUCASSE, T. GÎRBA.*Using Smalltalk as a Reflective Executable Meta-Language*, in "International Conference on Model Driven Engineering Languages and Systems (Models/UML 2006)", Berlin, Germany, LNCS, Springer-Verlag, 2006, vol. 4199, p. 604–618 [*DOI :* 10.1007/11880240_42], http://scg.unibe.ch/archive/papers/Duca06dMOOSEMODELS2006.pdf

[31] S. DUCASSE, T. GÎRBA, M. LANZA, S. DEMEYER.*Moose: a Collaborative and Extensible Reengineering Environment*, in "Tools for Software Maintenance and Reengineering", Milano, RCOST / Software Technology Series, Franco Angeli, 2005, p. 55–71, http://scg.unibe.ch/archive/papers/Duca05aMooseBookChapter.pdf

[32] S. DUCASSE, O. NIERSTRASZ, N. SCHÄRLI, R. WUYTS, A. P. BLACK.*Traits: A Mechanism for fine-grained Reuse*, in "ACM Transactions on Programming Languages and Systems (TOPLAS)", March 2006, vol. 28, n^{o} 2, p. 331–388 [*DOI :* 10.1145/1119479.1119483], http://scg.unibe.ch/archive/papers/Duca06bTOPLASTraits.pdf

[33] S. DUCASSE, R. WUYTS, A. BERGEL, O. NIERSTRASZ.*User-Changeable Visibility: Resolving Unanticipated Name Clashes in Traits*, in "Proceedings of 22nd International Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA'07)", New York, NY, USA, ACM Press, October 2007, p. 171–190 [*DOI :* 10.1145/1297027.1297040], http://scg.unibe.ch/archive/papers/Duca07b-FreezableTrait.pdf

[34] A. DUNSMORE, M. ROPER, M. WOOD.*Object-Oriented Inspection in the Face of Delocalisation*, in "Proceedings of ICSE '00 (22nd International Conference on Software Engineering)", ACM Press, 2000, p. 467–476

[35] K. FISHER, J. REPPY.*Statically typed traits*, University of Chicago, Department of Computer Science, December 2003, n^{o} TR-2003-13

[36] P. W. L. FONG, C. ZHANG.*Capabilities as alias control: Secure cooperation in dynamically extensible systems*, Department of Computer Science, University of Regina, 2004

[37] M. FURR, J.-H. AN, J. S. FOSTER.*Profile-guided static typing for dynamic scripting languages*, in "OOPSLA'09", 2009

[38] A. GOLDBERG.*Smalltalk 80: the Interactive Programming Environment*, Addison Wesley, Reading, Mass., 1984

[39] L. GONG.*New security architectural directions for Java*, in "compcon", 1997, vol. 0, 97, http://dx.doi.org/10.1109/CMPCON.1997.584679

[40] M. HICKS, S. NETTLES.*Dynamic software updating*, in "ACM Transactions on Programming Languages and Systems", nov 2005, vol. 27, n^{o} 6, p. 1049–1096, http://dx.doi.org/10.1145/1108970.1108971

[41] G. KICZALES, J. DES RIVIÈRES, D. G. BOBROW.*The Art of the Metaobject Protocol*, MIT Press, 1991

[42] G. KICZALES, L. RODRIGUEZ.*Efficient Method Dispatch in PCL*, in "Proceedings of ACM conference on Lisp and Functional Programming", Nice, 1990, p. 99–105

[43] R. KOSCHKE.*Atomic Architectural Component Recovery for Program Understanding and Evolution*, Universität Stuttgart, 2000, http://www2.informatik.uni-stuttgart.de/cgi-bin/NCSTRL/NCSTRL_view.pl?id=DIS-2000-05&mod=0&engl=0&inst=PS

[44] S. LIANG, G. BRACHA.*Dynamic Class Loading in the Java Virtual Machine*, in "Proceedings of OOPSLA '98, ACM SIGPLAN Notices", 1998, p. 36–44

[45] L. LIQUORI, A. SPIWACK.*FeatherTrait: A Modest Extension of Featherweight Java*, in "ACM Transactions on Programming Languages and Systems (TOPLAS)", 2008, vol. 30, n^{o} 2, p. 1–32 [*DOI* : 10.1145/1330017.1330022], http://www-sop.inria.fr/members/Luigi.Liquori/PAPERS/toplas-07.pdf

[46] B. LIVSHITS, T. ZIMMERMANN.*DynaMine: finding common error patterns by mining software revision histories*, in "SIGSOFT Software Engineering Notes", September 2005, vol. 30, n^{o} 5, p. 296-305

[47] R. C. MARTIN.*Agile Software Development. Principles, Patterns, and Practices*, Prentice-Hall, 2002

[48] M. S. MILLER.*Robust Composition: Towards a Unified Approach to Access Control and Concurrency Control*, Johns Hopkins University, Baltimore, Maryland, USA, May 2006

[49] M. S. MILLER, C. MORNINGSTAR, B. FRANTZ.*Capability-based Financial Instruments*, in "FC '00: Proceedings of the 4th International Conference on Financial Cryptography", Springer-Verlag, 2001, vol. 1962, p. 349–378

[50] O. NIERSTRASZ, S. DUCASSE, N. SCHÄRLI.*Flattening Traits*, in "Journal of Object Technology", May 2006, vol. 5, n^{o} 4, p. 129–148, http://www.jot.fm/issues/issue_2006_05/article4

[51] P. J. QUITSLUND.*Java Traits — Improving Opportunities for Reuse*, OGI School of Science & Engineering, Beaverton, Oregon, USA, September 2004, n^{o} CSE-04-005

[52] J. REPPY, A. TURON.*A Foundation for Trait-based Metaprogramming*, in "International Workshop on Foundations and Developments of Object-Oriented Languages", 2006

[53] F. RIVARD.*Pour un lien d'instanciation dynamique dans les langages à classes*, in "JFLA96", Inria — collection didactique, January 1996

[54] J. H. SALTZER, M. D. SCHOROEDER.*The Protection of Information in Computer Systems*, in "Fourth ACM Symposium on Operating System Principles", IEEE, September 1975, vol. 63, p. 1278–1308

[55] N. SANGAL, E. JORDAN, V. SINHA, D. JACKSON.*Using Dependency Models to Manage Complex Software Architecture*, in "Proceedings of OOPSLA'05", 2005, p. 167–176

[56] N. SCHÄRLI, A. P. BLACK, S. DUCASSE.*Object-oriented Encapsulation for Dynamically Typed Languages*, in "Proceedings of 18th International Conference on Object-Oriented Programming Systems, Languages and Applications (OOPSLA'04)", October 2004, p. 130–149 [*DOI :* 10.1145/1028976.1028988], http://scg.unibe.ch/archive/papers/Scha04bOOEncapsulation.pdf

[57] N. SCHÄRLI, S. DUCASSE, O. NIERSTRASZ, A. P. BLACK.*Traits: Composable Units of Behavior*, in "Proceedings of European Conference on Object-Oriented Programming (ECOOP'03)", LNCS, Springer Verlag, July 2003, vol. 2743, p. 248–274 [*DOI :* 10.1007/B11832], http://scg.unibe.ch/archive/papers/Scha03aTraits.pdf

[58] C. SMITH, S. DROSSOPOULOU.*Chai: Typed Traits in Java*, in "Proceedings ECOOP 2005", 2005

[59] G. SNELTING, F. TIP.*Reengineering Class Hierarchies using Concept Analysis*, in "ACM Trans. Programming Languages and Systems", 1998

[60] K. J. SULLIVAN, W. G. GRISWOLD, Y. CAI, B. HALLEN.*The Structure and Value of Modularity in SOftware Design*, in "ESEC/FSE 2001", 2001

[61] D. VAINSENCHER.*MudPie: layers in the ball of mud*, in "Computer Languages, Systems & Structures", 2004, vol. 30, n^{o} 1-2, p. 5–19

[62] N. WILDE, R. HUITT.*Maintenance Support for Object-Oriented Programs*, in "IEEE Transactions on Software Engineering", December 1992, vol. SE-18, n^{o} 12, p. 1038–1044

Project-Team SEQUEL

Sequential Learning

IN COLLABORATION WITH: Centre de Recherche en Informatique, Signal et Automatique de Lille

IN PARTNERSHIP WITH:
Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Optimization, machine learning and statistical methods

Table of contents

Project-Team SEQUEL

Creation of the Project-Team: 2007 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
A3.3.1. - On-line analytical processing
A3.3.2. - Data mining
A3.3.3. - Big data analysis
A3.4. - Machine learning and statistics
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
A3.5.2. - Recommendation systems
A5.1. - Human-Computer Interaction
A5.10.7. - Learning
A9. - Artificial intelligence
A9.2. - Machine learning
A9.3. - Signal analysis
A9.4. - Natural language processing
A9.7. - AI algorithmics

Other Research Topics and Application Domains:

B2. - Health
B3.1. - Sustainable development
B3.5. - Agronomy
B4.4. - Energy delivery
B4.4.1. - Smart grids
B5.8. - Learning and training
B7.2.1. - Smart vehicles
B9.1.1. - E-learning, MOOC
B9.5. - Sciences
B9.5.6. - Data science

1. Team, Visitors, External Collaborators

Research Scientists

Émilie Kaufmann [CNRS, Researcher]
Odalric-Ambrym Maillard [Inria, Researcher, HDR]
Michal Valko [Inria, Researcher, until Mar 2019, HDR]
Jill-Jênn Vie [Inria, Researcher, from Oct 2019]

Faculty Member

Philippe Preux [Team leader, Université de Lille, Professor, HDR]

External Collaborators

Remi Bardenet [CNRS]
Olivier Pietquin [Google Brain, HDR]

Technical Staff

Guillaume Gautier [CNRS, Engineer]
Franck Valentini [Inria, Engineer, from May 2019]

PhD Students

Dorian Baudry [CNRS, from Nov 2019]
Lilian Besson [Ecole Normale Supérieure de Cachan, until Sep 2019]
Nicolas Carrara [Université de Lill, until Aug 2019]
Omar Darwiche Domingues [Inria]
Johan Ferret [Google, from Sep 2019]
Yannis Flet Berliac [Université de Lille]
Ronan Fruit [Inria, until Nov 2019]
Romain Gautron [Centre de coopération internationale en recherche agronomique, from Sep 2019]
Jean-Bastien Grill [DeepMind]
Nathan Grinsztajn [Ecole polytechnique, from Oct 2019]
Édouard Leurent [Renault]
Reda Ouhamma [Ecole polytechnique, from Sep 2019]
Pierre Perrault [Inria]
Hassan Saber [Inria]
Mathieu Seurin [Université de Lille]
Julien Seznec [Le livre scolaire]
Xuedong Shang [Université de Lille]
Florian Strub [Université de Lille, until May 2019]
Jean Tarbouriech [Facebook, from Apr 2019]
Kiewan Villatel [Criteo, until Jun 2019]

Post-Doctoral Fellows

Pierre Ménard [Inria, from Feb 2019]
Mohammad Sadegh Talebi Mazraeh Shahi [Inria]

Visiting Scientists

Rianne de Heide [CWI, The Netherlands, from Apr 2019 until Aug 2019]
Per Anders Jonsson [Pompeu Fabra University, Spain, from Aug 2019]
Chuan Zheng Lee [Stanford University, USA, from Jun 2019 until Oct 2019]
Arun Verma [IIT Bombay, India, from Jun 2019 until Nov 2019]
Kaige Yang [University College London, UK, from Oct 2019]

Administrative Assistant

Amelie Supervielle

2. Overall Objectives

2.1. Presentation

SEQUEL means “Sequential Learning”. As such, SEQUEL focuses on the task of learning in artificial systems (either hardware, or software) that gather information along time. Such systems are named *(learning) agents* (or learning machines) in the following. These data may be used to estimate some parameters of a model, which in turn, may be used for selecting actions in order to perform some long-term optimization task.

For the purpose of model building, the agent needs to represent information collected so far in some compact form and use it to process newly available data.

The acquired data may result from an observation process of an agent in interaction with its environment (the data thus represent a perception). This is the case when the agent makes decisions (in order to attain a certain objective) that impact the environment, and thus the observation process itself.

Hence, in SEQUEL, the term **sequential** refers to two aspects:

- The **sequential acquisition of data**, from which a model is learned (supervised and non supervised learning),
- the **sequential decision making task**, based on the learned model (reinforcement learning).

Examples of sequential learning problems include:

Supervised learning tasks deal with the prediction of some response given a certain set of observations of input variables and responses. New sample points keep on being observed.

Unsupervised learning tasks deal with clustering objects, these latter making a flow of objects. The (unknown) number of clusters typically evolves during time, as new objects are observed.

Reinforcement learning tasks deal with the control (a policy) of some system which has to be optimized (see [63]). We do not assume the availability of a model of the system to be controlled.

In all these cases, we mostly assume that the process can be considered stationary for at least a certain amount of time, and slowly evolving.

We wish to have any-time algorithms, that is, at any moment, a prediction may be required/an action may be selected making full use, and hopefully, the best use, of the experience already gathered by the learning agent.

The perception of the environment by the learning agent (using its sensors) is generally not the best one either to make a prediction, or to take a decision (we deal with Partially Observable Markov Decision Problem). So, the perception has to be mapped in some way to a better, and relevant, state (or input) space.

Finally, an important issue of prediction regards its evaluation: how wrong may we be when we perform a prediction? For real systems to be controlled, this issue can not be simply left unanswered.

To sum-up, in SEQUEL, the main issues regard:

- the learning of a model: we focus on models that map some input space $\mathbb{R}^P$ to $\mathbb{R}$,
- the observation to state mapping,
- the choice of the action to perform (in the case of sequential decision problem),
- the performance guarantees,
- the implementation of usable algorithms,

all that being understood in a *sequential* framework.

3. Research Program

3.1. In Short

SEQUEL is primarily grounded on two domains:

- the problem of decision under uncertainty,
- statistical analysis and statistical learning, which provide the general concepts and tools to solve this problem.

To help the reader who is unfamiliar with these questions, we briefly present key ideas below.

3.2. Decision-making Under Uncertainty

The phrase "Decision under uncertainty" refers to the problem of taking decisions when we do not have a full knowledge neither of the situation, nor of the consequences of the decisions, as well as when the consequences of decision are non deterministic.

We introduce two specific sub-domains, namely the Markov decision processes which model sequential decision problems, and bandit problems.

3.2.1. Reinforcement Learning

Sequential decision processes occupy the heart of the SEQUEL project; a detailed presentation of this problem may be found in Puterman's book [61].

A Markov Decision Process (MDP) is defined as the tuple $(\mathcal{X}, \mathcal{A}, P, r)$ where $\mathcal{X}$ is the state space, $\mathcal{A}$ is the action space, P is the probabilistic transition kernel, and $r : \mathcal{X} \times \mathcal{A} \times \mathcal{X} \to I\!R$ is the reward function. For the sake of simplicity, we assume in this introduction that the state and action spaces are finite. If the current state (at time t) is $x \in \mathcal{X}$ and the chosen action is $a \in \mathcal{A}$, then the Markov assumption means that the transition probability to a new state $x' \in \mathcal{X}$ (at time $t + 1$) only depends on (x, a). We write $p(x'|x, a)$ the corresponding transition probability. During a transition $(x, a) \to x'$, a reward $r(x, a, x')$ is incurred.

In the MDP $(\mathcal{X}, \mathcal{A}, P, r)$, each initial state x_0 and action sequence $a_0, a_1, ...$ gives rise to a sequence of states $x_1, x_2, ...$, satisfying $\mathbb{P}\left(x_{t+1} = x'|x_t = x, a_t = a\right) = p(x'|x, a)$, and rewards [0]$r_1, r_2, ...$ defined by $r_t = r(x_t, a_t, x_{t+1})$.

The history of the process up to time t is defined to be $H_t = (x_0, a_0, ..., x_{t-1}, a_{t-1}, x_t)$. A policy π is a sequence of functions $\pi_0, \pi_1, ...$, where π_t maps the space of possible histories at time t to the space of probability distributions over the space of actions $\mathcal{A}$. To follow a policy means that, in each time step, we assume that the process history up to time t is $x_0, a_0, ..., x_t$ and the probability of selecting an action a is equal to $\pi_t(x_0, a_0, ..., x_t)(a)$. A policy is called stationary (or Markovian) if π_t depends only on the last visited state. In other words, a policy $\pi = (\pi_0, \pi_1, ...)$ is called stationary if $\pi_t(x_0, a_0, ..., x_t) = \pi_0(x_t)$ holds for all $t \geq 0$. A policy is called deterministic if the probability distribution prescribed by the policy for any history is concentrated on a single action. Otherwise it is called a stochastic policy.

We move from an MD process to an MD problem by formulating the goal of the agent, that is what the sought policy π has to optimize. It is very often formulated as maximizing (or minimizing), in expectation, some functional of the sequence of future rewards. For example, an usual functional is the infinite-time horizon sum of discounted rewards. For a given (stationary) policy π, we define the value function $V^\pi(x)$ of that policy π at a state $x \in \mathcal{X}$ as the expected sum of discounted future rewards given that we state from the initial state x and follow the policy π:

$$V^\pi(x) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r_t | x_0 = x, \pi\right], \tag{3}$$

where $\mathbb{E}$ is the expectation operator and $\gamma \in (0, 1)$ is the discount factor. This value function V^π gives an evaluation of the performance of a given policy π. Other functionals of the sequence of future rewards may be considered, such as the undiscounted reward (see the stochastic shortest path problems [60]) and average reward settings. Note also that, here, we consider the problem of maximizing a reward functional, but a formulation in terms of minimizing some cost or risk functional would be equivalent.

[0]Note that for simplicity, we considered the case of a deterministic reward function, but in many applications, the reward r_t itself is a random variable.

In order to maximize a given functional in a sequential framework, one usually applies Dynamic Programming (DP) [58], which introduces the optimal value function $V^*(x)$, defined as the optimal expected sum of rewards when the agent starts from a state x. We have $V^*(x) = \sup_\pi V^\pi(x)$. Now, let us give two definitions about policies:

- We say that a policy π is optimal, if it attains the optimal values $V^*(x)$ for any state $x \in \mathcal{X}$, *i.e.*, if $V^\pi(x) = V^*(x)$ for all $x \in \mathcal{X}$. Under mild conditions, deterministic stationary optimal policies exist [59]. Such an optimal policy is written π^*.
- We say that a (deterministic stationary) policy π is greedy with respect to (w.r.t.) some function V (defined on $\mathcal{X}$) if, for all $x \in \mathcal{X}$,

  $$\pi(x) \in \arg\max_{a \in \mathcal{A}} \sum_{x' \in \mathcal{X}} p(x'|x,a) \left[r(x,a,x') + \gamma V(x') \right].$$

  where $\arg\max_{a \in \mathcal{A}} f(a)$ is the set of $a \in \mathcal{A}$ that maximizes $f(a)$. For any function V, such a greedy policy always exists because $\mathcal{A}$ is finite.

The goal of Reinforcement Learning (RL), as well as that of dynamic programming, is to design an optimal policy (or a good approximation of it).

The well-known Dynamic Programming equation (also called the Bellman equation) provides a relation between the optimal value function at a state x and the optimal value function at the successors states x' when choosing an optimal action: for all $x \in \mathcal{X}$,

$$V^*(x) = \max_{a \in \mathcal{A}} \sum_{x' \in \mathcal{X}} p(x'|x,a) \left[r(x,a,x') + \gamma V^*(x') \right]. \tag{4}$$

The benefit of introducing this concept of optimal value function relies on the property that, from the optimal value function V^*, it is easy to derive an optimal behavior by choosing the actions according to a policy greedy w.r.t. V^*. Indeed, we have the property that a policy greedy w.r.t. the optimal value function is an optimal policy:

$$\pi^*(x) \in \arg\max_{a \in \mathcal{A}} \sum_{x' \in \mathcal{X}} p(x'|x,a) \left[r(x,a,x') + \gamma V^*(x') \right]. \tag{5}$$

In short, we would like to mention that most of the reinforcement learning methods developed so far are built on one (or both) of the two following approaches ([64]):

- Bellman's dynamic programming approach, based on the introduction of the value function. It consists in learning a "good" approximation of the optimal value function, and then using it to derive a greedy policy w.r.t. this approximation. The hope (well justified in several cases) is that the performance V^π of the policy π greedy w.r.t. an approximation V of V^* will be close to optimality. This approximation issue of the optimal value function is one of the major challenges inherent to the reinforcement learning problem. **Approximate dynamic programming** addresses the problem of estimating performance bounds (*e.g.* the loss in performance $||V^* - V^\pi||$ resulting from using a policy π-greedy w.r.t. some approximation V- instead of an optimal policy) in terms of the approximation error $||V^* - V||$ of the optimal value function V^* by V. Approximation theory and Statistical Learning theory provide us with bounds in terms of the number of sample data used to represent the functions, and the capacity and approximation power of the considered function spaces.
- Pontryagin's maximum principle approach, based on sensitivity analysis of the performance measure w.r.t. some control parameters. This approach, also called **direct policy search** in the Reinforcement

Learning community aims at directly finding a good feedback control law in a parameterized policy space without trying to approximate the value function. The method consists in estimating the so-called **policy gradient**, *i.e.* the sensitivity of the performance measure (the value function) w.r.t. some parameters of the current policy. The idea being that an optimal control problem is replaced by a parametric optimization problem in the space of parameterized policies. As such, deriving a policy gradient estimate would lead to performing a stochastic gradient method in order to search for a local optimal parametric policy.

Finally, many extensions of the Markov decision processes exist, among which the Partially Observable MDPs (POMDPs) is the case where the current state does not contain all the necessary information required to decide for sure of the best action.

3.2.2. Multi-arm Bandit Theory

Bandit problems illustrate the fundamental difficulty of decision making in the face of uncertainty: A decision maker must choose between what seems to be the best choice ("exploit"), or to test ("explore") some alternative, hoping to discover a choice that beats the current best choice.

The classical example of a bandit problem is deciding what treatment to give each patient in a clinical trial when the effectiveness of the treatments are initially unknown and the patients arrive sequentially. These bandit problems became popular with the seminal paper [62], after which they have found applications in diverse fields, such as control, economics, statistics, or learning theory.

Formally, a K-armed bandit problem ($K \geq 2$) is specified by K real-valued distributions. In each time step a decision maker can select one of the distributions to obtain a sample from it. The samples obtained are considered as rewards. The distributions are initially unknown to the decision maker, whose goal is to maximize the sum of the rewards received, or equivalently, to minimize the regret which is defined as the loss compared to the total payoff that can be achieved given full knowledge of the problem, *i.e.*, when the arm giving the highest expected reward is pulled all the time.

The name "bandit" comes from imagining a gambler playing with K slot machines. The gambler can pull the arm of any of the machines, which produces a random payoff as a result: When arm k is pulled, the random payoff is drawn from the distribution associated to k. Since the payoff distributions are initially unknown, the gambler must use exploratory actions to learn the utility of the individual arms. However, exploration has to be carefully controlled since excessive exploration may lead to unnecessary losses. Hence, to play well, the gambler must carefully balance exploration and exploitation. Auer *et al.* [57] introduced the algorithm UCB (Upper Confidence Bounds) that follows what is now called the "optimism in the face of uncertainty principle". Their algorithm works by computing upper confidence bounds for all the arms and then choosing the arm with the highest such bound. They proved that the expected regret of their algorithm increases at most at a logarithmic rate with the number of trials, and that the algorithm achieves the smallest possible regret up to some sub-logarithmic factor (for the considered family of distributions).

3.3. Statistical analysis of time series

Many of the problems of machine learning can be seen as extensions of classical problems of mathematical statistics to their (extremely) non-parametric and model-free cases. Other machine learning problems are founded on such statistical problems. Statistical problems of sequential learning are mainly those that are concerned with the analysis of time series. These problems are as follows.

3.3.1. Prediction of Sequences of Structured and Unstructured Data

Given a series of observations $x_1, ..., x_n$ it is required to give forecasts concerning the distribution of the future observations $x_{n+1}, x_{n+2}, \cdots$; in the simplest case, that of the next outcome x_{n+1}. Then x_{n+1} is revealed and the process continues. Different goals can be formulated in this setting. One can either make some assumptions on the probability measure that generates the sequence $x_1, \cdots, x_n, \cdots$, such as that the outcomes are independent and identically distributed (i.i.d.), or that the sequence is a Markov chain, that it is a stationary process, etc. More generally, one can assume that the data is generated by a probability measure that belongs to a certain set $\mathcal{C}$. In these cases the goal is to have the discrepancy between the predicted and the "true" probabilities to go to zero, if possible, with guarantees on the speed of convergence.

Alternatively, rather than making some assumptions on the data, one can change the goal: the predicted probabilities should be asymptotically as good as those given by the best reference predictor from a certain pre-defined set.

Another dimension of complexity in this problem concerns the nature of observations x_i. In the simplest case, they come from a finite space, but already basic applications often require real-valued observations. Moreover, function or even graph-valued observations often arise in practice, in particular in applications concerning Web data. In these settings estimating even simple characteristics of probability distributions of the future outcomes becomes non-trivial, and new learning algorithms for solving these problems are in order.

3.3.2. Hypothesis testing

Given a series of observations of $x_1, \cdots, x_n, \cdots$ generated by some unknown probability measure μ, the problem is to test a certain given hypothesis H_0 about μ, versus a given alternative hypothesis H_1. There are many different examples of this problem. Perhaps the simplest one is testing a simple hypothesis "μ is Bernoulli i.i.d. measure with probability of 0 equals 1/2" versus "μ is Bernoulli i.i.d. with the parameter different from 1/2". More interesting cases include the problems of model verification: for example, testing that μ is a Markov chain, versus that it is a stationary ergodic process but not a Markov chain. In the case when we have not one but several series of observations, we may wish to test the hypothesis that they are independent, or that they are generated by the same distribution. Applications of these problems to a more general class of machine learning tasks include the problem of feature selection, the problem of testing that a certain behavior (such as pulling a certain arm of a bandit, or using a certain policy) is better (in terms of achieving some goal, or collecting some rewards) than another behavior, or than a class of other behaviors.

The problem of hypothesis testing can also be studied in its general formulations: given two (abstract) hypothesis H_0 and H_1 about the unknown measure that generates the data, find out whether it is possible to test H_0 against H_1 (with confidence), and if so, how can one do it.

3.3.3. Change Point Analysis

A stochastic process is generating the data. At some point, the process distribution changes. In the "offline" situation, the statistician observes the resulting sequence of outcomes and has to estimate the point or the points at which the change(s) occurred. In online setting, the goal is to detect the change as quickly as possible.

These are the classical problems in mathematical statistics, and probably among the last remaining statistical problems not adequately addressed by machine learning methods. The reason for the latter is perhaps in that the problem is rather challenging. Thus, most methods available so far are parametric methods concerning piece-wise constant distributions, and the change in distribution is associated with the change in the mean. However, many applications, including DNA analysis, the analysis of (user) behavior data, etc., fail to comply with this kind of assumptions. Thus, our goal here is to provide completely non-parametric methods allowing for any kind of changes in the time-series distribution.

3.3.4. Clustering Time Series, Online and Offline

The problem of clustering, while being a classical problem of mathematical statistics, belongs to the realm of unsupervised learning. For time series, this problem can be formulated as follows: given several samples $x^1 = (x_1^1, \cdots, x_{n_1}^1), \cdots, x^N = (x_N^1, \cdots, x_{n_N}^N)$, we wish to group similar objects together. While this is of course not a precise formulation, it can be made precise if we assume that the samples were generated by k different distributions.

The online version of the problem allows for the number of observed time series to grow with time, in general, in an arbitrary manner.

3.3.5. Online Semi-Supervised Learning

Semi-supervised learning (SSL) is a field of machine learning that studies learning from both labeled and unlabeled examples. This learning paradigm is extremely useful for solving real-world problems, where data is often abundant but the resources to label them are limited.

Furthermore, *online* SSL is suitable for adaptive machine learning systems. In the classification case, learning is viewed as a repeated game against a potentially adversarial nature. At each step t of this game, we observe an example $\mathbf{x_t}$, and then predict its label $\widehat{y}_t$.

The challenge of the game is that we only exceptionally observe the true label y_t. In the extreme case, which we also study, only a handful of labeled examples are provided in advance and set the initial bias of the system while unlabeled examples are gathered online and update the bias continuously. Thus, if we want to adapt to changes in the environment, we have to rely on indirect forms of feedback, such as the structure of data.

3.3.6. Online Kernel and Graph-Based Methods

Large-scale kernel ridge regression is limited by the need to store a large kernel matrix. Similarly, large-scale graph-based learning is limited by storing the graph Laplacian. Furthermore, if the data come online, at some point no finite storage is sufficient and per step operations become slow.

Our challenge is to design sparsification methods that give guaranteed approximate solutions with a reduced storage requirements.

4. Application Domains

4.1. Sequential decision making under uncertainty and prediction

The spectrum of applications of our research is very wide: it ranges from the core of our research, that is sequential decision making under uncertainty, to the application of components used to solve this decision making problem.

To be more specific, we work on computational advertising and recommendation systems; these problems are considered as a sequential matching problem in which resources available in a limited amount have to be matched to meet some users' expectations. The sequential approach we advocate paves the way to better tackle the cold-start problem, and non stationary environments. More generally, these approaches are applied to the optimization of budgeted resources under uncertainty, in a time-varying environment, including constraints on computational times (typically, a decision has to be made in less than 1 ms in a recommendation system). An other field of application of our research is related to education which we consider as a sequential matching problem between a student, and educational contents.

The algorithms to solve these tasks heavily rely on tools from machine learning, statistics, and optimization. Henceforth, we also apply our work to more classical supervised learning, and prediction tasks, as well as unsupervised learning tasks. The whole range of methods is used, from decision forests, to kernel methods, to deep learning. For instance, we have recently used deep learning on images. We also have a line of work related to software development studying how machine learning can improve the quality of software being developed. More generally, we apply our research to data science.

5. Highlights of the Year

5.1. Highlights of the Year

- Organization of the 1st Reinforcement Learning Summer Scool: 2 weeks of lectures, keynotes, and practical sessions fully dedicated to bandits and reinforcement learning. We received about 300 applications from all around the world and selected 110 participants.

- Julien Seznec and Michal Valko have obtained an oral at AI&Stats (2,5% acceptance rate) [32].
- This is the ultimate SEQUEL highlight: after 12 years, following Inria's policy, SEQUEL comes to an end. We have designed a new team-project which will be named SCOOL.

5.1.1. Awards

BEST PAPERS AWARDS :

[16]

M. ASADI, M. S. TALEBI, H. BOUREL, O.-A. MAILLARD.*Model-Based Reinforcement Learning Exploiting State-Action Equivalence*, in "ACML 2019, Proceedings of Machine Learning Research", Nagoya, Japan, 2019, vol. 101, p. 204 - 219, https://hal.archives-ouvertes.fr/hal-02378887

6. New Software and Platforms

6.1. gym-backgammon

Backgammon environment

KEYWORD: Artificial intelligence

FUNCTIONAL DESCRIPTION: This software program follows the openai gym API (https://gym.openai.com/), that is the interaction loop of reinforcement learning: the game is in a certain state, the agent selects an action, this action is simulated in the game, and the next state of the game as well as the return are returned to the agent. All these notions follows backgammon rules and should be understood as pertaining to the reinforcement learning vocabulary. As far as we are aware of, gym-backgammon is the only existing software of this type, and it is available in open source. Great care has been put into the debugging and the efficiency. This software program is developped in python. The inetraction is made according to a client-server model.

- Author: Alessio Della Libera
- Contact: Philippe Preux
- URL: https://github.com/dellalibera/gym-backgammon

6.2. gym-rubik

Rubik's cube environment

KEYWORD: Artificial intelligence

FUNCTIONAL DESCRIPTION: This software program follows the openai gym API (https://gym.openai.com/), that is the interaction loop of reinforcement learning: the game is in a certain state, the agent selects an action, this action is simulated in the game, and the next state of the game as well as the return are returned to the agent. All these notions follows Rubik's cube rules and should be understood as pertaining to the reinforcement learning vocabulary. Great care has been put into the debugging and the efficiency. This software program is developped in python.

- Author: Raphaël Avalos Martinez De Escobar
- Contact: Philippe Preux
- URL: https://github.com/raphaelavalos/gym-rubikscube

6.3. highway-env

An environment for autonomous driving decision-making

KEYWORDS: Generic modeling environment - Simulation - Autonomous Cars - Artificial intelligence

FUNCTIONAL DESCRIPTION: The environment is composed of several variants, each of which corresponds to driving scenes: highway, roundabout, intersection, merge, parking, etc. The road network is described by a graph, and is then populated with simulated vehicles. Vehicle kinematics follows a simple Bicycle model, and their behavior is determined by models derived from road traffic simulation literature. The ego-vehicle has access to a description of the scene through several types of observations, and its behavior is controlled through an action space, either discrete (change of lanes, of cruising speed) or continuous (accelerator pedal, steering wheel angle). The objective function to maximize is also described by the environment and may vary depending on the task to be solved. The interface of the library is inherited from the standard defined by OpenAI Gym, consisting of four main methods: gym.make(id), env.step(action), env.reset(), and env.render().

- Author: Edouard Leurent
- Contact: Edouard Leurent

7. New Results

7.1. Decision-making Under Uncertainty

7.1.1. Reinforcement Learning

Model-Based Reinforcement Learning Exploiting State-Action Equivalence, [16]

Leveraging an equivalence property in the state-space of a Markov Decision Process (MDP) has been investigated in several studies. This paper studies equivalence structure in the reinforcement learning (RL) setup, where transition distributions are no longer assumed to be known. We present a notion of similarity between transition probabilities of various state-action pairs of an MDP, which naturally defines an equivalence structure in the state-action space. We present equivalence-aware confidence sets for the case where the learner knows the underlying structure in advance. These sets are provably smaller than their corresponding equivalence-oblivious counterparts. In the more challenging case of an unknown equivalence structure, we present an algorithm called ApproxEquivalence that seeks to find an (approximate) equivalence structure, and define confidence sets using the approximate equivalence. To illustrate the efficacy of the presented confidence sets, we present C-UCRL, as a natural modification of UCRL2 for RL in undiscounted MDPs. In the case of a known equivalence structure, we show that C-UCRL improves over UCRL2 in terms of regret by a factor of SA/C, in any communicating MDP with S states, A actions, and C classes, which corresponds to a massive improvement when C SA. To the best of our knowledge, this is the first work providing regret bounds for RL when an equivalence structure in the MDP is efficiently exploited. In the case of an unknown equivalence structure, we show through numerical experiments that C-UCRL combined with ApproxEquivalence outperforms UCRL2 in ergodic MDPs.

Practical Open-Loop Optimistic Planning, [25]

We consider the problem of online planning in a Markov Decision Process when given only access to a generative model, restricted to open-loop policies-i.e. sequences of actions-and under budget constraint. In this setting, the Open-Loop Optimistic Planning (OLOP) algorithm enjoys good theoretical guarantees but is overly conservative in practice, as we show in numerical experiments. We propose a modified version of the algorithm with tighter upper-confidence bounds, KL-OLOP, that leads to better practical performances while retaining the sample complexity bound. Finally, we propose an efficient implementation that significantly improves the time complexity of both algorithms.

Budgeted Reinforcement Learning in Continuous State Space, [20] A Budgeted Markov Decision Process (BMDP) is an extension of a Markov Decision Process to critical applications requiring safety constraints. It relies on a notion of risk implemented in the shape of a cost signal constrained to lie below an-adjustable-threshold. So far, BMDPs could only be solved in the case of finite state spaces with known dynamics. This work extends the state-of-the-art to continuous spaces environments and unknown dynamics. We show that the solution to a BMDP is a fixed point of a novel Budgeted Bellman Optimality operator. This observation allows us to introduce natural extensions of Deep Reinforcement Learning algorithms to address large-scale BMDPs. We validate our approach on two simulated applications: spoken dialogue and autonomous driving.

Regret Bounds for Learning State Representations in Reinforcement Learning, [29]

We consider the problem of online reinforcement learning when several state representations (mapping histories to a discrete state space) are available to the learning agent. At least one of these representations is assumed to induce a Markov decision process (MDP), and the performance of the agent is measured in terms of cumulative regret against the optimal policy giving the highest average reward in this MDP representation. We propose an algorithm (UCB-MS) with O($\sqrt{T}$) regret in any communicating MDP. The regret bound shows that UCB-MS automatically adapts to the Markov model and improves over the currently known best bound of order O(T 2/3).

Planning in entropy-regularized Markov decision processes and games, [24]

We propose SmoothCruiser, a new planning algorithm for estimating the value function in entropy-regularized Markov decision processes and two-player games, given a generative model of the environment. SmoothCruiser makes use of the smoothness of the Bellman operator promoted by the regularization to achieve problem-independent sample complexity of order O(1/ϵ 4) for a desired accuracy ϵ, whereas for non-regularized settings there are no known algorithms with guaranteed polynomial sample complexity in the worst case.

7.1.1.1. Deep reinforcement learning

"I'm sorry Dave, I'm afraid I can't do that" Deep Q-Learning From Forbidden Actions, [42]

The use of Reinforcement Learning (RL) is still restricted to simulation or to enhance human-operated systems through recommendations. Real-world environments (e.g. industrial robots or power grids) are generally designed with safety constraints in mind implemented in the shape of valid actions masks or contingency controllers. For example, the range of motion and the angles of the motors of a robot can be limited to physical boundaries. Violating constraints thus results in rejected actions or entering in a safe mode driven by an external controller, making RL agents incapable of learning from their mistakes. In this paper, we propose a simple modification of a state-of-the-art deep RL algorithm (DQN), enabling learning from forbidden actions. To do so, the standard Q-learning update is enhanced with an extra safety loss inspired by structured classification. We empirically show that it reduces the number of hit constraints during the learning phase and accelerates convergence to near-optimal policies compared to using standard DQN. Experiments are done on a Visual Grid World Environment and Text-World domain.

MERL: Multi-Head Reinforcement Learning, [39]

A common challenge in reinforcement learning is how to convert the agent's interactions with an environment into fast and robust learning. For instance, earlier work makes use of domain knowledge to improve existing reinforcement learning algorithms in complex tasks. While promising, previously acquired knowledge is often costly and challenging to scale up. Instead, we decide to consider problem knowledge with signals from quantities relevant to solve any task, e.g., self-performance assessment and accurate expectations. $\mathcal{V}^{ex}$ is such a quantity. It is the fraction of variance explained by the value function V and measures the discrepancy between V and the returns. Taking advantage of $\mathcal{V}^{ex}$, we propose MERL, a general framework for structuring reinforcement learning by injecting problem knowledge into policy gradient updates. As a result, the agent is not only optimized for a reward but learns using problem-focused quantities provided by MERL, applicable out-of-the-box to any task. In this paper: (a) We introduce and define MERL, the multi-head reinforcement learning framework we use throughout this work. (b) We conduct experiments across a variety of standard benchmark environments, including 9 continuous control tasks, where results show improved performance.

(c) We demonstrate that MERL also improves transfer learning on a set of challenging pixel-based tasks. (d) We ponder how MERL tackles the problem of reward sparsity and better conditions the feature space of reinforcement learning agents.

Self-Educated Language Agent With Hindsight Experience Replay For Instruction Following, [45]

Language creates a compact representation of the world and allows the description of unlimited situations and objectives through compositionality. These properties make it a natural fit to guide the training of interactive agents as it could ease recurrent challenges in Reinforcement Learning such as sample complexity, generalization, or multi-tasking. Yet, it remains an open-problem to relate language and RL in even simple instruction following scenarios. Current methods rely on expert demonstrations, auxiliary losses, or inductive biases in neural architectures. In this paper, we propose an orthogonal approach called Textual Hindsight Experience Replay (THER) that extends the Hindsight Experience Replay approach to the language setting. Whenever the agent does not fulfill its instruction, THER learns to output a new directive that matches the agent trajectory, and it relabels the episode with a positive reward. To do so, THER learns to map a state into an instruction by using past successful trajectories, which removes the need to have external expert interventions to relabel episodes as in vanilla HER. We observe that this simple idea also initiates a learning synergy between language acquisition and policy learning on instruction following tasks in the BabyAI environment.

High-Dimensional Control Using Generalized Auxiliary Tasks, [47]

A long-standing challenge in reinforcement learning is the design of function approximations and efficient learning algorithms that provide agents with fast training, robust learning, and high performance in complex environments. To this end, the use of prior knowledge, while promising, is often costly and, in essence, challenging to scale up. In contrast, we consider problem knowledge signals, that are any relevant indicator useful to solve a task, e.g., metrics of uncertainty or proactive prediction of future states. Our framework consists of predicting such complementary quantities associated with self-performance assessment and accurate expectations. Therefore, policy and value functions are no longer only optimized for a reward but are learned using environment-agnostic quantities. We propose a generally applicable framework for structuring reinforcement learning by injecting problem knowledge in policy gradient updates. In this paper: (a) We introduce MERL, our multi-head reinforcement learning framework for generalized auxiliary tasks. (b) We conduct experiments across a variety of standard benchmark environments. Our results show that MERL improves performance for on- and off-policy methods. (c) We show that MERL also improves transfer learning on a set of challenging tasks. (d) We investigate how our approach addresses the problem of reward sparsity and pushes the function approximations into a better-constrained parameter configuration.

7.1.2. Multi-armed Bandit Theory

Asymptotically Optimal Algorithms for Budgeted Multiple Play Bandits, [15]

We study a generalization of the multi-armed bandit problem with multiple plays where there is a cost associated with pulling each arm and the agent has a budget at each time that dictates how much she can expect to spend. We derive an asymptotic regret lower bound for any uniformly efficient algorithm in our setting. We then study a variant of Thompson sampling for Bernoulli rewards and a variant of KL-UCB for both single-parameter exponential families and bounded, finitely supported rewards. We show these algorithms are asymptotically optimal, both in rateand leading problem-dependent constants, including in the thick margin setting where multiple arms fall on the decision boundary.

Non-Asymptotic Pure Exploration by Solving Games, [46]

Pure exploration (aka active testing) is the fundamental task of sequentially gathering information to answer a query about a stochastic environment. Good algorithms make few mistakes and take few samples. Lower bounds (for multi-armed bandit models with arms in an exponential family) reveal that the sample complexity is determined by the solution to an optimisation problem. The existing state of the art algorithms achieve asymptotic optimality by solving a plug-in estimate of that optimisation problem at each step. We interpret the optimisation problem as an unknown game, and propose sampling rules based on iterative strategies to estimate and converge to its saddle point. We apply no-regret learners to obtain the first finite confidence guarantees that are adapted to the exponential family and which apply to any pure exploration query and bandit structure. Moreover, our algorithms only use a best response oracle instead of fully solving the optimisation problem.

Rotting bandits are not harder than stochastic ones, [32]

In bandits, arms' distributions are stationary. This is often violated in practice, where rewards change over time. In applications as recommendation systems, online advertising, and crowdsourcing, the changes may be triggered by the pulls, so that the arms' rewards change as a function of the number of pulls. In this paper, we consider the specific case of non-parametric rotting bandits, where the expected reward of an arm may decrease every time it is pulled. We introduce the filtering on expanding window average (FEWA) algorithm that at each round constructs moving averages of increasing windows to identify arms that are more likely to return high rewards when pulled once more. We prove that, without any knowledge on the decreasing behavior of the arms, FEWA achieves similar anytime problem-dependent, O(log(KT)), and problem-independent, O($sqrtKT$), regret bounds of near-optimal stochastic algorithms as UCB1 of Auer et al. (2002a). This result substantially improves the prior result of Levine et al. (2017) which needed knowledge of the horizon and decaying parameters to achieve problem-independent bound of only O(K1/3T2/3). Finally, we report simulations confirming the theoretical improvements of FEWA.

7.1.3. Black-box Optimization

General parallel optimization without a metric, [34]

Hierarchical bandits are an approach for global optimization of extremely irregular functions. This paper provides new elements regarding POO, an adaptive meta-algorithm that does not require the knowledge of local smoothness of the target function. We first highlight the fact that the subroutine algorithm used in POO should have a small regret under the assumption of local smoothness with respect to the chosen partitioning, which is unknown if it is satisfied by the standard subroutine HOO. In this work, we establish such regret guarantee for HCT, which is another hierarchical optimistic optimization algorithm that needs to know the smoothness. This confirms the validity of POO. We show that POO can be used with HCT as a subroutine with a regret upper bound that matches the one of best-known algorithms using the knowledge of smoothness up to a $\sqrt{\log n}$ factor. On top of that, we propose a general wrapper, called GPO, that can cope with algorithms that only have simple regret guarantees. Finally, we complement our findings with experiments on difficult functions.

A simple dynamic bandit algorithm for hyper-parameter tuning, [33]

Hyper-parameter tuning is a major part of modern machine learning systems. The tuning itself can be seen as a sequential resource allocation problem. As such, methods for multi-armed bandits have been already applied. In this paper, we view hyper-parameter optimization as an instance of best-arm identification in infinitely many-armed bandits. We propose D-TTTS, a new adaptive algorithm inspired by Thompson sampling, which dynamically balances between refining the estimate of the quality of hyper-parameter configurations previously explored and adding new hyper-parameter configurations to the pool of candidates. The algorithm is easy to implement and shows competitive performance compared to state-of-the-art algorithms for hyper-parameter tuning.

7.1.4. Statistics for Machine Learning

Non-asymptotic analysis of a sequential rupture detection test and its application to non-stationary bandits, [36]

We study a strategy for online change-point detection based on generalized likelihood ratios (GLR) and that can be expressed with the binary relative entropy. This test is used to detect a change in the mean of a bounded distribution, and we propose a non-asymptotic control of its false alarm probability and detection delay. We then explain how it can be useful for sequential decision making by proposing the GLR-klUCB bandit strategy, which is efficient in piece-wise stationary multi-armed bandit models.

Sequential change-point detection: Laplace concentration of scan statistics and non-asymptotic delay bounds, [27] We consider change-point detection in a fully sequential setup, when observations are received one by one and one must raise an alarm as early as possible after any change. We assume that both the change points and the distributions before and after the change are unknown. We consider the class of piecewise-constant mean processes with sub-Gaussian noise, and we target a detection strategy that is uniformly good on this class (this constrains the false alarm rate and detection delay). We introduce a novel tuning of the GLR test that takes here a simple form involving scan statistics, based on a novel sharp concentration inequality using an extension of the Laplace method for scan-statistics that holds doubly-uniformly in time. This also considerably simplifies the implementation of the test and analysis. We provide (perhaps surprisingly) the first fully non-asymptotic analysis of the detection delay of this test that matches the known existing asymptotic orders, with fully explicit numerical constants. Then, we extend this analysis to allow some changes that are not-detectable by any uniformly-good strategy (the number of observations before and after the change are too small for it to be detected by any such algorithm), and provide the first robust, finite-time analysis of the detection delay.

Learning Multiple Markov Chains via Adaptive Allocation, [35]

We study the problem of learning the transition matrices of a set of Markov chains from a single stream of observations on each chain. We assume that the Markov chains are ergodic but otherwise unknown. The learner can sample Markov chains sequentially to observe their states. The goal of the learner is to sequentially select various chains to learn transition matrices uniformly well with respect to some loss function. We introduce a notion of loss that naturally extends the squared loss for learning distributions to the case of Markov chains, and further characterize the notion of being uniformly good in all problem instances. We present a novel learning algorithm that efficiently balances exploration and exploitation intrinsic to this problem, without any prior knowledge of the chains. We provide finite-sample PAC-type guarantees on the performance of the algorithm. Further, we show that our algorithm asymptotically attains an optimal loss.

7.1.5. DPP

On two ways to use determinantal point processes for Monte Carlo integration, [40]

This paper focuses on Monte Carlo integration with determinantal point processes (DPPs) which enforce negative dependence between quadrature nodes. We survey the properties of two unbiased Monte Carlo estimators of the integral of interest: a direct one proposed by Bardenet & Hardy (2016) and a less obvious 60-year-old estimator by Ermakov & Zolotukhin (1960) that actually also relies on DPPs. We provide an efficient implementation to sample exactly a particular multidimen-sional DPP called multivariate Jacobi ensemble. This let us investigate the behavior of both estima-tors on toy problems in yet unexplored regimes.

7.2. Applications

7.2.1. Autonomous car

Practical Open-Loop Optimistic Planning, [25]

We consider the problem of online planning in a Markov Decision Process when given only access to a generative model, restricted to open-loop policies-i.e. sequences of actions-and under budget constraint. In this setting, the Open-Loop Optimistic Planning (OLOP) algorithm enjoys good theoretical guarantees but is overly conservative in practice, as we show in numerical experiments. We propose a modified version of the algorithm with tighter upper-confidence bounds, KL-OLOP, that leads to better practical performances while retaining the sample complexity bound. Finally, we propose an efficient implementation that significantly improves the time complexity of both algorithms.

Budgeted Reinforcement Learning in Continuous State Space, [20]

A Budgeted Markov Decision Process (BMDP) is an extension of a Markov Decision Process to critical applications requiring safety constraints. It relies on a notion of risk implemented in the shape of a cost signal constrained to lie below an-adjustable-threshold. So far, BMDPs could only be solved in the case of finite state spaces with known dynamics. This work extends the state-of-the-art to continuous spaces environments and unknown dynamics. We show that the solution to a BMDP is a fixed point of a novel

Budgeted Bellman Optimality operator. This observation allows us to introduce natural extensions of Deep Reinforcement Learning algorithms to address large-scale BMDPs. We validate our approach on two simulated applications: spoken dialogue and autonomous driving.

7.2.2. Cognitive radio

Decentralized Spectrum Learning for IoT Wireless Networks Collision Mitigation, [28]

This paper describes the principles and implementation results of reinforcement learning algorithms on IoT devices for radio collision mitigation in ISM unlicensed bands. Learning is here used to improve both the IoT network capability to support a larger number of objects as well as the autonomy of IoT devices. We first illustrate the efficiency of the proposed approach in a proof-of-concept based on USRP software radio platforms operating on real radio signals. It shows how collisions with other RF signals present in the ISM band are diminished for a given IoT device. Then we describe the first implementation of learning algorithms on LoRa devices operating in a real LoRaWAN network, that we named IoTligent. The proposed solution adds neither processing overhead so that it can be ran in the IoT devices, nor network overhead so that no change is required to LoRaWAN. Real life experiments have been done in a realistic LoRa network and they show that IoTligent device battery life can be extended by a factor 2 in the scenarios we faced during our experiment.

GNU Radio Implementation of MALIN: "Multi-Armed bandits Learning for Internet-of-things Networks", [37]

We implement an IoT network in the following way: one gateway, one or several intelligent (i.e., learning) objects, embedding the proposed solution, and a traffic generator that emulates radio interferences from many other objects. Intelligent objects communicate with the gateway with a wireless ALOHA-based protocol, which does not require any specific overhead for the learning. We model the network access as a discrete sequential decision making problem, and using the framework and algorithms from Multi-Armed Bandit (MAB) learning, we show that intelligent objects can improve their access to the network by using low complexity and decentralized algorithms, such as UCB1 and Thompson Sampling. This solution could be added in a straightforward and costless manner in LoRaWAN networks, just by adding this feature in some or all the devices, without any modification on the network side.

7.2.3. Other

Accurate reconstruction of EBSD datasets by a multimodal data approach using an evolutionary algorithm, [14]

A new method has been developed for the correction of the distortions and/or enhanced phase differentiation in Electron Backscatter Diffraction (EBSD) data. Using a multi-modal data approach, the method uses segmented images of the phase of interest (laths, precipitates, voids, inclusions) on images gathered by backscattered or secondary electrons of the same area as the EBSD map. The proposed approach then search for the best transformation to correct their relative distortions and recombines the data in a new EBSD file. Speckles of the features of interest are first segmented in both the EBSD and image data modes. The speckle extracted from the EBSD data is then meshed, and the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is implemented to distort the mesh until the speckles superimpose. The quality of the matching is quantified via a score that is linked to the number of overlapping pixels in the speckles. The locations of the points of the distorted mesh are compared to those of the initial positions to create pairs of matching points that are used to calculate the polynomial function that describes the distortion the best. This function is then applied to un-distort the EBSD data, and the phase information is inferred using the data of the segmented speckle. Fast and versatile, this method does not require any human annotation and can be applied to large datasets and wide areas. Besides, this method requires very few assumptions concerning the shape of the distortion function. It can be used for the single compensation of the distortions or combined with the phase differentiation. The accuracy of this method is of the order of the pixel size. Some application examples in multiphase materials with feature sizes down to 1 μm are presented, including Ti-6Al-4V Titanium alloy, Rene 65 and additive manufactured Inconel 718 Nickel-base superalloys.

Energy Management for Microgrids: a Reinforcement Learning Approach, [41] This paper presents a framework based on reinforcement learning for energy management and economic dispatch of an islanded microgrid without any forecasting module. The architecture of the algorithm is divided in two parts: a learning phase trained by a reinforcement learning (RL) algorithm on a small dataset and the testing phase based on a decision tree induced from the trained RL. An advantage of this approach is to create an autonomous agent, able to react in real-time, considering only the past. This framework was tested on real data acquired at Ecole Polytechnique in France over a long period of time, with a large diversity in the type of days considered. It showed near optimal, efficient and stable results in each situation.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

8.1.1. Lelivrescolaire.fr

- Contract with http://Lelivrescolaire.fr; PI: Michal Valko

  Title: Sequential Machine Learning for Adaptive Educational Systems

  Duration: 3 years (Mar 2018 – Feb 2021)

  Abstract: This contract comes along the CIFRE grant on the same topic. Adaptive educational content are technologies which adapt to the difficulties encountered by students. With the rise of digital content in schools, the mass of data coming from education enables but also ask for machine learning methods. Since 2010, Lelivrescolaire.fr has been developing some learning materials for teachers and students through collaborative creation process. For instance, during the school year 2015/2016, students has achieved more than 8 000 000 exercises on its homework platform Afterclasse.fr. Our approach would be based on sequential machine learning: the algorithm learns to recommend some exercises which adapt to students gradually as they answer.
  Participants: Julien Seznec, Michal Valko.

8.1.2. Renault

- Contract with Renault; PI: Philippe Preux

  Title: Control of an autonomous vehicle

  Duration: 3 years (Dec 2017 – Nov 2020)

  Abstract: This contract comes along the CIFRE grant on the same topic. This work is done in collaboration with the NON-A team-project.
  Participants: Édouard Leurent, Odalric-Ambrym Maillard, Philippe Preux.

8.1.3. Critéo

- Contract with “Criteo”; PI: Philippe Preux

  Title: Computational advertizing

  Duration: 3 years (Dec 2017 – Jun 2019)

  Abstract: This contract comes along the CIFRE grant on the same topic. The goal is to investigate reinforcmeent learning and deep learning on the problem of ad selection on the Internet.

  Note: this contract came to its end because the PhD candidate quitted Critéo, hence aborting his PhD studies.
  Participants: Philippe Preux, Kiewan Villatel.

8.1.4. Share My Space

- Contract with “Share My Space”.

  Duration: 6 months
  Participant: Philippe Preux.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. With U. INSERM 1190, CHU Lille

Participants: Odalric-Ambrym Maillard, Philippe Preux, Philippe Preux.

Title: Bandits for Health (B4H)

Type: I-SITE Lille

Coordinator: Philippe Preux

Duration: 2019–2023

Abstract: B4H is a fundamental research project on a certain type of bandit algorithms, tailored to be applied to post-surgical patient follow-up. Bandit in a non-stationary environment will be studied. This work is performed in collaboration with Pr. F. Pattou and his group.

Title: No title

Type: Informal

Coordinator: Philippe Preux

Duration: 2019–2020

Abstract: This is mostly a data analysis work in order to study whether a certain disease may be predicted based on a certain dataset collected by U. INSERM 1190. Estelle Chatelain, a BiLille engineer, is involved in this project. This work is performed in collaboration with Pr. F. Pattou and his group.

9.1.2. With Service de Radiologie et Imagerie Musculosquelettique, CHU Lille

Participants: Philippe Preux, Franck Valentini.

Title: Radiology AI Demonstrator (RAID)

Type: CPER, Région Hauts-de-France

Coordinator: Philippe Preux

Duration: 2019–2020

Abstract: The goal of the RAID project is to assess the potential of deep learning for radio analysis and patient triage. Various applications are investigated.

9.2. National Initiatives

9.2.1. ANR BOLD

Participants: Émilie Kaufmann, Michal Valko, Pierre Ménard, Xuedong Shang, Omar Darwiche Domingues.

Title: Beyond Online Learning for better Decision making

Type: National Research Agency

Coordinator: Vianney Perchet (ENS Paris-Saclay / ENSAE)

Duration: 2019–2023

Abstract: Reactive machine learning algorithms adapt to data generating processes, typically do not require large computational power and, moreover, can be translated into offline (as opposed to online) algorithms if needed. Introduced in the 30s in the context of clinical trials, online ML algorithms have been gaining a lot of theoretical interest for the last 15 years because of their applications to the optimization of recommender systems, click through rates, planning in congested networks, to name just a few. However, in practice, such algorithms are not used as much as they should, because the traditional low-level modelling assumptions they are based upon are not appropriate, as it appears.

Instead of trying to complicate and generalise arbitrarily a framework unfit for potential applications, we will tackle this problem from another perspective. We will seek a better understanding of the simple original problem and extend it in the appropriate directions. There are currently three main barriers to a broader development of online learning, that this project aim at overcoming. 1) The classical "one step, one decision, one reward" paradigm is unfit. 2) Optimality is defined with respect to worst-case generic lower bounds and mechanics behind online learning are not fully understood. 3) Algorithms were designed in a non strategic or interactive environment.

The project gathers four parnters: ENS Paris-Saclay, University of Toulouse, Inria Lille and Université Paris Descartes.

9.2.2. ANR BoB

Participant: Michal Valko.

Title: Bayesian statistics for expensive models and tall data

Type: National Research Agency

Coordinator: CNRS (Rémi Bardenet)

Duration: 2016–2020

Abstract: Bayesian methods are a popular class of statistical algorithms for updating scientific beliefs. They turn data into decisions and models, taking into account uncertainty about models and their parameters. This makes Bayesian methods popular among applied scientists such as biologists, physicists, or engineers. However, at the heart of Bayesian analysis lie 1) repeated sweeps over the full dataset considered, and 2) repeated evaluations of the model that describes the observed physical process. The current trends to large-scale data collection and complex models thus raises two main issues. Experiments, observations, and numerical simulations in many areas of science nowadays generate terabytes of data, as does the LHC in particle physics for instance. Simultaneously, knowledge creation is becoming more and more data-driven, which requires new paradigms addressing how data are captured, processed, discovered, exchanged, distributed, and analyzed. For statistical algorithms to scale up, reaching a given performance must require as few iterations and as little access to data as possible. It is not only experimental measurements that are growing at a rapid pace. Cell biologists tend to have scarce data but large-scale models of tens of nonlinear differential equations to describe complex dynamics. In such settings, evaluating the model once requires numerically solving a large system of differential equations, which may take minutes for some tens of differential equations on today's hardware. Iterative statistical processing that requires a million sequential runs of the model is thus out of the question. In this project, we tackle the fundamental cost-accuracy trade-off for Bayesian methods, in order to produce generic inference algorithms that scale favorably with the number of measurements in an experiment and the number of runs of a statistical model. We propose a collection of objectives with different risk-reward trade-offs to tackle these two goals. In particular, for experiments with large numbers of measurements, we further develop existing subsampling-based Monte Carlo methods, while developing a novel decision theory framework that includes data constraints. For expensive models, we build an ambitious programme around Monte Carlo methods that leverage determinantal processes, a rich class of probabilistic tools that lead to accurate inference with limited model evaluations. In short, using innovative techniques such as subsampling-based Monte Carlo and determinantal point processes, we propose in this project to push the boundaries of the applicability of Bayesian inference.

9.2.3. ANR Badass

Participants: Odalric-Ambrym Maillard, Émilie Kaufmann.

Title: BAnDits for non-Stationarity and Structure

Type: National Research Agency

Coordinator: Inria Lille (O. Maillard)

Duration: 2016–2020

Abstract: Motivated by the fact that a number of modern applications of sequential decision making require developing strategies that are especially robust to change in the stationarity of the signal, and in order to anticipate and impact the next generation of applications of the field, the BADASS project intends to push theory and application of MAB to the next level by incorporating non-stationary observations while retaining near optimality against the best not necessarily constant decision strategy. Since a non-stationary process typically decomposes into chunks associated with some possibly hidden variables (states), each corresponding to a stationary process, handling non-stationarity crucially requires exploiting the (possibly hidden) structure of the decision problem. For the same reason, a MAB for which arms can be arbitrary non-stationary processes is powerful enough to capture MDPs and even partially observable MDPs as special cases, and it is thus important to jointly address the issue of non-stationarity together with that of structure. In order to advance these two nested challenges from a solid theoretical standpoint, we intend to focus on the following objectives: *(i)* To broaden the range of optimal strategies for stationary MABs: current strategies are only known to be provably optimal in a limited range of scenarios for which the class of distribution (structure) is perfectly known; also, recent heuristics possibly adaptive to the class need to be further analyzed. *(ii)* To strengthen the literature on pure sequential prediction (focusing on a single arm) for non-stationary signals via the construction of adaptive confidence sets and a novel measure of complexity: traditional approaches consider a worst-case scenario and are thus overly conservative and non-adaptive to simpler signals. *(iii)* To embed the low-rank matrix completion and spectral methods in the context of reinforcement learning, and further study models of structured environments: promising heuristics in the context of e.g. contextual MABs or Predictive State Representations require stronger theoretical guarantees.

This project will result in the development of a novel generation of strategies to handle non-stationarity and structure that will be evaluated in a number of test beds and validated by a rigorous theoretical analysis. Beyond the significant advancement of the state of the art in MAB and RL theory and the mathematical value of the program, this JCJC BADASS is expected to strategically impact societal and industrial applications, ranging from personalized health-care and e-learning to computational sustainability or rain-adaptive river-bank management to cite a few.

9.2.4. Grant of Fondation Mathématique Jacques Hadamard

Participants: Michal Valko, Ronan Fruit.

Title: Theoretically grounded efficient algorithms for high-dimensional and continuous reinforcement learning

Type: PGMO-IRMO, funded by Criteo

PI: Michal Valko

Criteo contact: Marc Abeille

Duration: 2018–2020

Abstract: While learning how to behave optimally in an unknown environment, a reinforcement learning (RL) agent must trade off the exploration needed to collect new information about the dynamics and reward of the environment, and the exploitation of the experience gathered so far to gain as much reward as possible. A good measure of the agent's performance is the regret, which measures the difference between the performance of optimal policy and the actual rewards accumulated by the agent. Two common approaches to the exploration-exploitation dilemma with provably good regret guarantees are the optimism in the face of uncertainty principle and Thompson Sampling. While these approaches have been successfully applied to small environments with a finite number of states and action (tabular scenario), existing approach for large or continuous environments either rely on heuristics and come with no regret guarantees, or can be proved to achieve small regret but cannot be implemented efficiently. In this project, we propose to make a significant contribution in the understanding of large and/or continuous RL problems by developing and analyzing new algorithms that perform well both in theory and practice.

This research line can have a practical impact in all the applications requiring continuous interaction with an unknown environment. Recommendation systems belong to this category and, by definition, they can be modeled has a sequence of repeated interaction between a learning agent and a large (possibly continuous) environment.

9.2.5. With CIRAD and CGIAR

Participants: Philippe Preux, Odalric-Ambrym Maillard, Romain Gautron.

Title: Crop management

Duration: 2019–2022

Abstract: We study how reinforcement learning may be used to provide recommendations of practices to small farm holders in under-developped countries. In such countries, agriculture remains mostly a non mechanized activity, dealing with fields of very small surface.

This is a very challenging application for RL: data is scarce, recommendations made to farmers should be of quality: we can not just learn by making millions of bad recommendations to people who use them to live and feed their family. Modeling the problem as an RL is yet an other challenge.

We feel that it is very interesting to challenge RL with such complex tasks. Solving games with RL is nice and fun, but we should assess RL abilities to solve real risky tasks.

This pioneering work is done within Romain Gautron's PhD, in collaboration with CIRAD, the CGIAR, and in relation with the Africa Rising program.

9.2.6. Project CNRS-INSERM REPOS

Participants: Émilie Kaufmann, Clémence Réda [INSERM].

Title: Repositionnement de médicaments basé sur leurs effets transcriptionnels par des approches de réseaux géniques

Type: Appel à projet Santé Numérique

PI: Pr. Andrée Delahaye-Duriez (INSERM, UMR1141)

Duration: 2019

Abstract: Drug repurposing consists in studying molecules already commercialized and find other therapies in which they may be efficient. The quality of therapeutic components is often assessed by their affinity to a given protein, but it can also be assessed in terms of their impact at the transciptomic level. The aim of this project is to develop a method for selecting which drugs could be used for a given disease based on their ability to inverse the transcriptomic signature of a pathological phenotype. We will propose a new method based on algorithms for sequential decision making (bandit algorithms) to adaptively select which drug should be explored, where exploring a drug means performing simulations to propagate the perturbation (using for example gene regulatory networks) and estimate the transcriptomic impact of the perturbation induced by the drug. These simulations will hinge on existing gene expression data that are already available for many drugs, but also on new transcriptomic data generated for a mouse model of a rare disease called the Ondine syndrom.

9.2.7. National Partners

- ENS Paris-Saclay
  - M. Valko collaborated with V. Perchet on structured bandit problem. They co-supervise a PhD student (P. Perrault) together
  - O-A. Maillard collaborates with V. Perchet on automated feature learning. They co-supervise a PhD student (R. Ouhamma) together
  - E. Kaufmann collaborated with V. Perchet and E. Boursier on Multi-Player bandits
- Institut de Mathématiques de Toulouse, then Ecole Normale Supérieure de Lyon
  - E. Kaufmann collaborated with Aurélien Garivier on sequential testing and structured bandit problems
- Centrale-Supélec Rennes:
  - E. Kaufmann co-advises Lilian Besson, who works at CentraleSupélec with Christophe Moy on MAB for cognitive radio and Internet-of-Things communications
- Participation to the Inria Project Lab (IPL) “HPC – Big Data”: Started in 2018, this IPL gathers a dozen Inria team-projects, mixing researchers in HPC with researchers in machine learning and data science. SEQUEL contribution in this project is about how we can take advantage of HPC for our computational needs regarding deep learning and deep reinforcement learning, and also how such learning algorithms might be redesigned or re-implemented in order to take advantage of HPC architectures.
- Participation to the Inria Project Lab (IPL) “HYAIAI”: Started in 2019, this IPL gathers Magnet and SEQUEL in Lille, Tau in Saclay, Lacodam in Rennes, Orpailleur and Multispeech in Nancy. The goal of this IPL is to study machine learning combining symbolic and numeric approaches, to obtain interpretable AI systems.
- PCIM (École Polytechnique)
  - Ph. Preux collaborates with Tanguy Levent (PhD student) on the control of smartgrids with reinforcement learning
- Defrost (Inria Lille)
  - Ph. Preux collaborates with Pierre Schegg (PhD student) on the control of soft robots with reinforcement learning

9.3. European Initiatives

9.3.1. Collaborations in European Programs, Except FP7 & H2020

9.3.1.1. DELTA

Participants: Michal Valko, Émilie Kaufmann, Omar Darwiche Domingues, Pierre Ménard.

Program: CHIST-ERA

Project acronym: DELTA

Project title: Dynamically Evolving Long-Term Autonomy

Duration: October 2017 - December 2021

Coordinator: Anders Jonsson (PI)

Inria Coordinator: Michal Valko

Other partners: UPF Spain, MUL Austria, ULG Belgium

Abstract: Many complex autonomous systems (e.g., electrical distribution networks) repeatedly select actions with the aim of achieving a given objective. Reinforcement learning (RL) offers a powerful framework for acquiring adaptive behaviour in this setting, associating a scalar reward with each action and learning from experience which action to select to maximise long-term reward. Although RL has produced impressive results recently (e.g., achieving human-level play in Atari games and beating the human world champion in the board game Go), most existing solutions only work under strong assumptions: the environment model is stationary, the objective is fixed, and trials end once the objective is met. The aim of this project is to advance the state of the art of fundamental research in lifelong RL by developing several novel RL algorithms that relax the above assumptions. The new algorithms should be robust to environmental changes, both in terms of the observations that the system can make and the actions that the system can perform. Moreover, the algorithms should be able to operate over long periods of time while achieving different objectives. The proposed algorithms will address three key problems related to lifelong RL: planning, exploration, and task decomposition. Planning is the problem of computing an action selection strategy given a (possibly partial) model of the task at hand. Exploration is the problem of selecting actions with the aim of mapping out the environment rather than achieving a particular objective. Task decomposition is the problem of defining different objectives and assigning a separate action selection strategy to each. The algorithms will be evaluated in two realistic scenarios: active network management for electrical distribution networks, and microgrid management. A test protocol will be developed to evaluate each individual algorithm, as well as their combinations.

9.4. International Initiatives

9.4.1. Inria International Partners

- É. Kaufmann visited CWI, Amsterdam for one week in February, working with Wouter Koolen, Rémy Degenne and Rianne De Heide. Pierre Ménard also collaborated with them.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

- Anders Jonsson, Pompeu Fabra University, Spain ,sabbatical year Sep 2019 – Jul 2020
- Kaige Yang, University College London, UK, Oct 9 & Jan 9 2020
- Rianne de Heide, CWI, The Netherlands, April 23 – August 3, 2019
- Chuan-Zheng Lee, Stanford University, USA, June – October 2019
- Arun Verma, IIT Bombay, June 1 – November 30, 2019

9.5.1.1. Internships

- Alessio Della Libera, from Jul 2019 until Sep 2019

  TD-Gammon, and his github with the gym-backgammon code

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events: Organisation

- the 1st Reinforcement Learning Summer Scool, July 1-12, 2019, Villeneuve d'Ascq
- the 3rd Vigil workshop at NeurIPS 2019

10.1.1.1. Member of the Organizing Committees

- F. Strub, co-organizer of the workshop “Visually Grounded Interaction and Language (ViGIL)” at NeurIPS 2019
- The whole SEQUEL team has organized RLSS

10.1.2. Scientific Events: Selection

10.1.2.1. Member of the Conference Program Committees

- Émilie Kaufmann: ALT
- Odalric-Ambrym Maillard: ICML, ECAI, SIF
- Philippe Preux: ECML, EGC, SFC

10.1.2.2. Reviewer

In 2019, we have reviewed submissions for: AI&Stats, NeurIPS, ALT, ICML, COLT, IJCAI, AAAI, CDC, ECAI

10.1.3. Journal

10.1.3.1. Reviewer - Reviewing Activities

- Journal of Machine Learning Research
- Journal of Artificial Intelligence Research
- The Annals of Statistics
- Bernoulli
- IEEE Transactions on Knowledge and Data Engineering
- Machine Learning
- Information and Inference: A Journal of the IMA

10.1.4. Invited Talks

- E. Kaufmann
  - “Beyond Classical Bandit Tools for Monte-Carlo Tree Search”, AAAI workshop on Reinforcement Learning for Games, Honolulu, Jan 2019
  - “New tools for Adaptive Testing and Applications to Bandit Problems”, Machine Learning and Optimization Working Group, Ecole des Ponts, Feb 2019
  - “Generalized Likelihood Ratios Tests applied to Sequential Decision Making”, Statistics Seminar, Agro ParisTech, Paris, May 2019
  - “Generalized Likelihood Ratios Tests applied to Sequential Decision Making”, Machine Learning Seminar, University of Leiden, The Netherlands, May 2019
  - “Quelques outils statistiques pour la prise de décision séquentielle”, Conférence plénière du GRETSI, Lille, Aug 2019
  - “Practical algorithm for multi-player bandits”, MAPLE workshop, Milan, Italy Sep 2019
  - “Practical algorithm for multi-player bandits”, Invited session of the Allerton Conference, Urbana-Champaign, USA Sep 2019
- Odalric-Ambrym Maillard:
  - “La prise de décision séquentielle au service de la société de demain”, Euratechnologie, Lille, Feb 2019
  - “Change of mean detection, non-asymptotic delay and aggregation”, 3rd non-stationary day, Institut Henry Poincaré, Paris, Mar 2019
  - “A tour of time-uniform concentration inequalities: Laplace, Peeling, Kernel”, Workshop on empirical Processes and Applications to Statistics, Besançon, May 2019

  - "A tour of time-uniform concentration inequalities: Laplace, Peeling, Kernel", CWI, Amsterdam, The Netherlands, Jun 2019
  - "Reinforcement Learning: successes and promises", Ecole Polytechnique, Palaiseau, Nov 2019
- Philippe Preux:
  - A brief introduction to supervised learning and reinforcement learning, 1st humAIn seminar, Villeneuve d'Ascq, Feb 2019
  - "Sous le contrôle des bandits", AFCE, June 2019
  - Explainability in machine learning, 3rd humAIn seminar, Lille, June 2019
  - "Learning to act", ENS-Paris-Saclay, Conférence de rentrée, Sep 2019
  - "Apprentissage par renforcement : mythe et réalité", FOOR, Tourcoing, Nov 2019
- Jill-Jênn Vie:
  - "IA, éducation et formation", Hermès, Paris, Oct 2019
  - "JJ Vie's Factorization IV", LaBRI, Bordeaux, Nov 2019
  - "Deep Learning for Anime & Manga", Paris Open Source Summit, Dec 2019
  - "Deep Learning for Recommender Systems", Université Cergy-Pontoise, Dec 2019
- R. Gautron, O-A. Maillard, Ph. Preux, "Reinforcement learning for crop-management: a sequential decision-making under uncertainty approach", CGIAR convention, Hydebarad, India, Oct 2019
- Ph. Preux, M. Seurin, "L'IA, les données, ... et l'Homme dans tout ça ?", congress "Les données et leurs usages dans les technologies du numérique", Douai, Oct 2019

10.1.5. Scientific Expertise

- Émilie Kaufmann:
  - member of the hiring committee for an assistant professor in probability/statistics at Université Paris-Sud
- Odalric-Ambrym Maillard:
  - member of the hiring committee for CRCN at Inria Lille
- Philippe Preux:
  - member of the hiring committee for CRCN at Inria Rennes
  - member of the hiring committee for CRCN at Inria (national)
  - evaluation of submissions to ANRT (he also declined many such invitations due to lack of time, *e.g.* with ANR)

10.1.6. Research Administration

- Odalric-Ambrym Maillard is:
  - member of the CER at Inria Lille
- Philippe Preux is:
  - "délégué scientifique adjoint" of the Inria center in Lille
  - member of the Inria evaluation committee (CE)
  - member of the Inria internal scientific committee (COSI)
  - member of the scientific committee of CRIStAL until Jan 2019
  - the head of the "Data Intelligence" thematic group at CRIStAL until Jan 2019

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Doctorat: Émilie Kaufmann and Odalric-Ambrym-Maillard, “Bandit algorithms I”, RLSS Summer School, Lille, 9h, July 2019

Master: Émilie Kaufmann, “Data Mining”, 36h, M1, Université de Lille, Jan-Apr 2019

Master: Émilie Kaufmann, “Reinforcement Learning”, 24h, M2, Ecole Centrale de Lille, Nov 2019-Jan 2020

Master: Odalric-Ambrym Maillard, “Reinforcement Learning”, 38h equivalent TD, M2, Ecole Polytechnique, Palaiseau, Jan-Mar 2019

Doctorat: Odalric-Ambrym Maillard, “Bandit algorithms II”, RLSS Summer School, Lille, 9h, July 2019

Doctorat: Philippe Preux, “Reinforcement Learning”, Fall School on AI (IA2) of the GDR IA (CNRS), Lyon, 3h, Oct 2019

Doctorat: Philippe Preux, “AI learns to act”, MOMI, Sophia-Antipolis, 1h30, Feb 2019

10.2.2. Supervision

HdR: Odalric-Ambrym Maillard, Mathematics of Sequential Decision Making, Université de Lille, Feb 11, 2019

PhD: Lilian Besson, Multi-players Bandit Algorithms for Internet of Things Networks, CentraleSupélec Rennes, Nov 20, 2019, supervisors: Christophe Moy (Université de Rennes) et Émilie Kaufmann

PhD: Ronan Fruit, Exploration–exploitation dilemma in Reinforcement Learning under various form of prior knowledge, Université de Lille, Nov 6, 2019, supervisor: Alessandro Lazaric

PhD: Nicolas Carrara, “Apprentissage par renforcement pour optimisation de systèmes de dialogue via l’adaptation à chaque utilisateur”, Université de Lille, Dec 18, 2019, supervisor: Oliier Pietquin

PhD in progress: Dorian Baudry, “Efficient Exploration for Structured Bandits and Reinforcement Learning”, since Nov 2019, supervisors: É. Kaufmann, O-A. Maillard

PhD in progress: Omar Darwiche Domingues, “Sequential Learning in Dynamic Environments”, since Oct 2018, supervisors: É. Kaufmann, M. Valko

PhD in progress: Johan Ferret, “Explainable Reinforcement Learning via Deep Neural Networks”, since Fall 2019, supervisor: Ph. Preux, O. Pietquin

PhD in progress: Yannis Flet-Berliac, “Deep reinforcement learning in stochastic and non stationary environments”, since Oct 2018, supervisor: Ph. Preux

PhD in progress: Guillaume Gautier, DPPs in ML, started Oct 2016, defense scheduled in March 2020. Supervisors: R. Bardenet, M. Valko.

PhD in progress: Jean-Bastien Grill, “Création et analyse d’algorithmes efficaces pour la prise de décision dans un environnement inconnu et incertain”, started Oct 2014, defended on Dec 19, 2019. Supervisors: R. Munos, M. Valko

PhD in progress: Nathan Grinsztajn, “Apprentissage par renforcement pour la résolution séquentielle de problèmes d’optimisation combinatoire incertains et partiellement définis”, since Fall 2019, supervisor: Ph. Preux

PhD in progress: Léonard Hussenot, “Adversarial reinforcement learning: attacks and robustness”, since Fall 2019, supervisor: Ph. Preux, O. Pietquin

PhD in progress: Édouard Leurent, “Autonomous vehicle control: application of machine learning to contextualized path planning”, since Oct 2017, supervisors: O-A. Maillard, D. Effimov (Valse), W. Perruquetti (CRIStAL)

PhD in progress: Reda Ouhamma, “Automated feature representation”, since Fall 2019, O-A. Maillard

PhD in progress: Pierre Perrault, “Online Learning on Streaming Graphs”, since Sep 2017, supervisors: M. Valko, V. Perchet

PhD in progress: Sarah Perrin, “Reinforcement Learning in Mean Field Games”, since Fall 2019, supervisors: O. Pietquin, R. Elie

PhD in progress: Hassan Saber, “Structured multi-armed bandits”, since Oct 2018, Structured Multi-armed bandits, supervisor: O-A. Maillard.

PhD in progress: Mathieu Seurin, “Multi-scale rewards in reinforcement learning”, since Oct 2017, supervisors: O. Pietquin, Ph. Preux

PhD in progress: Julien Seznec, “Sequential Learning for Educational Systems”, since Mar 2017, supervisors: M. Valko, A. Lazaric, J. Banon

PhD in progress: Xuedong Shang, “Adaptive methods for optimization in stochastic environments”, started Oct 2017, supervisors: É. Kaufmann, M. Valko

PhD in progress: Florian Strub, “Reinforcement Learning for visually grounded interaction”, since Jan 2016, defense scheduled for Jan 2020, supervisors: O. Pietquin and J. Mary

PhD in progress: Kiewan Villatel, “Deep Learning for Conversion Rate Prediction in Online Advertising”, started Oct 2017, aborted June 2019, supervisor: Ph. Preux

10.2.3. Juries

- Émilie Kaufmann:
  - Aristide Tossou, member of the jury, Chalmers University, Sweden, Nov 18, 2019
  - Rémi Degenne, member of the jury, Université Paris-Diderot, Dec 18, 2019
  - member of the Mathematics jury for the admission competition of ENS, section B/L
- Odalric-Ambrym Maillard:
  - Léonard Torossian, reviewer, Université Toulouse III, Dec 17, 2019.
- Philippe Preux:
  - Quentin Waymel (medical doctorate), member of the jury, Université de Lille, Jun 2019
  - Adrien Legrand, reviewer, Université de Picardie, Amiens, Nov 29, 2019
  - Erinc Merdivan, reviewer, Centrale-Supélec Metz, Dec 17, 2019
  - Nicolas Carrara, member of the jury, Université de Lille, Dec 18, 2019
- Michal Valko:
  - Aristide Tossou, opponent, Chalmers University, Sweden, Nov 18, 2019

10.3. Popularization

10.3.1. Articles and contents

- Philippe Preux:
  - interviewed by *Le Monde* pulished in Sep 2019
  - interview on I-SITE project B4H, Inria

10.3.2. Education

- Odalric-Ambrym Maillard:
  - “Reinforcement Learning: successes and promises”, Executive Master, Ecole Polytechnique, Palaiseau, Nov 2019

10.3.3. Interventions

- Philippe Preux:

Figure 1.

  - panel on “Promises and perils of AI”, CGIAR, Hyderabad, India, Oct 2019
  - panel on “AI and man”, Euratechnologies, Lille, Sep 2019
- Yannis Flet-Berliac and Philippe Preux: panel on “Who’s the pilot: man of software?”, FOOR, Le Fresnoy, Tourcoing, Nov 2019
- Yannis Flet-Berliac: “Princess of parallelograms” installation (with Thomas Depas), Le Fresnoy, Tourcoing, ”Damien & The Love Guru” gallery in Brussels, Belgium, Sep–Dec 2019

  Princess of parallelograms is a collaborative project between Yannis Flet-Berliac and a student from Le Fresnoy National Studio of Contemporary Arts. They created an interactive sculpture made of a variety of computer vision attributes: a support for anthropomorphic projections, a set of generated virtual masks, or a new form of photographic trap. When visitors stand in front of the device’s webcam, a Deep Convolutional Conditional-GAN Auto-Encoder model applies a filter on their face with virtual flesh, hair, and facial expressions in real-time. In the meantime, an emotion detection model trained on the FER-2013 dataset is running in the background. The system allows the users to actively interact with the installation. So far, the project has been exposed at Le Fresnoy and in Brussels.

11. Bibliography

Major publications by the team in recent years

[1] O. CAPPÉ, A. GARIVIER, O.-A. MAILLARD, R. MUNOS, G. STOLTZ.*Kullback-Leibler Upper Confidence Bounds for Optimal Sequential Allocation*, in "Annals of Statistics", 2013, vol. 41, n^{o} 3, p. 1516-1541, https://hal.archives-ouvertes.fr/hal-00738209

[2] A. CARPENTIER, M. VALKO.*Revealing Graph Bandits for Maximizing Local Influence*, in "Proceedings of the 19th International Conference on Artificial Intelligence and Statistics", Cadiz, Spain, A. GRETTON, C. C. ROBERT (editors), Proceedings of Machine Learning Research, PMLR, May 2016, vol. 51, p. 10-18, http://proceedings.mlr.press/v51/carpentier16a.html

[3] H. DE VRIES, F. STRUB, J. MARY, H. LAROCHELLE, O. PIETQUIN, A. COURVILLE.*Modulating early visual processing by language*, in "Conference on Neural Information Processing Systems", Long Beach, United States, December 2017, p. 6594-6604, https://hal.inria.fr/hal-01648683

[4] N. GATTI, A. LAZARIC, M. ROCCO, F. TROVÒ.*Truthful Learning Mechanisms for Multi–Slot Sponsored Search Auctions with Externalities*, in "Artificial Intelligence", October 2015, vol. 227, p. 93-139, https://hal.inria.fr/hal-01237670

[5] M. GHAVAMZADEH, Y. ENGEL, M. VALKO.*Bayesian Policy Gradient and Actor-Critic Algorithms*, in "Journal of Machine Learning Research", January 2016, vol. 17, n^o^ 66, p. 1-53, https://hal.inria.fr/hal-00776608

[6] H. KADRI, E. DUFLOS, P. PREUX, S. CANU, A. RAKOTOMAMONJY, J. AUDIFFREN.*Operator-valued Kernels for Learning from Functional Response Data*, in "Journal of Machine Learning Research (JMLR)", April 2016, vol. 17, n^o^ 20, p. 1-54, https://hal.archives-ouvertes.fr/hal-01221329

[7] E. KAUFMANN, O. CAPPÉ, A. GARIVIER.*On the Complexity of Best Arm Identification in Multi-Armed Bandit Models*, in "Journal of Machine Learning Research", January 2016, vol. 17, p. 1-42, https://hal.archives-ouvertes.fr/hal-01024894

[8] A. LAZARIC, M. GHAVAMZADEH, R. MUNOS.*Analysis of Classification-based Policy Iteration Algorithms*, in "Journal of Machine Learning Research", 2016, vol. 17, p. 1-30, https://hal.inria.fr/hal-01401513

[9] R. MUNOS.*From Bandits to Monte-Carlo Tree Search: The Optimistic Principle Applied to Optimization and Planning*, in "Foundations and Trends in Machine Learning", 2014, vol. 7, n^o^ 1, p. 1-129, http://dx.doi.org/10.1561/2200000038

[10] R. ORTNER, D. RYABKO, P. AUER, R. MUNOS.*Regret bounds for restless Markov bandits*, in "Journal of Theoretical Computer Science (TCS)", 2014, vol. 558, p. 62-76 [*DOI :* 10.1016/J.TCS.2014.09.026], https://hal.inria.fr/hal-01074077

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] N. CARRARA.*Reinforcement learning for Dialogue Systems optimization with user adaptation*, Ecole Doctoral Science pour l'Ingénieur Université Lille Nord-de-France, December 2019, https://tel.archives-ouvertes.fr/tel-02422691

[12] R. FRUIT.*Exploration-exploitation dilemma in Reinforcement Learning under various form of prior knowledge*, Université de Lille 1, Sciences et Technologies; CRIStAL UMR 9189, November 2019, https://tel.archives-ouvertes.fr/tel-02388395

[13] O.-A. MAILLARD.*Mathematics of Statistiscal Sequential Decision Making*, Université de Lille Nord de France, February 2019, Habilitation à diriger des recherches, https://hal.archives-ouvertes.fr/tel-02077035

Articles in International Peer-Reviewed Journal

[14] M.-A. CHARPAGNE, F. STRUB, T. M. POLLOCK.*Accurate reconstruction of EBSD datasets by a multimodal data approach using an evolutionary algorithm*, in "Materials Characterization", April 2019, vol. 150, p. 184-198, https://arxiv.org/abs/1903.02988 - A short version of this paper exists towards people working in Machine Learning, namely arxiv:1903.02982 [*DOI :* 10.1016/J.MATCHAR.2019.01.033], https://hal.archives-ouvertes.fr/hal-02062098

[15] A. R. LUEDTKE, E. KAUFMANN, A. CHAMBAZ.*Asymptotically Optimal Algorithms for Budgeted Multiple Play Bandits*, in "Machine Learning Journal", September 2019, vol. 108, n^o^ 11, p. 1919-1949, https://arxiv.org/abs/1606.09388 , https://hal.archives-ouvertes.fr/hal-01338733

International Conferences with Proceedings

[16] *Best Paper*
M. Asadi, M. S. Talebi, H. Bourel, O.-A. Maillard.*Model-Based Reinforcement Learning Exploiting State-Action Equivalence*, in "ACML 2019, Proceedings of Machine Learning Research", Nagoya, Japan, 2019, vol. 101, p. 204 - 219, https://hal.archives-ouvertes.fr/hal-02378887.

[17] P. Bartlett, V. Gabillon, J. Healey, M. Valko.*Scale-free adaptive planning for deterministic dynamics & discounted rewards*, in "International Conference on Machine Learning", Long Beach, United States, 2019, https://hal.inria.fr/hal-02387484

[18] P. Bartlett, V. Gabillon, M. Valko.*A simple parameter-free and adaptive approach to optimization under a minimal local smoothness assumption*, in "Algorithmic Learning Theory", Chicago, United States, 2019, https://hal.inria.fr/hal-01885368

[19] D. Calandriello, L. Carratino, A. Lazaric, M. Valko, L. Rosasco.*Gaussian process optimization with adaptive sketching: Scalable and no regret*, in "Conference on Learning Theory", Phoenix, United States, 2019, https://hal.inria.fr/hal-02144311

[20] N. Carrara, E. Leurent, R. Laroche, T. Urvoy, O.-A. Maillard, O. Pietquin.*Budgeted Reinforcement Learning in Continuous State Space*, in "Conference on Neural Information Processing Systems", Vancouver, Canada, Advances in Neural Information Processing Systems, December 2019, vol. 32, https://arxiv.org/abs/1903.01004 , https://hal.archives-ouvertes.fr/hal-02375727

[21] M. Dereziński, D. Calandriello, M. Valko.*Exact sampling of determinantal point processes with sublinear time preprocessing*, in "Neural Information Processing Systems", Vancouver, Canada, 2019, https://hal.inria.fr/hal-02387524

[22] C. Dimitrakakis, Y. Liu, D. Parkes, G. Radanovic.*Bayesian Fairness*, in "AAAI 2019 - Thirty-Third AAAI Conference on Artificial Intelligence", Honolulu, United States, January 2019, https://hal.inria.fr/hal-01953311

[23] G. Gautier, R. Bardenet, M. Valko.*On two ways to use determinantal point processes for Monte Carlo integration – Long version*, in "NeurIPS 2019 - Thirty-third Conference on Neural Information Processing Systems", Vancouver, Canada, Advances in Neural Information Processing Systems, 2019, https://hal.archives-ouvertes.fr/hal-02277739

[24] J.-B. Grill, O. D. Domingues, P. Ménard, R. Munos, M. Valko.*Planning in entropy-regularized Markov decision processes and games*, in "Neural Information Processing Systems", Vancouver, Canada, 2019, https://hal.inria.fr/hal-02387515

[25] E. Leurent, O.-A. Maillard.*Practical Open-Loop Optimistic Planning*, in "European Conference on Machine Learning", Würzburg, Germany, European Conference on Machine Learning, September 2019, https://arxiv.org/abs/1904.04700 , https://hal.archives-ouvertes.fr/hal-02375697

[26] A. LOCATELLI, A. CARPENTIER, M. VALKO.*Active multiple matrix completion with adaptive confidence sets*, in "International Conference on Artificial Intelligence and Statistics", Okinawa, Japan, 2019, https://hal.inria.fr/hal-02387468

[27] O.-A. MAILLARD.*Sequential change-point detection: Laplace concentration of scan statistics and non-asymptotic delay bounds*, in "Algorithmic Learning Theory", Chicago, United States, 2019, vol. 98, p. 1 - 23, https://hal.archives-ouvertes.fr/hal-02351665

[28] C. MOY, L. BESSON.*Decentralized Spectrum Learning for IoT Wireless Networks Collision Mitigation*, in "ISIoT 2019 - 1st International Workshop on Intelligent Systems for the Internet of Things", Santorin, Greece, May 2019, https://arxiv.org/abs/1906.00614 , https://hal.inria.fr/hal-02144465

[29] R. ORTNER, M. PIROTTA, R. FRUIT, A. LAZARIC, O.-A. MAILLARD.*Regret Bounds for Learning State Representations in Reinforcement Learning*, in "Conference on Neural Information Processing Systems", Vancouver, Canada, Conference on Neural Information Processing Systems, December 2019, https://hal.archives-ouvertes.fr/hal-02375715

[30] P. PERRAULT, V. PERCHET, M. VALKO.*Exploiting structure of uncertainty for efficient matroid semi-bandits*, in "International Conference on Machine Learning", Long Beach, United States, 2019, https://hal.inria.fr/hal-02387478

[31] P. PERRAULT, V. PERCHET, M. VALKO.*Finding the bandit in a graph: Sequential search-and-stop*, in "International Conference on Artificial Intelligence and Statistics", Okinawa, Japan, 2019, https://hal.inria.fr/hal-02387465

[32] J. SEZNEC, A. LOCATELLI, A. CARPENTIER, A. LAZARIC, M. VALKO.*Rotting bandits are not harder than stochastic ones*, in "International Conference on Artificial Intelligence and Statistics", Naha, Japan, 2019, https://hal.inria.fr/hal-01936894

[33] X. SHANG, E. KAUFMANN, M. VALKO.*A simple dynamic bandit algorithm for hyper-parameter tuning*, in "Workshop on Automated Machine Learning at International Conference on Machine Learning", Long Beach, United States, AutoML@ICML 2019 - 6th ICML Workshop on Automated Machine Learning, June 2019, https://hal.inria.fr/hal-02145200

[34] X. SHANG, E. KAUFMANN, M. VALKO.*General parallel optimization without a metric*, in "Algorithmic Learning Theory", Chicago, United States, 2019, vol. 98, https://hal.inria.fr/hal-02047225

[35] M. S. TALEBI, O.-A. MAILLARD.*Learning Multiple Markov Chains via Adaptive Allocation*, in "Advances in Neural Information Processing Systems 32 (NIPS 2019)", Vancouver, Canada, December 2019, https://hal.archives-ouvertes.fr/hal-02387345

National Conferences with Proceeding

[36] L. BESSON, E. KAUFMANN.*Non-asymptotic analysis of a sequential rupture detection test and its application to non-stationary bandits*, in "GRETSI 2019 - XXVIIème Colloque francophone de traitement du signal et des images", Lille, France, August 2019, https://hal.inria.fr/hal-02152243

Conferences without Proceedings

[37] L. BESSON, R. BONNEFOI, C. MOY.*GNU Radio Implementation of MALIN: "Multi-Armed bandits Learning for Internet-of-things Networks"*, in "IEEE WCNC 2019 - IEEE Wireless Communications and Networking Conference", Marrakech, Morocco, April 2019, https://arxiv.org/abs/1902.01734 , https://hal.inria.fr/hal-02006825

[38] R. BONNEFOI, L. BESSON, J. MANCO-VASQUEZ, C. MOY.*Upper-Confidence Bound for Channel Selection in LPWA Networks with Retransmissions*, in "The 1st International Workshop on Mathematical Tools and technologies for IoT and mMTC Networks Modeling", Marrakech, Morocco, Philippe Mary, Samir Perlaza, Petar Popovski, April 2019, https://arxiv.org/abs/1902.10615 - The source code (MATLAB or Octave) used for the simula-tions and the figures is open-sourced under the MIT License, atBitbucket.org/scee_ietr/ucb_smart_retrans, https://hal.inria.fr/hal-02049824

[39] Y. FLET-BERLIAC, P. PREUX.*MERL: Multi-Head Reinforcement Learning*, in "NeurIPS 2019 Deep Reinforcement Learning Workshop", Vancouver, Canada, December 2019, https://arxiv.org/abs/1909.11939 , https://hal.inria.fr/hal-02305105

[40] G. GAUTIER, R. BARDENET, M. VALKO.*On two ways to use determinantal point processes for Monte Carlo integration*, in "NEGDEPML 2019 - ICML Workshop on Negative Dependence in ML", Long Beach, CA, United States, June 2019, https://hal.archives-ouvertes.fr/hal-02160382

[41] T. LEVENT, P. PREUX, E. LE PENNEC, J. BADOSA, G. HENRI, Y. BONNASSIEUX.*Energy Management for Microgrids: a Reinforcement Learning Approach*, in "ISGT-Europe 2019 - IEEE PES Innovative Smart Grid Technologies Europe", Bucharest, France, IEEE, September 2019, p. 1-5 [*DOI :* 10.1109/ISGTEUROPE.2019.8905538], https://hal.archives-ouvertes.fr/hal-02382232

[42] M. SEURIN, P. PREUX, O. PIETQUIN.*"I'm sorry Dave, I'm afraid I can't do that" Deep Q-Learning From Forbidden Actions*, in "Workshop on Safety and Robustness in Decision Making (NeurIPS 2019)", Vancouver, Canada, December 2019, https://hal.inria.fr/hal-02387419

Other Publications

[43] L. BESSON, E. KAUFMANN.*The Generalized Likelihood Ratio Test meets klUCB: an Improved Algorithm for Piece-Wise Non-Stationary Bandits*, February 2019, https://arxiv.org/abs/1902.01575 - working paper or preprint, https://hal.inria.fr/hal-02006471

[44] E. BOURSIER, E. KAUFMANN, A. MEHRABIAN, V. PERCHET.*A Practical Algorithm for Multiplayer Bandits when Arm Means Vary Among Players*, May 2019, https://arxiv.org/abs/1902.01239 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02006069

[45] G. CIDERON, M. SEURIN, F. STRUB, O. PIETQUIN.*Self-Educated Language Agent With Hindsight Experience Replay For Instruction Following*, November 2019, https://arxiv.org/abs/1910.09451 - working paper or preprint [*DOI :* 10.09451], https://hal.archives-ouvertes.fr/hal-02386585

[46] R. DEGENNE, W. M. KOOLEN, P. MÉNARD.*Non-Asymptotic Pure Exploration by Solving Games*, December 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02402665

[47] Y. FLET-BERLIAC, P. PREUX.*High-Dimensional Control Using Generalized Auxiliary Tasks*, November 2019, working paper or preprint, https://hal.inria.fr/hal-02295705

[48] Y. FLET-BERLIAC, P. PREUX.*Samples Are Useful? Not Always: denoising policy gradient updates using variance explained*, September 2019, https://arxiv.org/abs/1904.04025 - working paper or preprint, https://hal.inria.fr/hal-02091547

[49] A. GARIVIER, H. HADIJI, P. MÉNARD, G. STOLTZ.*KL-UCB-switch: optimal regret bounds for stochastic bandits from both a distribution-dependent and a distribution-free viewpoints*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01785705

[50] A. GARIVIER, E. KAUFMANN.*Non-Asymptotic Sequential Tests for Overlapping Hypotheses and application to near optimal arm identification in bandit models*, May 2019, https://arxiv.org/abs/1905.03495 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02123833

[51] E. LEURENT, Y. BLANCO, D. EFIMOV, O.-A. MAILLARD.*Approximate Robust Control of Uncertain Dynamical Systems*, February 2019, https://arxiv.org/abs/1903.00220 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-01931744

[52] E. LEURENT, J. MERCAT.*Social Attention for Autonomous Decision-Making in Dense Traffic*, November 2019, https://arxiv.org/abs/1911.12250 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02383940

[53] O.-A. MAILLARD, T. A. MANN, R. ORTNER, S. MANNOR.*Active Roll-outs in MDP with Irreversible Dynamics*, July 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02177808

[54] X. SHANG, R. DE HEIDE, E. KAUFMANN, P. MÉNARD, M. VALKO.*Fixed-confidence guarantees for Bayesian best-arm identification*, October 2019, https://arxiv.org/abs/1910.10945 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02330187

[55] F. STRUB, M.-A. CHARPAGNE, T. M. POLLOCK.*Accurate reconstruction of EBSD datasets by a multimodal data approach using an evolutionary algorithm*, March 2019, https://arxiv.org/abs/1903.02988 - A short version of this paper exists towards people working in Machine Learning, namely arxiv:1903.02982 [*DOI :* 10.1016/J.MATCHAR.2019.01.033], https://hal.archives-ouvertes.fr/hal-02062104

[56] C. TRINH, E. KAUFMANN, C. VERNADE, R. COMBES.*Solving Bernoulli Rank-One Bandits with Unimodal Thompson Sampling*, December 2019, https://arxiv.org/abs/1912.03074 - working paper or preprint, https://hal.archives-ouvertes.fr/hal-02396943

References in notes

[57] P. AUER, N. CESA-BIANCHI, P. FISCHER.*Finite-time analysis of the multi-armed bandit problem*, in "Machine Learning", 2002, vol. 47, n^{o} 2/3, p. 235–256

[58] R. BELLMAN.*Dynamic Programming*, Princeton University Press, 1957

[59] D. BERTSEKAS, S. SHREVE.*Stochastic Optimal Control (The Discrete Time Case)*, Academic Press, New York, 1978

[60] D. BERTSEKAS, J. TSITSIKLIS.*Neuro-Dynamic Programming*, Athena Scientific, 1996

[61] M. PUTERMAN.*Markov Decision Processes: Discrete Stochastic Dynamic Programming*, John Wiley and Sons, 1994

[62] H. ROBBINS.*Some aspects of the sequential design of experiments*, in "Bull. Amer. Math. Soc.", 1952, vol. 55, p. 527–535

[63] R. SUTTON, A. BARTO.*Reinforcement learning: an introduction*, MIT Press, 1998

[64] P. WERBOS.*ADP: Goals, Opportunities and Principles*, IEEE Press, 2004, p. 3–44, Handbook of learning and approximate dynamic programming

Project-Team SPIRALS

Self-adaptation for distributed services and large software systems

IN COLLABORATION WITH: Centre de Recherche en Informatique, Signal et Automatique de Lille

IN PARTNERSHIP WITH:
Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Distributed Systems and middleware

Table of contents

Project-Team SPIRALS

Creation of the Team: 2014 January 01, updated into Project-Team: 2015 January 01

Keywords:

Computer Science and Digital Science:

A1.3. - Distributed Systems
A1.3.3. - Blockchain
A1.3.5. - Cloud
A1.4. - Ubiquitous Systems
A1.6. - Green Computing
A2.3.1. - Embedded systems
A2.3.2. - Cyber-physical systems
A2.4.2. - Model-checking
A2.5. - Software engineering
A2.5.2. - Component-based Design
A2.5.3. - Empirical Software Engineering
A2.5.4. - Software Maintenance & Evolution
A2.6.2. - Middleware
A3.1.3. - Distributed data
A3.1.4. - Uncertain data
A3.1.5. - Control access, privacy
A3.1.9. - Database
A3.2.1. - Knowledge bases
A3.2.4. - Semantic Web
A3.2.5. - Ontologies
A4.8. - Privacy-enhancing technologies
A7.2. - Logic in Computer Science
A9.1. - Knowledge

Other Research Topics and Application Domains:

B6.1. - Software industry
B6.4. - Internet of things
B6.5. - Information systems
B6.6. - Embedded systems
B8.5.2. - Crowd sourcing
B9.5.1. - Computer science
B9.5.6. - Data science
B9.10. - Privacy

1. Team, Visitors, External Collaborators

Research Scientists

Simon Bliudze [Inria, Researcher]

Pierre Bourhis [CNRS, Researcher]
Pierre Laperdrix [CNRS, Researcher, from Oct 2019]
Philippe Merle [Inria, Researcher, HDR]

Faculty Members

Lionel Seinturier [Team leader, Université de Lille, Professor, HDR]
Laurence Duchien [Université de Lille, Professor, HDR]
Clément Quinton [Université de Lille, Associate Professor]
Romain Rouvoy [Université de Lille, IUF, Professor, HDR]
Walter Rudametkin Ivey [Université de Lille, Associate Professor]

External Collaborator

Martin Monperrus [KTH Royal Institute of Technology, Sweden (on leave from Université de Lille), Professor]

Technical Staff

Jordan Bouchoucha [Université de Lille, Engineer, until Jun 2019]
Antoine Canda [Inria, Engineer]
Arthur d'Azémar [Inria, Engineer]
Julien Decaudin [Inria, Engineer, until Jun 2019]
Christophe Gourdin [Inria, Engineer, until Jun 2019]
Abderrahman Lahiaouni [Inria, Engineer, until Nov 2019]
Remy Raes [Inria, Engineer, from Sep 2019]
Larisa Safina [Inria, Engineer, from Oct 2019]
Faiez Zalila [Inria, Engineer, until Jul 2019]

PhD Students

Zeinab Abou Khalil [Université de Lille]
Mohammed Chakib Belgaid [Inria]
Sacha Brisset [Mantu]
Stéphanie Challita [Inria, until Mar 2019]
Benjamin Danglot [Inria, until Nov 2019]
Antonin Durey [Université de Lille]
Guillaume Fieni [Université de Lille]
Sarra Habchi [Inria]
Trinh Le Khanh [Université de Lille, from Oct 2019]
Lakhdar Meftah [Inria]
Vikas Mishra [Inria]
Zakaria Ournani [Orange]
Marion Tommasi [Inria]
Antoine Vastel [Université de Lille, until Oct 2019]

Post-Doctoral Fellows

Alexandre Garnier [Inria]
Louis Jachiet [CNRS, until Aug 2019]
Hoang Gia Nguyen [Inria]

Visiting Scientists

Fatima Belkouch Ameline [Université de Lille, Associate Professor, until Sep 2019]
Jonatan Enes Alvarez [University of A Coruña, Spain, PhD Student, from Apr 2019 until Jul 2019]
Alejandro Grez [Pontifical Catholic University of Chile, Apr 2019]

Administrative Assistants

Karine Lewandowski [Inria]
Nathalie Bonte [Inria]

2. Overall Objectives

2.1. Introduction

Our research is based on two complementary fields: distributed systems and software engineering. We aim at introducing more automation in the adaptation processes of software systems, that is, transitioning from the study of adaptive systems to self-adaptive systems. In particular, we work towards two directions: self-healing software systems with data mining solutions, and self-optimizing software systems with context monitoring. These two objectives are applied to two target environments: mobile computing and cloud computing.

2.2. Scientific Foundations

Distributed software services and systems are central to many human activities, such as communication, commerce, education, defense, etc. Distributed software services consist of an ever growing number of devices, often highly heterogeneous, from cloud platforms, sensor networks, to application servers, desktop machines, and mobile devices, such as smartphones. The future of this huge number of interconnected software services has been called the Internet of Services, a vision *"where everything that is needed to use software applications is available as a service on the Internet, such as the software itself, the tools to develop the software, the platform servers, storage and communication to run the software."*[0] This pervasiveness continuously leads to new usages that in turn foster the emergence of novel requirements and concepts for new software services. Hence, it is necessary to establish new paradigms to design and execute software programs in these highly interconnected and heterogeneous environments, and it is necessary to ensure not only that these software systems can be adapted to new usages, new infrastructures, and new execution environments in the long term, but also that after the adaptation process the services still perform as expected.

This research project focuses on defining *self-adaptive* software services and middleware. From the perspective of the Internet of Services, this project fits in the vision sketched by *e.g.*, the FP8 Expert Group Services in the Future Internet [73], the NESSI Research Priorities for the next Framework Programme for Research and Technological Development FP8 [77], the Roadmap for Advanced Cloud Technologies under H2020 [74], and research roadmaps, such as [82], [72], [61].

3. Research Program

3.1. Introduction

Our research program on self-adaptive software targets two key properties that are detailed in the remainder of this section: *self-healing* and *self-optimization*.

3.2. Objective #1: Self-healing - Mining software artifacts to automatically evolve systems

Software systems are under the pressure of changes all along their lifecycle. Agile development blurs the frontier between design and execution and requires constant adaptation. The size of systems (millions of lines of code) multiplies the number of bugs by the same order of magnitude. More and more systems, such as sensor network devices, live in "surviving" mode, in the sense that they are neither rebootable nor upgradable.

Software bugs are hidden in source code and show up at development-time, testing-time or worse, once deployed in production. Except for very specific application domains where formal proofs are achievable, bugs can not be eradicated. As an order of magnitude, on 16 Dec 2011, the Eclipse bug repository contains 366 922 bug reports. Software engineers and developers work on bug fixing on a daily basis. Not all developers spend the same time on bug fixing. In large companies, this is sometimes a full-time role to manage bugs, often referred to as *Quality Assurance* (QA) software engineers. Also, not all bugs are equal, some bugs are analyzed and fixed within minutes, others may take months to be solved [79].

[0] http://cordis.europa.eu/fp7/ict/ssai

In terms of research, this means that: (i) one needs means to automatically adapt the design of the software system through automated refactoring and API extraction, (ii) one needs approaches to automate the process of adapting source code in order to fix certain bugs, (iii) one needs to revisit the notion of error-handling so that instead of crashing in presence of errors, software adapts itself to continue with its execution, *e.g.*, in degraded mode.

There is no one-size-fits-all solution for each of these points. However, we think that novel solutions can be found by using **data mining and machine learning techniques tailored for software engineering** [80]. This body of research consists of mining some knowledge about a software system by analyzing the source code, the version control systems, the execution traces, documentation and all kinds of software development and execution artifacts in general. This knowledge is then used within recommendation systems for software development, auditing tools, runtime monitors, frameworks for resilient computing, etc.

The novelty of our approach consists of using and tailoring data mining techniques for analyzing software artifacts (source code, execution traces) in order to achieve the **next level of automated adaptation** (*e.g.*, automated bug fixing). Technically, we plan to mix unsupervised statistical learning techniques (*e.g.* frequent item set mining) and supervised ones (*e.g.* training classifiers such as decision trees). This research is currently not being performed by data mining research teams since it requires a high level of domain expertise in software engineering, while software engineering researchers can use off-the-shelf data mining libraries, such as Weka [58].

We now detail the two directions that we propose to follow to achieve this objective.

3.2.1. Learning from software history how to design software and fix bugs

The first direction is about mining techniques in software repositories (*e.g.*, CVS, SVN, Git). Best practices can be extracted by data mining source code and the version control history of existing software systems. The design and code of expert developers significantly vary from the artifacts of novice developers. We will learn to differentiate those design characteristics by comparing different code bases, and by observing the semantic refactoring actions from version control history. Those design rules can then feed the test-develop-refactor constant adaptation cycle of agile development.

Fault localization of bugs reported in bug repositories. We will build a solid foundation on empirical knowledge about bugs reported in bug repository. We will perform an empirical study on a set of representative bug repositories to identify classes of bugs and patterns of bug data. For this, we will build a tool to browse and annotate bug reports. Browsing will be helped with two kinds of indexing: first, the tool will index all textual artifacts for each bug report; second it will index the semantic information that is not present by default in bug management software—*i.e.*, "contains a stacktrace"). Both indexes will be used to find particular subsets of bug reports, for instance "all bugs mentioning invariants and containing a stacktrace". Note that queries with this kind of complexity and higher are mostly not possible with the state-of-the-art of bug management software. Then, analysts will use annotation features to annotate bug reports. The main outcome of the empirical study will be the identification of classes of bugs that are appropriate for automated localization. Then, we will run machine learning algorithms to identify the latent links between the bug report content and source code features. Those algorithms would use as training data the existing traceability links between bug reports and source code modifications from version control systems. We will start by using decision trees since they produce a model that is explicit and understandable by expert developers. Depending on the results, other machine learning algorithms will be used. The resulting system will be able to locate elements in source code related to a certain bug report with a certain confidence.

Automated bug fix generation with search-based techniques. Once a location in code is identified as being the cause of the bug, we can try to automatically find a potential fix. We envision different techniques: (1) infer fixes from existing contracts and specifications that are violated; (2) infer fixes from the software behavior specified as a test suite; (3) try different fix types one-by-one from a list of identified bug fix patterns; (4) search fixes in a fix space that consists of combinations of atomic bug fixes. Techniques 1 and 2 are explored in [54] and [78]. We will focus on the latter techniques. To identify bug fix patterns and atomic bug fixes, we will perform a large-scale empirical study on software changes (also known as changesets when referring to changes across multiple files). We will develop tools to navigate, query and annotate changesets in a version control system. Then, a grounded theory will be built to master the nature of fixes. Eventually, we will decompose change sets in atomic actions using clustering on changeset actions. We will then use this body of empirical knowledge to feed search-based algorithms (*e.g.* genetic algorithms) that will look for meaningful fixes in a large fix space. To sum up, our research on automated bug fixing will try not only to point to source code locations responsible of a bug, but to search for code patterns and snippets that may constitute the skeleton of a valid patch. Ultimately, a blend of expert heuristics and learned rules will be able to produce valid source code that can be validated by developers and committed to the code base.

3.2.2. Run-time self-healing

The second proposed research direction is about inventing a self-healing capability at run-time. This is complementary to the previous objective that mainly deals with development time issues. We will achieve this in two steps. First, we want to define frameworks for resilient software systems. Those frameworks will help to maintain the execution even in the presence of bugs—*i.e.* to let the system survive. As exposed below, this may mean for example to switch to some degraded modes. Next, we want to go a step further and to define solutions for automated runtime repair, that is, not simply compensating the erroneous behavior, but also determining the correct repair actions and applying them at run-time.

Mining best effort values. A well-known principle of software engineering is the "fail-fast" principle. In a nutshell, it states that as soon as something goes wrong, software should stop the execution before entering incorrect states. This is fine when a human user is in the loop, capable of understanding the error or at least rebooting the system. However, the notion of “failure-oblivious computing” [71] shows that in certain domains, software should run in a resilient mode (*i.e.* capable of recovering from errors) and/or best-effort mode—*i.e.* a slightly imprecise computation is better than stopping. Hence, we plan to investigate data mining techniques in order to learn best-effort values from past executions (*i.e.* somehow learning what is a correct state, or the opposite what is not a completely incorrect state). This knowledge will then be used to adapt the software state and flow in order to mitigate the error consequences, the exact opposite of fail-fast for systems with long-running cycles.

Embedding search based algorithms at runtime. Harman recently described the field of search-based software engineering [59]. We believe that certain search based approaches can be embedded at runtime with the goal of automatically finding solutions that avoid crashing. We will create software infrastructures that allow automatically detecting and repairing faults at run-time. The methodology for achieving this task is based on three points: (1) empirical study of runtime faults; (2) learning approaches to characterize runtime faults; (3) learning algorithms to produce valid changes to the software runtime state. An empirical study will be performed to analyze those bug reports that are associated with runtime information (*e.g.* core dumps or stacktraces). After this empirical study, we will create a system that learns on previous repairs how to produce small changes that solve standard runtime bugs (*e.g.* adding an array bound check to throw a handled domain exception rather than a spurious language exception). To achieve this task, component models will be used to (1) encapsulate the monitoring and reparation meta-programs in appropriate components and (2) support runtime code modification using scripting, reflective or bytecode generation techniques.

3.3. Objective #2: Self-optimization - Sharing runtime behaviors to continuously adapt software

Complex distributed systems have to seamlessly adapt to a wide variety of deployment targets. This is due to the fact that developers cannot anticipate all the runtime conditions under which these systems are immersed. A major challenge for these software systems is to develop their capability to continuously reason about themselves and to take appropriate decisions and actions on the optimizations they can apply to improve themselves. This challenge encompasses research contributions in different areas, from environmental monitoring to real-time symptoms diagnosis, to automated decision making. The variety of distributed systems, the number of optimization parameters, and the complexity of decisions often resign the practitioners to design monolithic and static middleware solutions. However, it is now globally acknowledged that the development of dedicated building blocks does not contribute to the adoption of sustainable solutions. This is confirmed by the scale of actual distributed systems, which can—for example—connect several thousands of devices to a set of services hosted in the Cloud. In such a context, the lack of support for smart behaviors at different levels of the systems can inevitably lead to its instability or its unavailability. In June 2012, an outage of Amazon's Elastic Compute Cloud in North Virginia has taken down Netflix, Pinterest, and Instagram services. During hours, all these services failed to satisfy their millions of customers due to the lack of integration of a self-optimization mechanism going beyond the boundaries of Amazon.

The research contributions we envision within this area will therefore be organized as a reference model for engineering **self-optimized distributed systems** autonomously driven by *adaptive feedback control loops*, which will automatically enlarge their scope to cope with the complexity of the decisions to be taken. This solution introduces a multi-scale approach, which first privileges local and fast decisions to ensure the homeostasis [0] property of a single node, and then progressively propagates symptoms in the network in order to reason on a longer term and a larger number of nodes. Ultimately, domain experts and software developers can be automatically involved in the decision process if the system fails to find a satisfying solution. The research program for this objective will therefore focus on the study of mechanisms for **monitoring, taking decisions, and automatically reconfiguring software at runtime and at various scales**. As stated in the self-healing objective, we believe that there is no one-size-fits-all mechanism that can span all the scales of the system. We will therefore study and identify an optimal composition of various adaptation mechanisms in order to produce long-living software systems.

The novelty of this objective is to exploit the wisdom of crowds to define new middleware solutions that are able to continuously adapt software deployed in the wild. We intend to demonstrate the applicability of this approach to distributed systems that are deployed from mobile phones to cloud infrastructures. The key scientific challenges to address can be summarized as follows: *How does software behave once deployed in the wild? Is it possible to automatically infer the quality of experience, as it is perceived by users? Can the runtime optimizations be shared across a wide variety of software? How optimizations can be safely operated on large populations of software instances?*

The remainder of this section further elaborates on the opportunities that can be considered within the frame of this objective.

3.3.1. Monitoring software in the wild

Once deployed, developers are generally no longer aware of how their software behave. Even if they heavily use testbeds and benchmarks during the development phase, they mostly rely on the bugs explicitly reported by users to monitor the efficiency of their applications. However, it has been shown that contextual artifacts collected at runtime can help to understand performance leaks and optimize the resilience of software systems [81]. Monitoring and understanding the context of software at runtime therefore represent the first building block of this research challenge. Practically, we intend to investigate crowd-sensing approaches, to smartly collect and process runtime metrics (*e.g.*, request throughput, energy consumption, user context). Crowd-sensing can be seen as a specific kind of crowdsourcing activity, which refers to the capability of lifting a (large) diffuse group of participants to delegate the task of retrieving trustable data from the field. In particular, crowd-sensing covers not only *participatory sensing* to involve the user in the sensing task (*e.g.*, surveys), but also *opportunistic sensing* to exploit mobile sensors carried by the user (*e.g.*, smartphones).

While reported metrics generally enclose raw data, the monitoring layer intends to produce meaningful indicators like the *Quality of Experience* (QoE) perceived by users. This QoE reflects representative symptoms of software requiring to trigger appropriate decisions in order to improve its efficiency. To diagnose these symptoms, the system has to process a huge variety of data including runtime metrics, but also history of logs to explore the sources of the reported problems and identify opportunities for optimizations. The techniques we envision at this level encompass machine learning, principal component analysis, and fuzzy logic [70] to provide enriched information to the decision level.

[0] Homeostasis is the property of a system that regulates its internal environment and tends to maintain a stable, relatively constant condition of properties [Wikipedia].

3.3.2. Collaborative decision-making approaches

Beyond the symptoms analysis, decisions should be taken in order to improve the *Quality of Service* (QoS). In our opinion, collaborative approaches represent a promising solution to effectively converge towards the most appropriate optimization to apply for a given symptom. In particular, we believe that exploiting the wisdom of the crowd can help the software to optimize itself by sharing its experience with other software instances exhibiting similar symptoms. The intuition here is that the body of knowledge that supports the optimization process cannot be specific to a single software instance as this would restrain the opportunities for improving the quality and the performance of applications. Rather, we think that any software instance can learn from the experience of others.

With regard to the state-of-the-art, we believe that a multi-levels decision infrastructure, inspired from distributed systems like Spotify [57], can be used to build a decentralized decision-making algorithm involving the surrounding peers before requesting a decision to be taken by more central control entity. In the context of collaborative decision-making, peer-based approaches therefore consist in quickly reaching a consensus on the decision to be adopted by a majority of software instances. Software instances can share their knowledge through a micro-economic model [51], that would weight the recommendations of experienced instances, assuming their age reflects an optimal configuration.

Beyond the peer level, the adoption of algorithms inspired from evolutionary computations, such as genetic programming, at an upper level of decision can offer an opportunity to test and compare several alternative decisions for a given symptom and to observe how does the crowd of applications evolves. By introducing some diversity within this population of applications, some instances will not only provide a satisfying QoS, but will also become naturally resilient to unforeseen situations.

3.3.3. Smart reconfigurations in the large

Any decision taken by the crowd requires to propagate back to and then operated by the software instances. While simplest decisions tend to impact software instances located on a single host (*e.g.*, laptop, smartphone), this process can also exhibit more complex reconfiguration scenarios that require the orchestration of various actions that have to be safely coordinated across a large number of hosts. While it is generally acknowledged that centralized approaches raise scalability issues, we think that self-optimization should investigate different reconfiguration strategies to propagate and apply the appropriate actions. The investigation of such strategies can be addressed in two steps: the consideration of *scalable data propagation protocols* and the identification of *smart reconfiguration mechanisms*.

With regard to the challenge of scalable data propagation protocols, we think that research opportunities encompass not only the exploitation of gossip-based protocols [56], but also the adoption of publish/subscribe abstractions [64] in order to decouple the decision process from the reconfiguration. The fundamental issue here is the definition of a communication substrate that can accommodate the propagation of decisions with relaxed properties, inspired by *Delay Tolerant Networks* (DTN), in order to reach weakly connected software instances. We believe that the adoption of asynchronous communication protocols can provide the sustainable foundations for addressing various execution environments including harsh environments, such as developing countries, which suffer from a partial connectivity to the network. Additionally, we are interested in developing the principle of *social networks of applications* in order to seamlessly group and organize software instances according to their similarities and acquaintances. The underlying idea is that grouping application instances can contribute to the identification of optimization profiles not only contributing to the monitoring layer, but also interested in similar reconfigurations. Social networks of applications can contribute to the anticipation of reconfigurations by exploiting the symptoms of similar applications to improve the performance of others before that problems actually happen.

With regard to the challenge of smart reconfiguration mechanisms, we are interested in building on our established experience of adaptive middleware [75] in order to investigate novel approaches to efficient application reconfigurations. In particular, we are interested in adopting seamless micro-updates and micro-reboot techniques to provide in-situ reconfiguration of pieces of software. Additionally, the provision of safe and secured reconfiguration mechanisms is clearly a key issue that requires to be carefully addressed in order to avoid malicious exploitation of dynamic reconfiguration mechanisms against the software itself. In this area, although some reconfiguration mechanisms integrate transaction models [65], most of them are restricted to local reconfigurations, without providing any support for executing distributed reconfiguration transactions. Additionally, none of the approached published in the literature include security mechanisms to preserve from unauthorized or malicious reconfigurations.

4. Application Domains

4.1. Introduction

Although our research is general enough to be applied to many application domains, we currently focus on applications and distributed services for the retail industry and for the digital home. These two application domains are supported by a strong expertise in mobile computing and in cloud computing that are the two main target environments on which our research prototypes are built, for which we are recognized, and for which we have already established strong collaborations with the industrial ecosystem.

4.2. Distributed software services for the retail industry

This application domain is developed in relation with the PICOM (*Pôle de compétivité Industries du Commerce*) cluster. We have established strong collaborations with local companies in the context of former funded projects, such as Cappucino and Macchiato, which focused on the development of a new generation of mobile computing platforms for e-commerce. We are also involved in the Datalyse and OCCIware funded projects that define cloud computing environments with applications for the retail industry. Finally, our activities in terms of crowd-sensing and data gathering on mobile devices with the APISENSE® platform share also applications for the retail industry.

4.3. Distributed software services for the digital home

We are developing new middleware solutions for the digital home, in particular through our long standing collaboration with Orange Labs. We are especially interested in developing energy management and saving solutions with the POWERAPI software library for distributed environments such the ones that equip digital homes. We are also working to bridge the gap between distributed services hosted on home gateways and distributed services hosted on the cloud to be able to smoothly transition between both environments. This work is especially conducted with the SALOON platform.

5. Highlights of the Year

5.1. Highlights of the Year

In 2019, Christophe Gourdin and Philippe Merle have created the XScalibur company. XScalibur is a startup company that sells an innovative software solution to design, deploy and monitor software systems in a multi-cloud environment. The company is the result of a transfert activity initiated by Philippe Merle and Christophe Gourdin around the OCCIware Studio software tool suite. This model-driven based solution for cloud management is the result of several years of research [26], [15] and has especially been developed in the context of the OCCIware collaborative project from 2014 to 2017. XScalibur has been selected in March 2019 by the Alliancy magazine in the top-13 of startup companies in the domain of cloud computing to "follow closely"[0].

[0] https://www.alliancy.fr/a-laffiche/cloud/2019/03/13/13-start-up-du-cloud-a-suivre-de-pres

In 2019, Laurence Duchien has been general chair of the 13th edition of the European Conference on Software Architecture (ECSA) [49], [48] and program co-chair of the 23rd edition of the International Systems and Software Product Line Conference (SPLC) [47], [46]. These two events have been co-located in Paris from 9 to 13 September 2019. The fact that a member of the Spirals project-team was proposed by the software architecture research community to serve in these two major events is a testimony of the recognition and of the visibility of our research activities in this domain.

5.1.1. Awards

Walter Rudametkin and Pierre Laperdrix were awarded in January 2019 the *Prix Inria CNIL protection de la vie privée*. The award was announced during the 12th edition of the International Computers, Privacy and Data Protection (CPDP) conference, and rewards research undertaken with a view to creating a trustworthy digital society. The award was granted thanks to the work of Walter Rudametkin, Pierre Laperdrix, and Benoit Baudry [63] on browser fingerprinting.

Thomas Durieux was awarded in June 2019 a honorable mention (*accessit*) at *Prix de thèse GDR GPL* for his PhD work on software automated repair that was defended in September 2018 [55]. GDR GPL (*Génie de la Programmation et du Logiciel*) is the group that gathers the French research community on software engineering and programming languages. This is the third time that a PhD student from the Spirals project-team wins either this prize or an honorable mention (Clément Quinton won the prize in 2014, and Maria Gomez won a honorable mention in 2017).

Lakhdar Meftah and Romain Rouvoy won a Best Paper award at the 19th International Conference on Distributed Applications and Interoperable Systems (DAIS 2019) [37]. This award distinguishes their work on improving user privacy in crowdsourced mobile datasets. They propose a decentralized approach, named Fougere, to convey data samples from user devices to third-party servers. By introducing an a priori data anonymization process, they show that Fougere defeats state-of-the-art location-based privacy attacks with little impact on the quality of crowd-sourced datasets. This work takes place in the context of the Inria Project Lab BetterNet and involves Isabelle Chrisment, who is heading the BetterNet IPL, and who is co-supervising with Romain Rouvoy the PhD thesis of Lakhdar Meftah [13] that was defended in December 2019.

Philippe Merle won a Best Demo Paper award at the 5th IEEE International Conference on Network Softwarization (NetSoft 2019) [39]. This award distinguishes his work on the formal verification of virtualized network and cloud environments. Philippe Merle and his co-authors demonstrate a lightweight toolchain for validating descriptors of network functions compliant with the latest ETSI NFV standards, visualizing these descriptors in the form of various diagrams (i.e. network-oriented, function-oriented, and UML2-based), analyzing these descriptors formally with Alloy Analyzer, and deploying these virtualized network functions on OpenStack. This work takes place in the context of a funded project that ended in 2019 with the Orange telecom operator and that involved, in addition to Philippe Merle, Jean-Bernard Stefani from the Inria Grenoble - Rhône-Alpes research center.

Laurence Duchien was awarded the rank of Chevalière de l'Ordre National du Mérite (JO du 29 mai 2019).

BEST PAPERS AWARDS :

[37]
L. MEFTAH, R. ROUVOY, I. CHRISMENT.*FOUGERE: User-Centric Location Privacy in Mobile Crowdsourcing Apps*, in "19th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS)", Kongens Lyngby, Denmark, J. PEREIRA, L. RICCI (editors), Distributed Applications and Interoperable Systems, Springer International Publishing, 2019, vol. LNCS-11534, p. 116-132 [*DOI :* 10.1007/978-3-030-22496-7_8], https://hal.inria.fr/hal-02121311

[39]
P. MERLE, A. NDEYE SYLLA, M. OUZZIF, F. KLAMM, K. GUILLOUARD.*A Lightweight Toolchain to Validate, Visualize, Analyze, and Deploy ETSI NFV Topologies*, in "NetSoft 2019 - The 5th IEEE International Conference on Network Softwarization", Paris, France, June 2019 [*DOI :* 10.1109/NETSOFT.2019.8806632], https://hal.archives-ouvertes.fr/hal-02124164

6. New Software and Platforms

6.1. APISENSE

KEYWORDS: Mobile sensing - Crowd-sensing - Mobile application - Crowd-sourcing - Android

FUNCTIONAL DESCRIPTION: APISENSE platform is a software solution to collect various contextual information from Android devices (client application) and automatically upload collected data to a server (deployed as a SaaS). APISENSE is based on a Cloud computing infrastructure to facilitate datasets collection from significant populations of mobile users for research purposes.

- Participants: Antoine Veuiller, Christophe Ribeiro, Julien Duribreux, Nicolas Haderer, Romain Rouvoy, Romain Sommerard and Lakhdar Meftah
- Partner: Université de Lille
- Contact: Romain Rouvoy
- URL: https://apisense.io

6.2. PowerAPI

KEYWORDS: Energy efficiency - Energy management

FUNCTIONAL DESCRIPTION: PowerAPI is a library for monitoring the energy consumption of software systems.

PowerAPI differs from existing energy process-level monitoring tool in its software orientation, with a fully customizable and modular solution that let the user to precisely define what he/she wants to monitor. PowerAPI is based on a modular and asynchronous event-driven architecture using the Akka library. PowerAPI offers an API which can be used to define requests about energy spent by a process, following its hardware resource utilization (in term of CPU, memory, disk, network, etc.).

- Participants: Adel Noureddine, Loïc Huertas, Maxime Colmant, Romain Rouvoy, Mohammed Chakib Belgaid and Arthur D'azemar
- Contact: Romain Rouvoy
- URL: http://powerapi.org

6.3. Saloon

KEYWORDS: Feature Model - Software Product Line - Cloud computing - Model-driven engineering - Ontologies

FUNCTIONAL DESCRIPTION: Saloon is a framework for the selection and configuration of Cloud providers according to application requirements. The framework enables the specification of such requirements by defining ontologies. Each ontology provides a unified vision of provider offers in terms of frameworks, databases, languages, application servers and computational resources (i.e., memory, storage and CPU frequency). Furthermore, each provider is related to a Feature Model (FM) with attributes and cardinalities, which captures its capabilities. By combining the ontology and FMs, the framework is able to match application requirements with provider capabilities and select a suitable one. Specific scripts to the selected provider are generated in order to enable its configuration.

- Participants: Clément Quinton, Daniel Romero Acero, Laurence Duchien, Lionel Seinturier and Romain Rouvoy
- Partner: Université de Lille
- Contact: Clément Quinton
- URL: https://gitlab.irisa.fr/drome00A/saloon

6.4. SPOON

KEYWORDS: Java - Code analysis

FUNCTIONAL DESCRIPTION: Spoon is an open-source library that enables you to transform (see below) and analyze Java source code (see example). Spoon provides a complete and fine-grained Java metamodel where any program element (classes, methods, fields, statements, expressions...) can be accessed both for reading and modification. Spoon takes as input source code and produces transformed source code ready to be compiled.

- Participants: Gérard Paligot, Lionel Seinturier, Martin Monperrus, Nicolas Petitprez and Simon Urli
- Contact: Martin Monperrus
- URL: http://spoon.gforge.inria.fr

7. New Results

7.1. Browser fingerprinting

We obtained new results on the concept of browser fingerprinting. This is a major technique of Internet security that is widely used for many purposes such as tracking activities, enhancing authentication, detecting bots, just to name a few. These results contribute to the enhancement of security for distributed software systems.

Our contributions to browser fingerprinting include the following three elements. First, we collected 122K fingerprints from 2 346 browsers and studied their stability over more than 2 years. We showed that, despite frequent changes in the fingerprints, a significant fraction of browsers can be tracked over a long period of time. Second, we designed a test suite to evaluate fingerprinting countermeasures. We applied our test suite to 7 countermeasures, some of them claiming to generate consistent fingerprints, and show that all of them can be identified, which can make their users more identifiable. Third, we explored the use of browser fingerprinting for crawler detection. We measured its use in the wild, as well as the main detection techniques. Since fingerprints are collected on the client-side, we also evaluated its resilience against an adversarial crawler developer that tries to modify its crawler fingerprints to bypass security checks.

These results have been obtained in the context of the PhD thesis of Antoine Vastel [14] defended in October 2019.

7.2. Test amplification

With respect to self-healing, we proposed a new algorithm for test amplification. Test amplification consists of exploiting the knowledge of test methods, in which developers embed input data and expected properties, in order to enhance these tests [22].

We proposed a new approach based on test inputs transformation and assertions generation to amplify test suites, and implemented this approach in the DSpot software tool that we created [21]. By evaluating DSpot on open-source projects from GitHub, we showed that we improve the mutation score of test suites. These improvements have been proposed to developers through pull requests: their feedbacks show that they value the output of DSpot by accepting to integrate amplified test methods into their test suite. This proves that DSpot can improve the quality of the test suite of real projects. We also showed that DSpot can generate amplified test methods that specify behavioral changes, and can generate amplified test methods to improve the ability to detect potential regressions.

These results have been obtained in the context of the STAMP H2020 project and in the context of the PhD thesis of Benjamin Danglot [11] defended in November 2019.

7.3. Understanding mobile-specific code smells

With respect to self-healing, we obtained new results in the domain of code smells for mobile software systems. Code smells are well-known concepts in software engineering. They refer to bad design and development practices commonly observed in software systems.

We obtained three new results that contribute to a better understanding of mobile code smells. First, we studied the expansion of code smells in different mobile platforms. Then, we conducted a large-scale study to analyze the change history of mobile apps and discern the factors that favor the introduction and survival of code smells. To consolidate these studies, we also performed a user study to investigate developers' perception of code smells and the adequacy of static analyzers as a solution for coping with them. Finally, we performed a qualitative study to question the established foundation about the definition and detection of mobile code smells. The results of these studies revealed important research findings. Notably, we showed that pragmatism, prioritization, and individual attitudes are not relevant factors for the accrual of mobile code smells. The problem is rather caused by ignorance and oversight, which are prevalent among mobile developers. Furthermore, we highlighted several flaws in the code smell definitions that are currently adopted by the research community. These results allowed us to elaborate some recommendations for researchers and tool makers willing to design detection and refactoring tools for mobile code smells [33], [34]. On top of that, our results opened perspectives for research works about the identification of mobile code smells and development practices in general.

These results have been obtained in the context of the PhD thesis of Sarra Habchi [12] defended in December 2019.

7.4. Towards privacy-sensitive mobile crowdsourcing

We obtained new results in the domain of data privacy for crowdsourced data.

We proposed an anonymous data collection library for mobile apps, a software library that improves the user's privacy without compromising the overall quality of the crowdsourced dataset. In particular, we proposed a decentralized approach, named FOUGERE, to convey data samples from user devices using peer-to-peer (P2P) communications to third-party servers, thus introducing an a priori data anonymization process that is resilient to location-based attacks. To validate the approach, we proposed a testing framework to test this P2P communication library, named PeerFleet. Beyond the identification of P2P-related errors, PeerFleet also helps to tune the discovery protocol settings to optimize the deployment of P2P apps. We validated FOUGERE using 500 emulated devices that replay a mobility dataset and use FOUGERE to collect location data. We evaluated the overhead, the privacy and the utility of FOUGERE. We showed that FOUGERE defeats the state-of-the-art location-based privacy attacks with little impact on the quality of the collected data [38], [5].

These results have been obtained in the context of the PhD thesis of Lakhdar Meftah [13] defended in December 2019.

8. Bilateral Contracts and Grants with Industry

8.1. ip-label

Participant: Romain Rouvoy [contact person].

A software exploitation license (2014–ongoing) of the APISENSE® crowd-sensing platform has been sold to the ip-label company. They use this platform as a solution to monitor the quality of the GSM signal in the wild. The objective is to provide developers and stakeholders with a feedback on the quality of experience of GSM connection depending on their location.

8.2. Davidson

Participants: Mohammed Chakib Belgaid, Romain Rouvoy [contact person], Lionel Seinturier.

This collaboration (2017–20) aims at proposing new solutions for optimizing the energy footprint of ICT software infrastructures. We want to be able to measure and assess the energy footprint of ICT systems while preserving various quality of service parameters, such as performance and security. We aim at proposing a testbed for assessing the energy footprint of various programming languages. This testbed will also incorporate frameworks for web and mobile programming. Finally, we want to be able to issue recommendations to developers in order to assist them in improving the energy footprint of their programs. This collaboration will take advantage of the POWERAPI software library.

The PhD of Mohammed Chakib Belgaid takes place in the context of this collaboration.

8.3. Orange #1

Participants: Philippe Merle [contact person], Lionel Seinturier.

This collaboration (2017–19) aims at defining a computational model for software infrastructures layered on top of virtualized and interconnected cloud resources. This computational model provides application programming and management facilities to distributed applications and services [39], [44]. This computational model defines a pivot model that enables the interoperability of various existing and future standards for cloud systems such as OCCI and TOSCA. This pivot model is defined with the Alloy specification language [62]. This collaboration takes advantage of the expertise that we are developing since several years on reconfigurable component-based software systems [75], on cloud systems [69], and on the Alloy specification language [67].

This collaboration with Orange Labs is a joint project with Jean-Bernard Stefani from the Spades Inria project-team.

8.4. Orange #2

Participants: Zakaria Ournani, Romain Rouvoy [contact person], Lionel Seinturier.

This collaboration (2018–21) aims at proposing new solutions for modeling the energy efficiency of software systems and to design and implement new methods for measuring and reducing the energy consumption of software systems at development time. We especially target software systems deployed on cloud environments.

The CIFRE PhD of Zakaria Ournani takes place in the context of this collaboration.

8.5. Amaris (now Mantu)

Participants: Sacha Brisset, Romain Rouvoy [contact person], Lionel Seinturier.

This collaboration (2018–21) aims at proposing new solutions for automatically spotting and fixing recurrent user experience issues in web applications. We are interested in developing an autonomic framework that learns and classifies the behaviors and figures out causality links between data such as web GUI events, support tickets and user feedback, source version management events (e.g. recent commits). The ultimate objective is to implement an AI-powered recommendation system to guide the maintenance and even to automatically predict and solve user issues.

The CIFRE PhD of Sacha Brisset takes place in the context of this collaboration.

9. Partnerships and Cooperations

9.1. Regional Initiatives

9.1.1. Région Hauts-de-France

9.1.1.1. CIRRUS

Participants: Stéphanie Challita, Guillaume Fieni, Alexandre Garnier, Christophe Gourdin, Philippe Merle [contact person], Romain Rouvoy, Lionel Seinturier, Faiez Zalila.

CIRRUS is a 3-year (2017–20) joint team with the Scalair cloud operator and architect company funded by the Hauts-de-France region. The CIRRUS joint team is developing novel solutions in the domains of the on demand configuration of heterogeneous cloud resources, the management of cloud elasticity for all deployed services (SaaS, PaaS, IaaS) in order to guarantee quality of service and user quality of experience, and the taming of financial costs of cloud infrastructures.

9.1.1.2. Alloy@Scale

Participants: Abderrahman Lahiaouni, Philippe Merle [contact person], Romain Rouvoy, Lionel Seinturier.

Alloy@Scale is a 12-month (2018–19) project funded in the context of CPER Data program. Alloy@Scale aims at overcoming the limits of the formal verification of large software systems specified with the Alloy formal specification language. For that, the program combines the Grid'5000 infrastructure and the Docker container technology.

9.1.1.3. Rigorous Component-Based Design of Correct-by-Construction Software and Systems: Application to Cloud Computing

Participants: Simon Bliudze [contact person], Larisa Safina.

This 24-month (2019–20) project is funded in the context of the STaRS program. It aims at the development of methods and tools for rigorous design of cloud computing platforms and applications, which can be proven to be correct by construction. First results have been published in [17], [31], [3].

9.1.1.4. Indoor Analytics

Participants: Pierre Bourhis, Remy Raes, Romain Rouvoy [contact person], Lionel Seinturier.

Indoor Analytics is a 32-month (2019–21) project funded in the context of CPER Data program. Indoor Analytics aims at collaborating with the Mapwize company on the development of novel analytics for indoor location systems. In particular, Mapwize and Spirals target the joint delivery of an open-source software solution devoted to the acquisition, storage and processing of location events at scale.

9.1.1.5. COMMODE

Participants: Pierre Bourhis, Laurence Duchien, Clément Quinton [contact person].

COMMODE (Knowledge COMpilation for feature MODEls) is a 24-month (2019–21) project funded in the context of CPER Data program. COMMODE aims at using techiques from knowledge compilation, a subarea of artificial intelligence, for feature models, a representation of software products used in software engineering.

9.1.2. Inria Lille - Nord Europe

9.1.2.1. North European Lab LLEX

Participants: Benjamin Danglot, Martin Monperrus, Lionel Seinturier [contact person].

North European Lab LLEX (2017–19) is an international initiative supported by the Inria Lille - Nord Europe Center that takes place in the context of a collaboration between Inria and KTH. LLEX deals with research on automated diagnosis and repair of software bugs. Automated software repair is the process of fixing software bugs automatically. An automated software repair system fixes software bugs with no human intervention. The goal of automated software repair is to save maintenance costs and to enable systems to be more resilient to bugs and unexpected situations. This research may dramatically improve the quality of software systems. This initiative led to several results that have been published [24], [21], [45], [36], [10] and to the PhD thesis of Benajamin Danglot [11] that have been defended in November 2019.

9.1.2.2. ADT FingerKit

Participants: Antoine Canda, Walter Rudametkin Ivey [contact person], Antoine Vastel.

ADT FingerKit (2018–20) is a technology development initiative supported by the Inria Lille - Nord Europe Center that focuses on the design and development of a new and enhanced version of the AmIUnique platform. AmIUnique is a data collection and analysis platform to better understand, analyze and vulgarize the uses and threats of browser fingerprinting. This initiative led by Inria is a key asset to better understand novel techniques that threatens the user privacy on Internet. This ADT builds on our first results with the PhD thesis of Antoine Vastel [14].

9.1.2.3. ADT e-Lens

Participants: Arthur d'Azémar, Guillaume Fieni, Romain Rouvoy [contact person].

ADT e-Lens (2018–20) is a technology development initiative supported by the Inria Lille - Nord Europe Center that aims at extending the PowerAPI energy monitoring library that we develop in the team since 2011. The extension deals with the integration of new power models (for GPU, disk, network interface), the implementation of a self-optimization algorithm, the port of the platform to embedded systems running with Raspberry Pi, ROS and Android, and the implementation of an active learning algorithm for power models. This ADT builds on our results with the defended PhD theses of Adel Noureddine [68] and Maxime Colmant [52], and with the ongoing PhD thesis of Guillaume Fieni.

9.2. National Initiatives

9.2.1. ANR

9.2.1.1. ANR BottleNet

Participants: Romain Rouvoy [contact person], Walter Rudametkin Ivey, Lionel Seinturier.

BottleNet is a 48-month project (2015–19) funded by ANR. The objective of BottleNet is to deliver methods, algorithms, and software systems to measure Internet *Quality of Experience* (QoE) and diagnose the root cause of poor Internet QoE. Our goal calls for tools that run directly at users' devices. We plan to collect network and application performance metrics directly at users' devices and correlate it with user perception to model Internet QoE, and to correlate measurements across users and devices to diagnose poor Internet QoE. This data-driven approach is essential to address the challenging problem of modeling user perception and of diagnosing sources of bottlenecks in complex Internet services. BottleNet will lead to new solutions to assist users, network and service operators as well as regulators in understanding Internet QoE and the sources of performance bottleneck. Several results and publications have been obtained in the context of this project [6], [5], [76]. The paper [5] won the Best Paper award at the 19th International Conference on Distributed Applications and Interoperable Systems (DAIS 2019). The PhD thesis of Lakhdar Meftah [13], supervised by Romain Rouvoy and Isabelle Chrisment (Inria Nancy), was defended in December 2019 in the context of this project. This project is in relation with the Inria IPL BetterNet.

9.2.1.2. ANR SATAS

Participants: Alexandre Garnier, Philippe Merle [contact person], Romain Rouvoy, Lionel Seinturier.

SATAS is a 48-month project (2015–20) funded by ANR. SATAS aims to advance the state of the art in massively parallel SAT solving with a particular eye to the applications driving progress in the field. The final goal of the project is to be able to provide a "pay as you go" interface to SAT solving services, with a particular focus on their power consumption. This project will extend the reach of SAT solving technologies, daily used in many critical and industrial applications, to new application areas, which were previously considered too hard, and lower the cost of deploying massively parallel SAT solvers on the cloud. Our results from this project have been published in the following papers [53], [60].

9.2.1.3. ANR Headwork

Participants: Pierre Bourhis [contact person], Marion Tommasi.

Headwork is a 48-month project (2016–21) funded by ANR. The main objective of Headwork is to develop data-centric workflows for programming crowd sourcing systems in a flexible declarative manner. The problem of crowd sourcing systems is to fill a database with knowledge gathered by thousands or more human participants. A particular focus is to be put on the aspects of data uncertainty and for the representation of user expertise. This project is coordinated by D. Gross-Amblard from the Druid Team (Rennes 1). Other partners include the Dahu team (Inria Saclay), Sumo (Inria Bretagne), and Links (Inria Lille) with J. Nierhen and M. Sakho. Our results from this project have been published in the following paper [1].

9.2.1.4. ANR Delta

Participant: Pierre Bourhis [contact person].

Delta is a 48-month project (2016–21) funded by ANR. The project focuses on the study of logic, transducers and automata. In particular, it aims at extending classical framework to handle input/output, quantities and data. This project is coordinated by M. Zeitoun from LaBRI. Other partners include LIF (Marseille), IRIF (Paris-Diderot), and D. Gallois from the Inria Lille Links team. Several results and publications have been obtained in the context of this project [18], [16], [28], [27].

9.2.1.5. ANR CQFD

Participant: Pierre Bourhis [contact person].

CQFD is a 48-month project (2018–22) funded by ANR. The project focuses on the complex ontological queries over federated heterogeneous data. The project targets to set the foundations, to provide efficient algorithms, and to provide query rewriting oriented evaluation mechanisms, for ontology-mediated query answering over heterogeneous data models. This project is coordinated by Federico Ulliana from Inria Sophia Antipolis. Other partners include LaBRI, Inria Saclay, IRISA, LTCI, and LIG.

9.2.1.6. ANR FP-Locker

Participants: Vikas Mishra, Walter Rudametkin Ivey [contact person].

FP-Locker is a 42-month project (2019–23) funded by ANR in the context of the JCJC program. This project proposes to investigate advanced browser fingerprinting as a configurable authentication mechanism. We argue that it has the potential to be the only authentication mechanism when used in very low-security, public websites; it can be used to block bots and other fraudulent users from otherwise open websites. It also has the potential to be used as a second factor authentication mechanism, or as an additional factor in Multi-Factor Authentication (MFA) schemes. Besides strengthening a session's initial authentication, it can also be used for continuous session authentication to protect against session hijacking. In many contexts, fingerprinting is fully transparent to users, meaning that contrary to authentication processes that rely on external verification cards, code generating keys, special apps, SMS verification codes, users do not have to do anything to improve their security. In more restricted contexts, administrators can enforce different policies, for example, enrolling fingerprints from devices that connect from trusted IP addresses (e.g., an internal network), and then verifying these fingerprints when the same users connect from untrusted IP addresses. Consequently, we plan to design an architecture and implement it to be able to plug the browser fingerprinting authentication process to an existing authentication system.

9.2.1.7. ANR Koala

Participants: Pierre Bourhis, Clément Quinton [contact person].

Koala is a 42-month project (2019–23) funded by ANR in the context of the JCJC program. The project aims to deliver a series of innovative tools, methods and software to deal with the complexity of fog computing environments configurations and adaptations. In particular, we take a step back on the current limitations of existing approaches (e.g., lack of expressiveness and scalability) and address them placing knowledge as a first-class citizen. We plan to tackle configuration issues from a novel perspective in the field of variability management, using recent techniques from the area of knowledge compilation. Specifically, we will investigate the best-suited d-DNNF representation for each reasoning operation, and we plan to provide new variability modeling mechanisms (e.g., dimensions, priorities and scopes) required in a fog context. Regarding adaptation concerns, we want to leverage machine learning techniques to improve adaptation management and evolution under uncertainty, relying on a continuously enriched and reusable knowledge base. In particular, we plan to propose an approach for suggesting evolution scenarios in a predictive manner, relying on an evolution-aware knowledge base acquired at run-time through machine learning feedback.

9.2.2. Competitivity Clusters

9.2.2.1. FUI StoreConnect

Participants: Romain Rouvoy, Lionel Seinturier [contact person].

StoreConnect is a 36-month project (2016–19) funded by FUI and labelled by the PICOM (Pôle des Industries du COMmerce) competitivity cluster. The partners are Tevolys, Ubudu (leader), Smile, STIME, Leroy Merlin, Insiteo, Inria Spirals, Inria Fun, Inria Stars. The goal of the project is to define a modular multi-sensors middleware platform for indoor geolocation. Several results and publications have been obtained in the context of this project [5], [38], [66].

9.2.3. Inria National Initiatives

9.2.3.1. Inria IPL BetterNet

Participants: Lakhdar Meftah, Romain Rouvoy [contact person].

BetterNet (2016–19) 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. Our observatory can be defined as a vantage point, where: (1) tools, models and algorithms/heuristics will be provided to collect data, (2) acquired data will be analyzed, and shared appropriately with scientists, stakeholders and civil society, and (3) new value-added services will be proposed to end-users. IPL BetterNet is led by Isabelle Chrisment (Inria Madynes), with the participation of the Diana, Dionysos, Inria Chile, Muse, and Spirals Inria project-teams, as well as the ARCEP French agency and the ip-label company. Our results in the context of this project have been published in [66].

9.2.4. CNRS Momentum

9.2.4.1. Manage Your Data Without Information Leakage

Participants: Pierre Bourhis [contact person], Louis Jachiet.

"Gérer vos données sans fuite d'information" is a 3-year (2018–20) project granted in the context of the CNRS-Momentum call for projects. Data manipulated by modern applications are stored in large databases. To protect these pieces of data, security policies limit a user's access to what she is allowed to see. However, by using the semantics of the data, a user can deduce information that she was not supposed to have access to. The goal of this project is to establish methods and tools for understanding and detecting such data leaks. Several results and publications have been obtained in the context of this project [32], [29], [30].

9.3. European Initiatives

9.3.1. FP7 & H2020 Projects

Program: H2020 ICT-10-2016.

Project acronym: STAMP.

Project title: Software Testing Amplification.

Duration: 36 months (2016–19).

Coordinator: Inria.

Other partners: ActiveEon (France), Atos (Spain), Engineering (Italy), OW2 (France), SINTEF (Norway), TellU (Norway), TU Delft (The Netherlands), XWiki (France).

Abstract: By leveraging advanced research in automatic test generation, STAMP aims at pushing automation in DevOps one step further through innovative methods of test amplification. It will reuse existing assets (test cases, API descriptions, dependency models), in order to generate more test cases and test configurations each time the application is updated. Acting at all steps of development cycle, STAMP techniques aim at reducing the number and cost of regression bugs at unit level, configuration level and production stage.

Participants: Benjamin Danglot, Martin Monperrus [contact person].

Program: H2020 JU Shift2Rail.

Project acronym: X2Rail-1.

Project title: Start-up activities for Advanced Signalling and Automation System.

Duration: 36 months (2016–19).

Coordinator: Siemens.

Other partners: 19 partners, among others Bombardier, Siemens, Thales, IRT Railenium.

Abstract: Our contribution to the project is focused on adaptive communication middleware for cyber-physical railway systems.

Participants: Lionel Seinturier [contact person].

9.3.2. Collaborations in European Programs, Except FP7 & H2020

Program: EUREKA Celtic-Plus.

Project acronym: SENDATE.

Project title: SEcure Networking for a DATa Center Cloud in Europe.

Duration: 36 months (2016–19).

Coordinator: Nokia.

Other partners: 50+ partners in Finland, France, Germany, Norway, and Sweden. Selected partners involved: Nokia, Orange.

Abstract: The project addresses the convergence of telecommunication networks and IT in the context of distributed data centers. We are involved in the TANDEM subproject that targets the infrastructure of such a distributed system. More specifically, we are studying new approaches in terms of software engineering and component-based solutions for enabling this convergence of network and IT.

Participants: Lionel Seinturier [contact person].

9.4. International Initiatives

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

9.4.1.1. SOMCA

Title: Self-Optimization of Service Oriented Architectures for Mobile and Cloud Applications

International Partner (Institution - Laboratory - Researcher):

Université du Québec à Montréal (Canada) - LATECE - Naouel MOHA

Start year: 2017

See also: http://sofa.uqam.ca/somca.php

The long-term goal of this research program is to propose a novel and innovative methodology embodied in a software platform, to support the runtime detection and correction of anti-patterns in large-scale service-oriented distributed systems in order to continuously optimize their quality of service. One originality of this program lies in the dynamic nature of the service-oriented environments and the application on emerging frameworks for embedded and distributed systems (e.g., Android/iOS for mobile devices, PaaS/SaaS for Cloud environments), and in particular mobile systems interacting with remote services hosted on the Cloud.

9.4.2. Participation in Other International Programs

9.4.2.1. Partnership for joint Curriculum Development and Research in Energy Informatics (PACE)

Participants: Mohammed Chakib Belgaid, Arthur d'Azémar, Guillaume Fieni, Alexandre Garnier, Zakaria Ournani, Clément Quinton, Romain Rouvoy [contact person], Lionel Seinturier.

PACE is a 3-year (2019–21) project funded by the Research Council of Norway. The goal of the project is to establish a sustained education and research-oriented collaboration between four partner universities in energy informatics and green computing that will strengthen quality academic relations and mutually improve each other's quality of research and researcher training both at PhD and master level. Partner universities are: University of Oslo (Norway), University of Stavanger (Norway), TU Munich (Germany), Université de Lille.

9.5. International Research Visitors

9.5.1. Visits of International Scientists

9.5.1.1. Internships

Jonatan Enes, PhD Student in Computer Science from University of A Coruña, visited us for 3 months from April to July.

Alejandro Grez, from Pontifical Catholic University of Chile, visited us for 1 month in April.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

Simon Bliudze

- co-chair International Workshop on Methods and Tools for Rigorous System Design (MeTRiD)

Laurence Duchien

- general chair European Conference on Software Engineering (ECSA)

Walter Rudametkin Ivey

- Atelier sur la Protection de la Vie Privée (APVP)

10.1.1.2. Member of the Organizing Committees

Pierre Bourhis

- Atelier sur la Protection de la Vie Privée (APVP)
- Knowledge Comparison Workshop (KOCOON)

Clément Quinton

- publicity chair International Systems and Software Product Line Conference (SPLC)
- proceedings chair European Conference on Software Engineering (ECSA)

Romain Rouvoy

- poster & demonstration chair International Symposium on Reliable Distributed Systems (SRDS)

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

Laurence Duchien

- co-chair International Systems and Software Product Line Conference (SPLC)

10.1.2.2. Member of the Conference Program Committees

Simon Bliudze

- International Conference on Coordination Models and Languages (Coordination)
- International Conference on Formal Methods in Software Engineering (FormaliSE)
- International Conference on Formal Aspects of Component Software (FACS)
- International Conference on Verification and Evaluation of Computer and Communication Systems (VECoS)
- International Symposium on Model-Based Safety and Assessment (IMBSA)
- Workshop on Formal Approaches for Advanced Computing Systems (FAACS)
- International Workshop on Foundations of Coordination Languages and Self-Adaptative Systems (FOCLASA)
- Embedded Systems and the Internet of Things track at Euromicro Software Engineering and Advanced Applications conference (ES-IoT@SEAA)

Pierre Bourhis

- International Conference on Database Theory (ICDT)
- ACM SIGMOD International Conference on Management of Data (PODS)
- International Joint Conferences on Artificial Intelligence (IJCAI)
- Conférence sur la Gestion de Données – Principes, Technologies et Applications (BDA)

Laurence Duchien

- International Conference on Software Architecture (ICSA)
- International Workshop Series on Conducting Empirical Studies in Industry (CESI) in conjonction with the ICSE conference
- International Workshop on Software Qualities And Their Dependencies (SQADE) in conjunction with the ESEC/FSE conference
- Belgium-Netherlands Software Evolution Workshop

Pierre Laperdrix

- USENIX Workshop on Offensive Technologies (WOOT)

Philippe Merle

- International Conference on Cooperative Information Systems (CoopIS)
- International Conference on Evolving Internet (INTERNET)
- International Conference on Cloud Computing, GRIDs, and Virtualization (CLOUD COMPUTING)
- International Conference on Advanced Service Computing (SERVICE COMPUTATION)
- International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM)
- Workshop on CrossCloud Infrastructures & Platforms (CrossCloud)
- International Workshop on Adaptive and Reflective Middleware (ARM)

Clément Quinton

- International Workshop on Variability Modelling of Software-Intensive Systems (VaMoS)
- ACM Symposium on Applied Computing (SAC) SA-TTA Track
- International Systems and Software Product Line Conference (SPLC)
- SPLC Demonstration and Tools Track
- International Workshop on Software Product Line Teaching (SPLTea)
- Women in Software Architecture (WSA ECSA)

Romain Rouvoy

- International Conference on Service Oriented Computing (ICSOC)
- ACM Symposium on Applied Computing (SAC) DADS Track
- International Conference on Ambient Systems, Networks and Technologies (ANT)
- ACM International Conference on Distributed and Event-based Systems (DEBS)
- International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP)
- International Workshop on Advances in Mobile App Analysis (A-Mobile)
- Eurosys Doctoral Workshop

Lionel Seinturier

- IEEE/ACM International Conference on Mobile Software Engineering and Systems (MOBILESoft)
- ACM Symposium on Applied Computing (SAC) SA-TTA Track
- International Conference on Software Technologies (ICSOFT)
- International Conference on Service Oriented Computing (ICSOC)
- ACM International Conference on Management of Emergent Digital EcoSystems (MEDES)
- International Conference on Computational Collective Intelligence (ICCCI)

Walter Rudametkin Ivey

- International Symposium on Stabilization, Safety, and Security of Distributed Systems (SSS) Track D: Security and Privacy

10.1.3. Journal

10.1.3.1. Member of the Editorial Boards

Lionel Seinturier is editor for software engineering of the ISTE-Wiley Computer Science and Information Technology book collection.

10.1.3.2. Reviewer - Reviewing Activities

Simon Bliudze: Science of Computer Programming (SCICO), Journal of Logical and Algebraic Methods in Programming (JLAMP)

Laurence Duchien: Journal of Systems and Software (JSS).

Pierre Laperdrix: IEEE Access, Emerald Journal of Intellectual Capital.

Philippe Merle: ACM Computing Surveys (CSUR), Elsevier Future Generation Computer Systems (FGCS), IEEE Transactions on Services Computing (TSC), IEEE Access, Wiley Software Practice and Experience (SPE).

Clément Quinton: Elsevier Information and Software Technology (IST), Science of Computer Programming (SCICO).

Romain Rouvoy: Elsevier Information and Software Technology (IST), Elsevier Journal of Systems and Software (JSS), ACM Transactions on Internet of Things (TOIT), ACM Computing Surveys (CSUR), IEEE Transactions on Software Engineering (TSE).

Walter Rudametkin Ivey: ACM Transactions on the Web (TWEB), IEEE Transactions on Services Computing (TSC), IEEE Communications Magazine.

Lionel Seinturier: ACM Transactions on Software Engineering and Methodology (TOSEM), Springer Journal of Internet Services and Applications (JISA), Wiley Journal of Software Evolution and Process (JSME).

10.1.4. Invited Talks

Simon Bliudze gave a tutorial on the Rigorous Component-based Design in BIP at the 6th International Symposium on Model-Based Safety and Assessment (IMBSA) and an invited lecture on Component-Based Design of Concurrent Software in BIP for a group of Master 2 students at the Aristotle University of Thessaloniki (Greece).

Pierre Bourhis gave an invited talk on Reasoning on Leak of Information for Database Views at the Emerging Challenges in Databases and AI Research Seminar in Santa Cruz, Chile.

Romain Rouvoy gave an invited talk on "*Enabling In situ Location Privacy*" during the IRIXYS workshop organized in Lyon in June 2019. He was also a keynote speaker during the national days of the GDR GPL in June 2019 with a talk entitled "*Quels défis pour le développement durable des logiciels ?*".

10.1.5. Leadership within the Scientific Community

Romain Rouvoy is the co-head of the "Groupe de Travail Génie Logiciel pour les Systèmes Cyber-physiques" of the GDR GPL.

10.1.6. Scientific Expertise

Laurence Duchien was member of the Hcéres committee for the UMR Laboratoire d'Informatique de Grenoble (LIG). She was expert for Canada NSERC, and DGRI (Ministère de la Recherche) program PHC. She was chair of the recruitment committee for Professor in Computer Science at ENS Lyon. She was member of the recruitment committee for two associate professor positions at the University of Luxembourg.

Romain Rouvoy was vice-chair of the recruitment committee for research scientists at Inria Lille - Nord Europe. He was member of the recruitment committee for associate researchers at University of Luxembourg. He was member of scientific committee 25 - Infrastructures (networks, high performance computing and storage), software sciences and technologies of the French Research Agency (ANR). He was scientific expert for DRRT IDF.

Lionel Seinturier was chair of the recruitment committee for Professor in Computer Science at CNAM Paris. He was member of the promotion committee to Tenured Associate Professor, Department of Electrical and Computer Engineering, University of Waterloo, Canada. He was member of the IEEE Technical Council on Software Engineering Distinguished Educational Award Committee. He was scientific expert for Chile CONICYT, Belgium Innoviris, DRRT IDF.

10.1.7. Research Administration

Simon Bliudze was member of steering committees of International Interaction and Concurrency Experience workshop (ICE) and the International Symposium on Formal Approaches to Parallel and Distributed Systems (4PAD). He is lead guest editor for a special issue on Verification and Evaluation of Computer and Communication Systems (VECoS) of Innovations in Systems and Software Engineering: A NASA Journal (ISSE, Springer) and guest editor for a special issue on Methods and Tools for Rigorous System Design (MeTRiD) of the International Journal on Software Tools for Technology Transfer (STTT, Springer).

Laurence Duchien is member of the CNRS CoCNRS section 6 committee, and of the "bureau" of this committee. She is chairing the scientific and technical council of the PICOM business and research cluster for the retail industry. She is member of the scientific councils of IMT Atlantique (Nantes), Labex CIMI (Toulouse), IRT SystemX (Saclay). She is member of the steering committees of the European Conference on Software Architecture (ECSA) and of the Systems and Software Product Lines Conference (SPLC). She was guest editor for a special issue on the Configurable Systems of Springer Empirical Software Engineering journal. She was member of the committee for the Women in Science and Engineering Award from IEEE TCSE. She was member of a panel on soft skills for PhD students on humanities and social sciences for MESHS Lille. She is in charge of the Career development & Intersectoral secondments in the PEARL Project ("Programme for EArly-stage Researchers in Lille") at I-SITE Université Lille Nord Europe.

Philippe Merle is elected member of the Inria scientific board (CS), elected member of the Inria technical committee (CTI), elected member of the Inria national committee on "hygiène, de sécurité et des conditions de travail" (CNHSCT), president of the CUMI (Comité des Utilisateurs des Moyens Informatiques), permanent secretary of the CLHSCT (Comité Local d'Hygiène, de Sécurité et de Conditions de Travail), and elected member of the centre committee for the Inria Lille - Nord Europe research center. He is member of the steering committee of the Inria's continuous integration service. He is member of SPI doctoral school council of University of Lille.

Romain Rouvoy is an elected member of the "bureau" of the French chapter of the ACM Special Interest Group in Operating Systems (SIGOPS / ASF) and elected member of the administrative council of SpecifCampus. He is a member of the steering committee of DAIS (International Conference on Distributed Applications and Interoperable Systems).

Walter Rudametkin Ivey is member of the CDT (Comité de Développement Technologique) of the Inria Lille - Nord Europe research center.

Lionel Seinturier is member of the BSC (Bureau Scientifique du Centre) for the Inria Lille - Nord Europe research center. Until August 2019, he was Scientific Advisor for the evaluation of ICT research laboratories at the Hcéres. Since December 2019, he is chair of CNU27, the Computer Science section of the "Conseil National des Universités".

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Simon Bliudze is, in addition to his tenure junior research position at Inria, chargé d'enseignement at École Polytechnique, Palaiseau, France, in the Department of Computer Sciences (DIX).

- INF411: Les bases de la programmation et de l'algorithmique, 40h, 2nd year of the Engineering cycle
- INF442: Traitement des données massives, 40h, 2nd year of the Engineering cycle

Pierre Bourhis is, in addition to his tenure junior research position at CNRS, chargé d'enseignement spécialité Sciences des données at École Polytechnique, Palaiseau, France, in the Department of Computer Sciences (DIX).

- Info553: Bases de données, 18h, Cycle Polytechnique
- Modal Graphe Géant, 36h

Laurence Duchien teaches at the Université de Lille in the FST faculty. She is project leader for doctoral studies at Université de Lille.

- Software engineering project, 60h, Level M2, Master MIAGE FI
- Software engineering project, 50h, Level M2, Master MIAGE FC/FA
- Research initiation, 20h, Level M2, Master of Computer Science

Clément Quinton teaches at the Université de Lille in the FST faculty.

- Introduction to Computer Science, 46.5h, Level L1, Licence of Computer Science
- Object-oriented programming, 36h, Level L2, Licence of Computer Science
- Object-oriented design, 42h, Level L3, Licence of Computer Science
- Design of distributed applications, 42h, Level M1, Master of Computer Science
- Advanced design of distributed applications, 37.5h, Level M2, Master MIAGE
- Infrastructure and frameworks for the Internet, 33.75h, Level M2, Master of Computer Science
- Software product lines, 7.5h, Level M2, Master of Computer Science
- Suivi de stages et de projets, 30h, Licence and Master of Computer Science

Romain Rouvoy teaches at the Université de Lille in the FST faculty. He heads the Master of Computer Science program at the Université de Lille.

- Design of distributed applications, 12h, Level M1, Master of Computer Science
- Object-oriented design, 4h, Level L3, Licence of Computer Science
- Suivi de projets, 20h, Level M2, Master of Computer Science

Walter Rudametkin Ivey teaches at the Université de Lille in the Polytech engineering school.

- GIS4 Programmation par Objets, 32h
- GIS4 Architectures Logicielles, 26h
- GIS2A3 (apprentissage) Projet programmation par Objet, 24h
- IMA2A4 (apprentissage) Conception Modélisation Objet, 24h
- IMA3 Programmation Avancée, 62h
- GBIAAL4 Bases de données, 22h
- GIS5 Suivi de projets, 42h
- GIS2A (apprentissage) Suivi d'apprentis, 28h

Lionel Seinturier teaches at the Université de Lille in the FST faculty. Until July 2019, he headed the Computer Science Department at the Faculty of Science and Technology of the Université de Lille.

- Conception d'Applications Réparties, 50h, Level M1, Master MIAGE
- Infrastructures et Frameworks Internet, 70h, Level M2 E-Services IAGL TIIR, Master of Computer Science

10.2.2. Supervision

- PhD: Antoine Vastel, Cartographie de la qualité d'expérience pour l'accès à l'internet mobile, Université de Lille, October 2019, Romain Rouvoy & Walter Rudametkin. [14]
- PhD: Benjamin Danglot, Automatic Unit Test Amplification for DevOps, Université de Lille, November 2019, Martin Monperrus & Lionel Seinturier. [11]
- PhD: Lakhdar Meftah, Cartography of the Quality of Experience for Mobile Internet Access, Université de Lille, December 2019, Romain Rouvoy, co-supervision with Isabelle Chrisment (Inria Madynes). [13]
- PhD: Sarra Habchi, Une supervision de contexte sensible à la confidentialité pour les développements logiciels en crowdsource, Université de Lille, December 2019, Romain Rouvoy. [12]
- PhD in progress: Mohammed Chakib Belgaid, Développement durable des logiciels vers une optimisation énergétique de bout en bout des systèmes logiciels, January 2018, Romain Rouvoy & Lionel Seinturier.
- PhD in progress: Vikas Mishra, Collaborative Strategies to Protect Against Browser Fingerprinting, October 2018, Romain Rouvoy & Walter Rudametkin & Lionel Seinturier.
- PhD in progress: Marion Tommasi, Collaborative Data-centric Workflows: Towards Knowledge Centric Workflows and Integrating Uncertain Data, October 2018, Pierre Bourhis & Lionel Seinturier.
- PhD in progress: Antonin Durey, Leveraging Browser Fingerprinting to Fitht Fraud on the Web, October 2018, Romain Rouvoy & Walter Rudametkin & Lionel Seinturier.
- PhD in progress: Sacha Brisset, Automatic Spotting and fixing or Recurrent user Experience issues. Detecting and Fixing Anomalies by applying Machine Learning on user Experience Data, November 2018, Lionel Seinturier & Romain Rouvoy & Renaud Pawlak & Yoann Couillec.
- PhD in progress: Zakaria Ournani, Eco-conception des logiciels : modélisation de l'efficience énergétique des logiciels et conception d'outils pour mesurer et réduire leur consommation d'énergie, November 2018, Romain Rouvoy.
- PhD in progress: Zeinab Abou Khalil, November 2017, Laurence Duchien & Clément Quinton, co-supervision with Tom Mens (University of Mons, Belgium).
- PhD in progress: Guillaume Fieni, GreenData : Vers un traitement efficient et éco-responsable des grandes masses de données numériques, October 2017, Romain Rouvoy & Lionel Seinturier.
- PhD in progress: Trình Lê Khánh, Design of Correct-by-Construction Self-Adaptive Cloud Applications using Formal Methods, October 2019, Philippe Merle & Simon Bliudze.

10.2.3. Juries

Simon Bliudze

- Wajeb Saab (École polytechnique fédérale de Lausanne, Switzerland), examiner
- Rim El Ballouli (Université Grenoble Alpes), examiner

Laurence Duchien

- Stepan Shevtsov (Linnaeus University, Sweden and KU Leuven, Belgium), reviewer
- HDR, Marius Bilasco (Université de Lille), chair
- Thibault Raffaillac (Université de Lille), chair

Romain Rouvoy

- Hanyang Cao (Université de Bordeaux), reviewer
- Mousa Hayam (INSA Lyon), reviewer
- Stefan Contiu (Université de Bordeaux), reviewer
- Adrien Luxey (Université de Rennes 1), reviewer

Lionel Seinturier

- HDR, Thomas Polacsek (Université de Toulouse), reviewer
- Xinxiu Tao (Université Grenoble Alpes), reviewer
- Benjamin Benni (Université Nice Sophia Antipolis), reviewer

10.3. Popularization

10.3.1. Interventions

Pierre Laperdrix gave a talk in May at the RuhrSec IT security conference organized by the Ruhr University Bochum. The aim of the conference is to popularize different topics to an audience of non-experts so that they are familiar with the latest trends in computer security.

11. Bibliography

Major publications by the team in recent years

[1] S. Abiteboul, P. Bourhis, V. Vianu.*Explanations and Transparency in Collaborative Workflows*, in "PODS 2018 - 37th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles Of Database Systems", Houston, Texas, United States, June 2018, https://hal.inria.fr/hal-01744978

[2] Y. Al-Dhuraibi, F. Paraiso, N. Djarallah, P. Merle.*Elasticity in Cloud Computing: State of the Art and Research Challenges*, in "IEEE Transactions on Services Computing (TSC)", March 2018, vol. 11, n^{o} 2, p. 430-447 [*DOI :* 10.1109/TSC.2017.2711009], https://hal.inria.fr/hal-01529654

[3] S. Bliudze, S. Furic, J. Sifakis, A. Viel.*Rigorous Design of Cyber-Physical Systems*, in "Software and Systems Modeling", 2019, vol. 18, n^{o} 3, p. 1613–1636 [*DOI :* 10.1007/s10270-017-0642-5], https://hal.inria.fr/hal-01636392

[4] S. Habchi, X. Blanc, R. Rouvoy.*On Adopting Linters to Deal with Performance Concerns in Android Apps*, in "ASE18 - Proceedings of the 33rd IEEE/ACM International Conference on Automated Software Engineering", Montpellier, France, ACM Press, September 2018, vol. 11 [*DOI :* 10.1145/3238147.3238197], https://hal.inria.fr/hal-01829135

[5] L. Meftah, R. Rouvoy, I. Chrisment.*FOUGERE: User-Centric Location Privacy in Mobile Crowdsourcing Apps*, in "19th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS)", Kongens Lyngby, Denmark, J. Pereira, L. Ricci (editors), Distributed Applications and Interoperable Systems, Springer International Publishing, 2019, vol. LNCS-11534, p. 116-132 [*DOI :* 10.1007/978-3-030-22496-7_8], https://hal.inria.fr/hal-02121311

[6] M. NASERI, N. P. BORGES JR., A. ZELLER, R. ROUVOY.*AccessiLeaks: Investigating Privacy Leaks Exposed by the Android Accessibility Service*, in "PETS 2019 - The 19th Privacy Enhancing Technologies Symposium", Stockholm, Sweden, July 2019, https://hal.inria.fr/hal-01929049

[7] E. STACHTIARI, A. MAVRIDOU, P. KATSAROS, S. BLIUDZE, J. SIFAKIS.*Early validation of system requirements and design through correctness-by-construction*, in "Journal of Systems and Software", 2018, vol. 145, p. 52-78 [*DOI :* 10.1016/J.JSS.2018.07.053], https://hal.archives-ouvertes.fr/hal-01873999

[8] S. URLI, Z. YU, L. SEINTURIER, M. MONPERRUS.*How to Design a Program Repair Bot? Insights from the Repairnator Project*, in "40th International Conference on Software Engineering, Track Software Engineering in Practice", Gothenburg, Sweden, May 2018, p. 95-104, update for oadoi on Nov 02 2018 [*DOI :* 10.1145/3183519.3183540], https://hal.inria.fr/hal-01691496

[9] A. VASTEL, P. LAPERDRIX, W. RUDAMETKIN, R. ROUVOY.*FP-STALKER: Tracking Browser Fingerprint Evolutions*, in "IEEE S&P 2018 - 39th IEEE Symposium on Security and Privacy", San Francisco, United States, B. PARNO, C. KRUEGEL (editors), Proceedings of the 39th IEEE Symposium on Security and Privacy (S&P), IEEE, May 2018, p. 728-741 [*DOI :* 10.1109/SP.2018.00008], https://hal.inria.fr/hal-01652021

[10] Z. YU, C. BAI, L. SEINTURIER, M. MONPERRUS.*Characterizing the Usage, Evolution and Impact of Java Annotations in Practice*, in "IEEE Transactions on Software Engineering", 2019 [*DOI :* 10.1109/TSE.2019.2910516], https://hal.inria.fr/hal-02091516

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] B. DANGLOT.*Automatic Unit Test Amplification For DevOps*, Université de Lille, November 2019, https://tel.archives-ouvertes.fr/tel-02396530

[12] S. HABCHI.*Understanding Mobile-Specific Code Smells*, Université de Lille, December 2019, https://tel.archives-ouvertes.fr/tel-02414928

[13] L. MEFTAH.*Towards Privacy-sensitive Mobile Crowdsourcing*, University of Lille, December 2019, https://tel.archives-ouvertes.fr/tel-02399716

[14] A. VASTEL.*Tracking Versus Security: Investigating the Two Facets of Browser Fingerprinting*, Université de Lille Nord de France, October 2019, https://tel.archives-ouvertes.fr/tel-02343930

Articles in International Peer-Reviewed Journal

[15] Y. AL-DHURAIBI, F. ZALILA, N. DJARALLAH, P. MERLE.*Model-Driven Elasticity Management with OCCI*, in "IEEE transactions on cloud computing", 2019, vol. 99, 1, forthcoming [*DOI :* 10.1109/TCC.2019.2923686], https://hal.inria.fr/hal-02375362

[16] A. AMARILLI, P. BOURHIS, M. MONET, P. SENELLART.*Evaluating Datalog via Tree Automata and Cycluits*, in "Theory of Computing Systems", 2019, vol. 63, n^{o} 7, p. 1620-1678, https://arxiv.org/abs/1808.04663 [*DOI :* 10.1007/s00224-018-9901-2], https://hal.inria.fr/hal-01891811

[17] E. BARANOV, S. BLIUDZE.*Expressiveness of component-based frameworks: A study of the expressiveness of BIP*, in "Acta Informatica", 2019, forthcoming [*DOI :* 10.1007/S00236-019-00337-7], https://hal.archives-ouvertes.fr/hal-02149934

[18] M. BENEDIKT, P. BOURHIS, G. GOTTLOB, P. SENELLART.*Monadic Datalog, Tree Validity, and Limited Access Containment*, in "ACM Transactions on Computational Logic", October 2019, vol. 21, n^o 1, p. 6:1-6:45 [*DOI :* 10.1145/3344514], https://hal.inria.fr/hal-02307999

[19] S. BLIUDZE, S. FURIC, J. SIFAKIS, A. VIEL.*Rigorous Design of Cyber-Physical Systems : Linking Physicality and Computation*, in "Software and Systems Modeling", 2019, vol. 18, n^o 3, p. 1613–1636, forthcoming [*DOI :* 10.1007/S10270-017-0642-5], https://hal.inria.fr/hal-01636392

[20] H. BRABRA, A. MTIBAA, F. PETRILLO, P. MERLE, L. SLIMAN, N. MOHA, W. GAALOUL, Y.-G. GUEHENEUC, B. BENATALLAH, F. GARGOURI.*On semantic detection of cloud API (anti)patterns*, in "Information and Software Technology", March 2019, vol. 107, p. 65 - 82 [*DOI :* 10.1016/J.INFSOF.2018.10.012], https://hal.inria.fr/hal-02375380

[21] B. DANGLOT, O. L. VERA-PÉREZ, B. BAUDRY, M. MONPERRUS.*Automatic Test Improvement with DSpot: a Study with Ten Mature Open-Source Projects*, in "Empirical Software Engineering", 2019, p. 1-35 [*DOI :* 10.1007/S10664-019-09692-Y], https://hal.inria.fr/hal-01923575

[22] B. DANGLOT, O. L. VERA-PÉREZ, Z. YU, A. ZAIDMAN, M. MONPERRUS, B. BAUDRY.*A Snowballing Literature Study on Test Amplification*, in "Journal of Systems and Software", August 2019, vol. 157, p. 1-16 [*DOI :* 10.1016/J.JSS.2019.110398], https://hal.inria.fr/hal-02290742

[23] T. HARTMANN, F. FOUQUET, A. MOAWAD, R. ROUVOY, Y. LE TRAON.*GreyCat: Efficient What-If Analytics for Data in Motion at Scale*, in "Information Systems", 2019, vol. 83, p. 101-117, forthcoming [*DOI :* 10.1016/J.IS.2019.03.004], https://hal.inria.fr/hal-02059882

[24] M. MONPERRUS, S. URLI, T. DURIEUX, M. MARTINEZ, B. BAUDRY, L. SEINTURIER.*Repairnator patches programs automatically*, in "Ubiquity", July 2019, vol. July, n^o 2, p. 1-12, https://arxiv.org/abs/1910.06247 [*DOI :* 10.1145/3349589], https://hal.inria.fr/hal-02267512

[25] Z. YU, C. BAI, L. SEINTURIER, M. MONPERRUS.*Characterizing the Usage, Evolution and Impact of Java Annotations in Practice*, in "IEEE Transactions on Software Engineering", 2019, forthcoming [*DOI :* 10.1109/TSE.2019.2910516], https://hal.inria.fr/hal-02091516

[26] F. ZALILA, S. CHALLITA, P. MERLE.*Model-Driven Cloud Resource Management with OCCIware*, in "Future Generation Computer Systems", October 2019, vol. 99, p. 260 - 277 [*DOI :* 10.1016/J.FUTURE.2019.04.015], https://hal.inria.fr/hal-02375378

International Conferences with Proceedings

[27] A. AMARILLI, P. BOURHIS, S. MENGEL, M. NIEWERTH.*Constant-Delay Enumeration for Nondeterministic Document Spanners*, in "ICDT", Lisbon, Portugal, 22nd International Conference on Database Theory (ICDT 2019), March 2019, https://arxiv.org/abs/1807.09320 - 25 pages including 17 pages of main material. Integrates all reviewer feedback. Outside of possible minor formatting differences, this paper is exactly the same as the ICDT'19 paper except that it contains 6 pages of technical appendix [*DOI :* 10.4230/LIPICS.ICDT.2019.19], https://hal.telecom-paristech.fr/hal-02338344

[28] A. AMARILLI, P. BOURHIS, S. MENGEL, M. NIEWERTH.*Enumeration on Trees with Tractable Combined Complexity and Efficient Updates*, in "PODS", Amsterdam, France, Proceedings of the 38th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, ACM Press, June 2019, p. 89-103, https://arxiv.org/abs/1812.09519 [*DOI :* 10.1145/3294052.3319702], https://hal.archives-ouvertes.fr/hal-02326243

[29] M. BENEDIKT, P. BOURHIS, L. JACHIET, M. THOMAZO.*Reasoning about disclosure in data integration in the presence of source constraints*, in "IJCAI 2019 - 28th International Joint Conference on Artificial Intelligence", Macao, China, August 2019, https://arxiv.org/abs/1906.00624 , https://hal.inria.fr/hal-02145369

[30] M. BIENVENU, P. BOURHIS.*Mixed-World Reasoning with Existential Rules under Active-Domain Semantics*, in "Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI)", Macao, Macau SAR China, August 2019 [*DOI :* 10.24963/IJCAI.2019/216], https://hal.archives-ouvertes.fr/hal-02342129

[31] 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

[32] P. BOURHIS, M. LECLÈRE, M.-L. MUGNIER, S. TISON, F. ULLIANA, L. GALOIS.*Oblivious and Semi-Oblivious Boundedness for Existential Rules*, in "IJCAI 2019 - International Joint Conference on Artificial Intelligence", Macao, China, August 2019, https://hal-lirmm.ccsd.cnrs.fr/lirmm-02148142

[33] S. HABCHI, N. MOHA, R. ROUVOY.*The Rise of Android Code Smells: Who Is to Blame?*, in "MSR 2019 - Proceedings of the 16th International Conference on Mining Software Repositories", Montréal, Canada, B. ADAMS, S. HAIDUC (editors), May 2019, https://hal.inria.fr/hal-02054788

[34] S. HABCHI, R. ROUVOY, N. MOHA.*On the Survival of Android Code Smells in the Wild*, in "MOBILESoft 2019 - 6th IEEE/ACM International Conference on Mobile Software Engineering and Systems", Montréal, Canada, D. POSHYVANYK, I. MALAVOLTA (editors), May 2019, https://hal.inria.fr/hal-02059097

[35] J. LEVEAU, X. BLANC, L. RÉVEILLÈRE, J.-R. FALLERI, R. ROUVOY.*Fostering the Diversity of Exploratory Testing in Web Applications*, in "IEEE International Conference on Software Testing, Verification and Validation (ICST)", Porto, Portugal, C. PASAREANU, A. ZELLER (editors), March 2020, https://hal.inria.fr/hal-02398969

[36] F. MADEIRAL, S. URLI, M. MAIA, M. MONPERRUS.*Bears: An Extensible Java Bug Benchmark for Automatic Program Repair Studies*, in "SANER 2019 - 26th IEEE International Conference on Software Analysis, Evolution and Reengineering", Hangzhou, China, February 2019, https://arxiv.org/abs/1901.06024 - Proceedings of the 26th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER '19), https://hal.inria.fr/hal-01990052

[37] *Best Paper*
L. MEFTAH, R. ROUVOY, I. CHRISMENT.*FOUGERE: User-Centric Location Privacy in Mobile Crowdsourcing Apps*, in "19th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS)", Kongens Lyngby, Denmark, J. PEREIRA, L. RICCI (editors), Distributed Applications and Interoperable Systems, Springer International Publishing, 2019, vol. LNCS-11534, p. 116-132 [*DOI :* 10.1007/978-3-030-22496-7_8], https://hal.inria.fr/hal-02121311.

[38] L. MEFTAH, R. ROUVOY, I. CHRISMENT.*Testing Nearby Peer-to-Peer Mobile Apps at Large*, in "MOBILESoft 2019 - 6th IEEE/ACM International Conference on Mobile Software Engineering and Systems", Montréal, Canada, D. POSHYVANYK, I. MALAVOLTA (editors), May 2019, https://hal.inria.fr/hal-02059088

[39] *Best Paper*
P. MERLE, A. NDEYE SYLLA, M. OUZZIF, F. KLAMM, K. GUILLOUARD.*A Lightweight Toolchain to Validate, Visualize, Analyze, and Deploy ETSI NFV Topologies*, in "NetSoft 2019 - The 5th IEEE International Conference on Network Softwarization", Paris, France, June 2019 [*DOI :* 10.1109/NETSOFT.2019.8806632], https://hal.archives-ouvertes.fr/hal-02124164.

[40] V. MISHRA, P. LAPERDRIX, A. VASTEL, W. RUDAMETKIN, R. ROUVOY, M. LOPATKA.*Don't count me out: On the relevance of IP addresses in the tracking ecosystem*, in "The Web Conference", Tapeï, Taiwan, B. CRISPO, N. NIKIFORAKIS (editors), Proceedings of The Web Conference (WWW'20), April 2020, vol. Security, Privacy, and Trust, https://hal.inria.fr/hal-02435622

[41] M. NASERI, N. P. BORGES JR., A. ZELLER, R. ROUVOY.*AccessiLeaks: Investigating Privacy Leaks Exposed by the Android Accessibility Service*, in "PETS 2019 - The 19th Privacy Enhancing Technologies Symposium", Stockholm, Sweden, July 2019, https://hal.inria.fr/hal-01929049

[42] Z. OURNANI, M. C. BELGAID, R. ROUVOY, P. RUST, J. PENHOAT, L. SEINTURIER.*Taming Energy Consumption Variations in Systems Benchmarking*, in "International Conference on Performance Engineering (ICPE)", Edmonton, Canada, C. TRUBIANI, A. IOSUP (editors), Proceedings of the 11th ACM/SPEC International Conference on Performance Engineering (ICPE'20), April 2020, https://hal.inria.fr/hal-02403379

Conferences without Proceedings

[43] Z. ABOU KHALIL, E. CONSTANTINOU, T. MENS, L. DUCHIEN, C. QUINTON.*A Longitudinal Analysis of Bug Handling Across Eclipse Releases*, in "ICSME 2019 - 35th IEEE International Conference on Software Maintenance and Evolution", Cleveland, United States, September 2019, https://hal.archives-ouvertes.fr/hal-02179172

[44] 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

[45] M. WHITE, M. TUFANO, M. MARTINEZ, M. MONPERRUS, D. POSHYVANYK.*Sorting and Transforming Program Repair Ingredients via Deep Learning Code Similarities*, in "SANER 2019 - 26th IEEE International

Conference on Software Analysis, Evolution and Reengineering", Hangzhou, China, February 2019, https://arxiv.org/abs/1707.04742 , https://hal.archives-ouvertes.fr/hal-01581170

Scientific Books (or Scientific Book chapters)

[46] *SPLC '19: Proceedings of the 23rd International Systems and Software Product Line Conference - Volume A*, 2019, https://hal.archives-ouvertes.fr/hal-02386136

[47] *Proceedings of the 23rd International Systems and Software Product Line Conference - Volume B (SPLC) : Paris, France — September 09 - 13, 2019*, 2019, https://hal.archives-ouvertes.fr/hal-02386175

[48] *Proceedings of the 13th European Conference on Software Architecture, ECSA 2019, Companion Proceedings (Proceedings Volume 2),*, ACM Press, Paris, France, September 2019 [*DOI :* 10.1145/3344948], https://hal.archives-ouvertes.fr/hal-02378574

Books or Proceedings Editing

[49] T. BURES, L. DUCHIEN, P. INVERARDI (editors). *13th European Conference on Software Architecture, ECSA 2019*, Lecture Notes in Computer Science, Springer, September 2019, n° 11681 [*DOI :* 10.1007/978-3-030-29983-5], https://hal.archives-ouvertes.fr/hal-02378494

Other Publications

[50] P. BOURHIS, L. HÉLOUËT, R. SINGH, Z. MIKLÓS.*Data Centric Workflows for Crowdsourcing*, September 2019, working paper or preprint, https://hal.inria.fr/hal-01976280

References in notes

[51] L. CAPRA, W. EMMERICH, C. MASCOLO.*A micro-economic approach to conflict resolution in mobile computing*, in "SIGSOFT FSE", 2002, p. 31-40

[52] M. COLMANT.*Multi-Dimensional Analysis of Software Power Consumptions in Multi-Core Architectures*, Université Lille 1 - Sciences et Technologies, November 2016, https://tel.archives-ouvertes.fr/tel-01403559

[53] M. COLMANT, P. FELBER, R. ROUVOY, L. SEINTURIER.*WattsKit: Software-Defined Power Monitoring of Distributed Systems*, in "17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)", Madrid, Spain, F. CAPELLO, G. FOX, J. GARCIA-BLAS (editors), Proceedings of the 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid), IEEE, January 2017, 10, https://hal.inria.fr/hal-01439889

[54] V. DALLMEIER, A. ZELLER, B. MEYER.*Generating Fixes from Object Behavior Anomalies*, in "Proceedings of the 2009 IEEE/ACM International Conference on Automated Software Engineering", Washington, DC, USA, ASE '09, IEEE Computer Society, 2009, p. 550–554, http://dx.doi.org/10.1109/ASE.2009.15

[55] T. DURIEUX.*From Runtime Failures to Patches: Study of Patch Generation in Production*, University of Lille, September 2018, https://tel.archives-ouvertes.fr/tel-01883676

[56] F. FOUQUET, E. DAUBERT, N. PLOUZEAU, O. BARAIS, J. BOURCIER, J.-M. JÉZÉQUEL.*Dissemination of Reconfiguration Policies on Mesh Networks*, in "Proceedings of 12th IFIP International Conference on

Distributed Applications and Interoperable Systems (DAIS'12)", Stockholm, Sweden, K. M. GÖSCHKA, S. HARIDI (editors), Lecture Notes in Computer Science, Springer, June 2012, vol. 7272, p. 16-30

[57] M. GOLDMANN, G. KREITZ.*Measurements on the spotify peer-assisted music-on-demand streaming system*, in "Peer-to-Peer Computing", T. ASAMI, T. HIGASHINO (editors), IEEE, 2011, p. 206-211

[58] M. HALL, E. FRANK, G. HOLMES, B. PFAHRINGER, P. REUTEMANN, I. H. WITTEN.*The WEKA data mining software: an update*, in "SIGKDD Explorer Newsletter", 2009, vol. 11, n^{o} 1, p. 10–18

[59] M. HARMAN.*Automated patching techniques: the fix is in*, in "Communications of the ACM", May 2010, vol. 53, n^{o} 5, p. 108–108, http://doi.acm.org/10.1145/1735223.1735248

[60] A. HAVET, V. SCHIAVONI, P. FELBER, M. COLMANT, R. ROUVOY, C. FETZER.*GenPack: A Generational Scheduler for Cloud Data Centers*, in "5th IEEE International Conference on Cloud Engineering (IC2E)", Vancouver, Canada, I. GUPTA, J. LIU (editors), Proceedings of the 5th IEEE International Conference on Cloud Engineering (IC2E), IEEE, April 2017, 10, https://hal.inria.fr/hal-01403486

[61] V. ISSARNY, N. GEORGANTAS, S. HACHEM, A. ZARRAS, P. VASSILIADIS, M. AUTILI, M. A. GEROSA, A. BEN HAMIDA.*Service-Oriented Middleware for the Future Internet: State of the Art and Research Directions*, in "Journal of Internet Services and Applications", May 2011, vol. 2, n^{o} 1, p. 23-45, http://dx.doi.org/10.1007/s13174-011-0021-3

[62] D. JACKSON.*Alloy: A Lightweight Object Modelling Notation*, in "ACM Transactions on Software Engineering and Methodology", April 2002, vol. 11, n^{o} 2, p. 256-290

[63] P. LAPERDRIX, W. RUDAMETKIN, B. BAUDRY.*Beauty and the Beast: Diverting modern web browsers to build unique browser fingerprints*, in "37th IEEE Symposium on Security and Privacy (S&P 2016)", San Jose, United States, May 2016, https://hal.inria.fr/hal-01285470

[64] W. LI, S. HU, J. LI, H.-A. JACOBSEN.*Community Clustering for Distributed Publish/Subscribe Systems*, in "CLUSTER", IEEE, 2012, p. 81-89

[65] M. LÉGER, T. LEDOUX, T. COUPAYE.*Reliable Dynamic Reconfigurations in a Reflective Component Model*, in "CBSE", L. GRUNSKE, R. REUSSNER, F. PLASIL (editors), Lecture Notes in Computer Science, Springer, 2010, vol. 6092, p. 74-92

[66] L. MEFTAH, M. GOMEZ, R. ROUVOY, I. CHRISMENT.*AndroFleet: Testing WiFi Peer-to-Peer Mobile Apps in the Large*, in "ASE 2017 - 32nd IEEE/ACM International Conference on Automated Software Engineering ", Urbana-Champaign, Illinois, United States, ASE 2017 - The 32nd IEEE/ACM International Conference on Automated Software Engineering - Tool demonstration, October 2017, https://hal.inria.fr/hal-01574466

[67] P. MERLE, J.-B. STEFANI.*A formal specification of the Fractal component model in Alloy*, Inria, November 2008, n^{o} 6721, 44, http://hal.inria.fr/inria-00338987

[68] A. NOUREDDINE.*Towards a Better Understanding of the Energy Consumption of Software Systems*, Université des Sciences et Technologie de Lille - Lille I, March 2014, https://tel.archives-ouvertes.fr/tel-00961346

[69] F. PARAISO, P. MERLE, L. SEINTURIER.*soCloud: A service-oriented component-based PaaS for managing portability, provisioning, elasticity, and high availability across multiple clouds*, in "Springer Computing", May 2016, vol. 98, n^{o} 5, p. 539-565 [*DOI :* 10.1007/S00607-014-0421-X], https://hal.inria.fr/hal-01019385

[70] L. PROVENSI, F. ELIASSEN, R. VITENBERG, R. ROUVOY.*Improving Context Interpretation by Using Fuzzy Policies: The Case of Adaptive Video Streaming*, in "28th ACM Symposium on Applied Computing (SAC) - 8th Track on Dependable and Adaptive Distributed Systems (DADS)", Coimbra, Portugal, K. M. GÖSCHKA, R. OLIVEIRA, P. PIETZUCH, G. RUSSELLO (editors), ACM, March 2013, vol. 1, p. 415-422, Best paper award, http://hal.inria.fr/hal-00799136

[71] M. RINARD, C. CADAR, D. DUMITRAN, D. ROY, T. LEU, W. BEEBEE JR.*Enhancing server availability and security through failure-oblivious computing*, in "Proceedings of the 6th conference on Symposium on Operating Systems Design & Implementation (OSDI'06)", USENIX Association, 2004, p. 21–21

[72] M. SALEHIE, L. TAHVILDARI.*Self-adaptive software: Landscape and research challenges*, in "ACM Transactions on Autonomous and Adaptive Systems", May 2009, vol. 4, n^{o} 2, p. 14:1–14:42, http://doi.acm.org/10.1145/1516533.1516538

[73] H. SCHAFFERS, M. SHARPE.*Services in the Future Internet*, April 2011, FP8 Expert Group, EU Directorate-General Information Society & Media, http://cordis.europa.eu/fp7/ict/ssai/docs/softwareconsultationreportfeb2011.pdf

[74] L. SCHUBERT, K. JEFFERY, B. NEIDECKER-LUTZ.*A Roadmap for Advanced Cloud Technologies under H2020*, Dec 2012, http://cordis.europa.eu/fp7/ict/ssai/docs/cloud-expert-group/roadmap-dec2012-vfinal.pdf

[75] L. SEINTURIER, P. MERLE, R. ROUVOY, D. ROMERO, V. SCHIAVONI, J.-B. STEFANI.*A Component-Based Middleware Platform for Reconfigurable Service-Oriented Architectures*, in "Software: Practice and Experience", May 2012, vol. 42, n^{o} 5, p. 559-583 [*DOI :* 10.1002/SPE.1077], http://hal.inria.fr/inria-00567442

[76] R. SOMMERARD, R. ROUVOY.*Towards Privacy-Preserving Data Dissemination in Crowd-Sensing Middleware Platform*, in "11èmes journées francophones Mobilité et Ubiquité (UbiMob'16)", Lorient, France, July 2016, 6, https://hal.inria.fr/hal-01332588

[77] J. URBAN.*NESSI Research Priorities for the next Framework Programme for Research and Technological Development FP8*, May 2011, http://www.nessi-europe.eu/files/Docs/NESSI%20SRA_update_May_2011_V1-0.pdf

[78] Y. WEI, Y. PEI, C. A. FURIA, L. S. SILVA, S. BUCHHOLZ, B. MEYER, A. ZELLER.*Automated fixing of programs with contracts*, in "Proceedings of the 19th international symposium on Software testing and analysis", New York, NY, USA, ISSTA '10, ACM, 2010, p. 61–72, http://doi.acm.org/10.1145/1831708.1831716

[79] C. WEISS, R. PREMRAJ, T. ZIMMERMANN, A. ZELLER.*How long will it take to fix this bug?*, in "Proceedings of the Fourth International Workshop on Mining Software Repositories", IEEE Computer Society, 2007, 1

[80] T. XIE, S. THUMMALAPENTA, D. LO, C. LIU.*Data Mining for Software Engineering*, in "IEEE Computer", August 2009, vol. 42, n^{o} 8, p. 55-62, http://doi.ieeecomputersociety.org/10.1109/MC.2009.256

[81] W. XU, L. HUANG, A. FOX, D. A. PATTERSON, M. I. JORDAN.*Detecting large-scale system problems by mining console logs*, in "Proceedings of the 22nd ACM Symposium on Operating Systems Principles (SOSP'09)", Big Sky, Montana, USA, J. N. MATTHEWS, T. E. ANDERSON (editors), ACM, October 2009, p. 117-132

[82] R. DE LEMOS, H. GIESE, H. A. MULLER, M. SHAW, J. ANDERSSON, L. BARESI, B. BECKER, N. BENCOMO, Y. BRUN, B. CUKIC, R. DESMARAIS, S. DUSTDAR, G. ENGELS, K. GEIHS, K. M. GÖSCHKA, A. GORLA, V. GRASSI, P. INVERARDI, G. KARSAI, J. KRAMER, M. LITOIU, A. LOPES, J. MAGEE, S. MALEK, S. MANKOVSKII, R. MIRANDOLA, J. MYLOPOULOS, O. NIERSTRASZ, M. PEZZÈ, C. PREHOFE, W. SCHÄFER, R. SCHLICHTING, B. SCHMERL, D. B. SMITH, J. P. SOUSA, G. TAMURA, L. TAHVILDARI, N. M. VILLEGAS, T. VOGEL, D. WEYNS, K. WONG, J. WUTTKE.*Software Engineering for Self-Adaptive Systems: A Second Research Roadmap*, in "Software Engineering for Self-Adaptive Systems", R. DE LEMOS, H. GIESE, H. MÜLLER, M. SHAW (editors), Dagstuhl Seminar Proceedings, Springer, 2013, vol. 7475, p. 1-26, http://hal.inria.fr/inria-00638157

Project-Team VALSE

Finite-time control and estimation for distributed systems

IN COLLABORATION WITH: Centre de Recherche en Informatique, Signal et Automatique de Lille

IN PARTNERSHIP WITH:
Ecole Centrale de Lille

Université de Lille

RESEARCH CENTER
Lille - Nord Europe

THEME
Optimization and control of dynamic systems

Table of contents

Project-Team VALSE

Creation of the Project-Team: 2019 November 01

Valse is a dance in triple time: finite-time/fixed-time/hyperexponential.

Keywords:

Computer Science and Digital Science:

A5.9.2. - Estimation, modeling
A6.4.1. - Deterministic control
A6.4.4. - Stability and Stabilization
A6.4.5. - Control of distributed parameter systems
A9.5. - Robotics

Other Research Topics and Application Domains:

B1.1.8. - Mathematical biology
B2.1. - Well being
B5.6. - Robotic systems
B7.2.1. - Smart vehicles

1. Team, Visitors, External Collaborators

Research Scientists

Denis Efimov [Team leader, Inria, Researcher, HDR]
Andrey Polyakov [Inria, Researcher, HDR]
Rosane Ushirobira [Inria, Researcher, HDR]

Faculty Member

Jean-Pierre Richard [École Centrale de Lille, Professor, HDR]

External Collaborators

Leonid Fridman [Inria International Chair, HDR]
Gerald Dherbomez [CNRS Research Engineer]

Technical Staff

Fiodar Hancharou [Inria, from Nov 2019]

PhD Students

Youness Braidiz [École Centrale de Lille, from Feb 2019]
Nelson de Figueiredo Barroso [Inria]
Deesh Dileep [École Centrale de Lille & KUL (Belgium), from Mar 2019]
Alex Dos Reis de Souza [Inria]
Anatolii Khalin [Inria, from Nov 2019]
Tatiana Kharkovskaia [École Centrale de Lille & ITMO University (Russia)]
Wenjie Mei [Université de Lille & CSC (China), from Oct 2019]
Artem Nekhoroshikh [École Centrale de Lille & ITMO University (Russia), from Sep 2019]
Haik Jan Silm [École Centrale de Lille & KUL (Belgium)]
Jijju Thomas [École Centrale de Lille & TUe (The Netherlands)]
Quentin Voortman [École Centrale de Lille & TUe (The Netherlands)]
Siyuan Wang [École Centrale de Lille & CSC (China)]
Yue Wang [École Centrale de Lille & CSC (China), until Feb 2019]

Post-Doctoral Fellows

Cyrille Chenavier [Inria]
Nicolas Espitia [Inria, until Sep 2019]
Tonametl Sanchez Ramirez [Inria, until Sep 2019]

Visiting Scientists

Aleksandr Aleksandrov [SPbSU (Russia), from Mar 2019 until Apr 2019]
Mariana Ballesteros Escamilla [Cinvestav (Mexico), until Apr 2019]
David Cruz Ortiz [Cinvestav (Mexico), until Apr 2019]
Jose Franco Jaramillo [Technological Institute of La Laguna (Mexico), from Oct 2019 until Nov 2019]
Emilia Fridman [Tel Aviv University (Israel), from Jul 2019 until Sep 2019]
Jesus Mendoza Avila [UNAM (Mexico), from Sep 2019 until Dec 2019]
Jaime Moreno Perez [UNAM (Mexico), from Dec 2019]
Junfeng Zhang [HDU (China), from Aug 2019]
Konstantin Zimenko [ITMO University (Russia), from Sep 2019 until Nov 2019]
Stanislav Aranovskiy [Centrale Supelec Rennes, Jun 2019]
Alexander Medvedev [Upsalla University (Sweden), from Nov 2019 until Dec 2019]
Sergiy Zhuk [IBM Research (Ireland), from Dec 2019]
Jhonatan Epperlein [IBM Research (Ireland), from Dec 2019]
Youri Orlov [CICESE (Mexico), from Dec 2019]
Michael Ruderman [University of Agder (Norway), from Dec 2019]

Administrative Assistant

Nathalie Bonte [Inria]

2. Overall Objectives

2.1. Overall Objectives

Valse team studies the estimation and control problems arising in the analysis and the design of distributed, uncertain and interconnected *dynamical* systems:

- Using the concepts of *finite-time/fixed-time/hyperexponential* convergence and stability, the main idea is to separate and hierarchize in time the control and estimation processes, which are distributed in space. This greatly simplifies their analysis and the design for large-scale solutions.
- The main areas of investigation and application are *Internet of things* and *cyber-physical systems*.
- The team aims to draw up algorithms for *decentralized* finite-time control and estimation. The methodology to be developed includes extensions of the theory of *homogeneous* systems and of finite-time/fixed-time/hyperexponential convergence and stability notions. A particular attention is given to applications in real-world scenarios.
- It is a joint proposal with the CNRS CRIStAL UMR 9189 (Centrale Lille and the University of Lille).

3. Research Program

3.1. Research Program

Valse team works in the domains of control science: dynamical systems, stability analysis, estimation and automatic control. *Our developments are focused on the theoretical and applied aspects related to control and estimation of large-scale multi-sensor and multi-actuator systems based on the use of the theories of finite-time/fixed-time/hyperexponential convergence and homogeneous systems*. The Lyapunov function method and other methods of analysis of dynamical systems form a basis for the studies in Valse team.

The key idea of research program for the team is that a fast (non-asymptotic) convergence of the regulation and estimation errors increases the reliability of intelligent distributed actuators and sensors in complex scenarios, such as interconnected cyber-physical systems (CPSs).

The expertise of Valse's members in theoretical developments of control and estimation theory (finite-time control and estimation algorithms in centralized context [84], [70], [81], [80], [77], homogeneity framework for differential equations [85], [72], [71], [73], [75], [86], [82], time-delay systems [74], [76], [89], distributed systems [83] and algebraic-based methods for estimation [87], [88]) is an essential ingredient to achieve our objective.

The generic chart of different goals and tasks included in the scientific work program of Valse, and interrelations between them, are presented in Fig. 1. We have selected three main objectives to pursuit with the related tasks to fulfill:

- The first objective consists in design of control and estimation solutions for CPS and IoT, which is the principal aim of Valse, it will contain the main outcomes of our research.
- The second objective is more theoretical, which is needed to make the basement for our design and analysis parts in the previous goal.
- The third objective deals with applications, which will drive the team and motivate the theoretical studies and selected design performances.

All these objectives are interconnected: from a particular problem in an IoT application, it is planned to design a control or estimation algorithm, which leads to development of theoretical tools; and *vice versa*, a new theoretical advance can provide a possibility for development of novel tools, which can be used in applications.

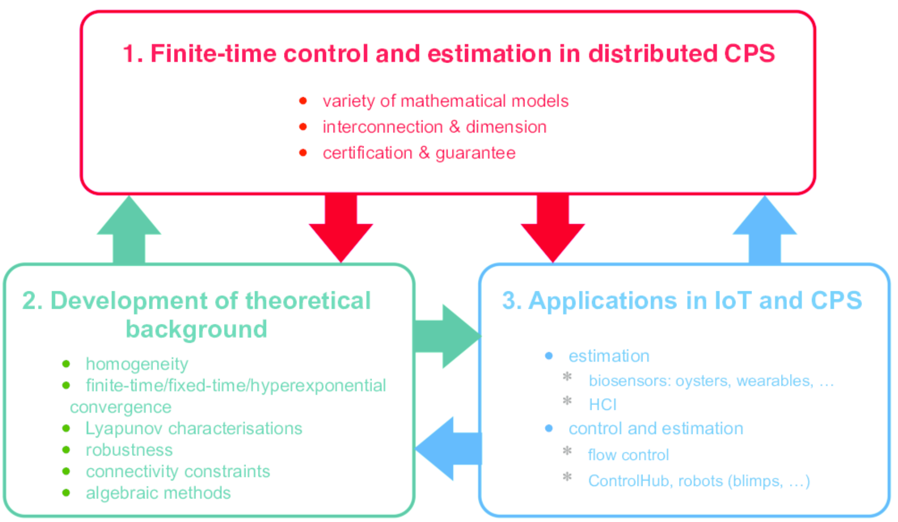

Figure 1. Structure of the objectives and tasks treated in Valse

To explain our motivation: *why to use finite-time?* Applying any method for control/estimation has a price in terms of its advantages and disadvantages. There is no universal framework that is the best always and everywhere. Finite-time may appear as a luxurious property for a physical system, requiring the use of nonlinear tools. Of course, if an asymptotic convergence and a linear model are enough for solving a given problem, then there is no reason to develop something else. However, most of the present problems in CPS and IoT are nonlinear (i.e. they have various local behaviors that cannot be collected in only one linear model). Design and analysis of various local linearized models and solutions are luxurious, too. The theory of homogeneity can go beyond linearity offering many new features, while not appearing as severe as other nonlinear tools and having almost all hints of the linear framework. Suppose that, thanks to the homogeneity theory, finite-time/fixed-time can be obtained with a limited difficulty, while adding the bonuses of a stronger robustness and a faster convergence compared to the linear case? *We are convinced that the price of going beyond linear control and estimation can be strongly dropped down by maturing the theory of homogeneity and finite/fixed-time convergence. And also, convinced that it will be compensated in terms of robustness and speed, which can be demanded in the new areas of application as IoT, for example.*

4. Application Domains

4.1. Application Domains

An objective of the team is the application of the developed control and estimation algorithms for different scenarios in IoT or CPSs. The participation in various potential applications allows Valse team to better understand the features of CPSs and their required performances, and to formulate properly the control and estimation problems that have to be solved. Here is a list of ongoing and potential applications addressed in the team:

- smart bivalve-based biosensor for water quality monitoring (ANR project WaQMoS, the developed sensor is shown in Fig. 2): in living beings, the presence of persistent external perturbations, may be difficult to measure, and important model uncertainties render the application of conventional techniques complicated; another issue for estimation is the consensus seeking between animals for a contamination detection [68];

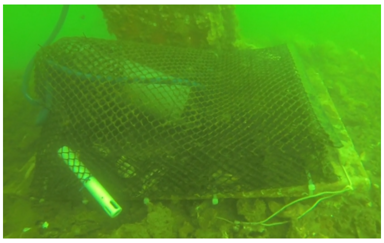

Figure 2. The valvometer used in ANR project WaQMoS

- control and estimation for flying vehicles, *e.g.* quadrotors or blimps given in Fig. 3 (PhD Centrale Lille): nonlinearity of the model and its uncertainty coupled with important aerodynamic perturbations have to be compensated by fast (finite- or fixed-time) and robust control and estimation algorithms;
- human behavior modeling and identification with posterior design of algorithms for human-computer interaction (ANR project TurboTouch): robust finite-time differentiators demonstrate good estimation capabilities needed for prediction in this application [88], [69];
- human physiological characteristics estimation (like emotion detection, galvanic skin response filtering, fatigue evaluation in collaborations with Neotrope and Ellcie Healthy): intelligent robust filtering and finite-time distributed estimation are key features in these scenarios;
- path planning for autonomous vehicles taking into account the behavior of humans (PhD CIFRE with SEQUEL team and Renault): application of interval and finite-time adaptive estimation and prediction techniques allows for treating the uncertainty of the environment by reducing computational complexity of reinforcement learning [79] [0];

[0] The examples of interval prediction algorithm application can be consulted here.

Figure 3. Blimp and quadrotor robots

- flow control (in the framework of ContrATech subprogram of CPER ELSAT, see also [78]): the case of control and estimation of a distributed-parameter system with very fast and uncertain dynamics, where finite-time solutions developed by Valse are necessary (an example of results is given in Fig. 4);

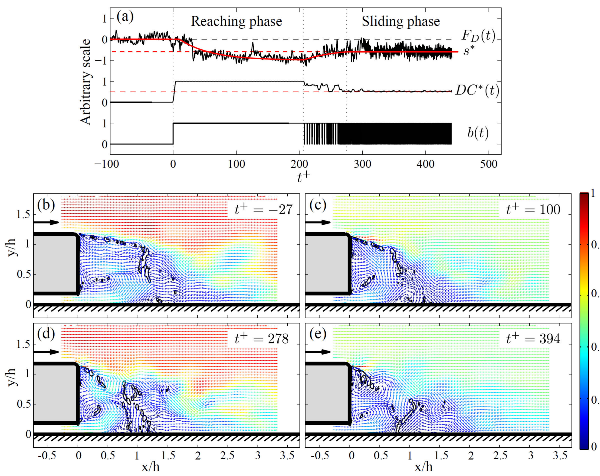

Figure 4. Particle Image Velocimetry on flow control for an Ahmed body (LAMIH wind tunnel)

- control of synthetic microbial communities (in the framework of IPL COSY, the experimental platform is shown in Fig. 5): here again, the problem is an important uncertainty of the model, which can be handled by robust sliding mode control algorithms, or by applying adaptive finite-time estimation and identification tools;

It is worth to highlight a widespread distribution of various scientific domains in the list of applications for the team given above. Such an *interdisciplinarity* for Valse is unsurprising, since the control theory is a science of systems whose interest today is, by nature, to interface with other disciplines and their fields of application. This is also well aligned with the domain of CPSs, which by its origin requires multidisciplinary competences.

5. Highlights of the Year

5.1. Highlights of the Year

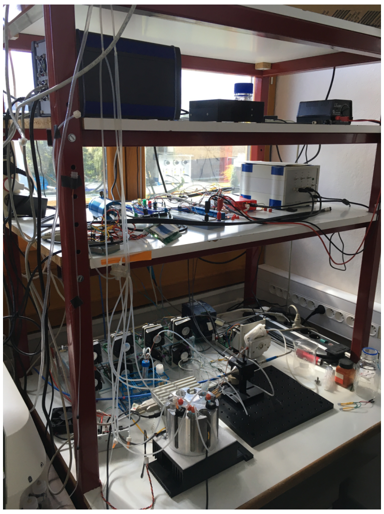

Figure 5. Chemostat platform at Inria, Grenoble

- This year Valse published 6 papers in Automatica and 4 in IEEE Transaction on Automatic Control (the top journals in the domain of control theory).
- A. Polyakov wrote a book Generalized Homogeneity in Systems and Control.

6. New Software and Platforms

6.1. Platforms

6.1.1. ControlHub

ControlHub is the platform for the rapid prototyping of control algorithms of cyber-physical systems. It allows the End-User to have a remote access to real experimental setups in order to validate mathematical control and estimation algorithms.

7. New Results

7.1. Stabilization of multistable systems

Participants: Rosane Ushirobira, Denis Efimov.

The problem of robust stabilization of affine nonlinear multistable systems with respect to disturbance inputs was studied in [29]. The results are obtained using the framework of input-to-state stability and integral input-to-state stability for systems with multiple invariant sets. The notions of corresponding control Lyapunov functions as well as the small control property are extended within the multistability framework. It is verified that the universal control formula can be applied. In a similar vein, input-to-state stability and stabilization of passive multistable systems were investigated in [53].

7.2. Robustness of homogeneous systems

Participants: Andrey Polyakov, Denis Efimov.

In [32], we studied the finite-time stability of a class of nonlinear systems $\dot{x} = f(x) = H(x)b(x)$, where H is homogeneous and b is bounded. We defined the homogeneous extension of the non-homogeneous function f and used this extension to prove that, under some conditions on b, if the system $\dot{x} = f(x)$ is globally asymptotically stable, then it is finite-time stable. In [31], a theoretical basement of the previous result has been given showing robust stability of the system $\dot{x} = f(x) = H(x)b(x)$ by considering b as a perturbation.

7.3. Finite-time control and estimation in chemostat

Participants: Denis Efimov, Andrey Polyakov.

In [34], the problem of robust stabilization of the concentration of two different species competing for a single limiting substrate was addressed. Such a stabilization is performed by means of discontinuous feedback control laws that ensure coexistence of all species. The control laws are designed considering bounded uncertainties on the kinetic rates and full state measurements. The problem of robust finite-time estimation of the state for a chemostat was considered in [35].

7.4. LKF for neutral type time-delay systems

Participant: Denis Efimov.

The problem of existence of a Lyapunov-Krasovskii functional (LKF) for nonlinear neutral type time-delay systems was revisited in [37] considering the uniform stability analysis and the LKF in a Sobolev space of absolutely continuous functions with bounded derivatives.

7.5. Finite-time/Fixed-time control of PDEs

Participant: Andrey Polyakov.

The paper [40] deals with continuous boundary time-varying feedbacks for fixed-time stabilization of constant-parameter reaction-diffusion systems. IT was shown that the time of convergence can be prescribed and is independent of the initial condition of the system. The design of time-varying feedbacks was carried out by using the backstepping approach for which suitable characterizations for time-varying kernels were derived. The work [39] considered the problem of local finite-time stabilization of the viscous Burgers equation. A boundary switched linear control with state dependent switching law was designed based on the backstepping approach. The strategy builds on discontinuous kernels, which render the control function to be piecewise continuous. It was proven that such a control stabilizes locally the viscous Burgers equation and that the settling time depends on initial conditions.

7.6. Interval prediction for LPV systems

Participant: Denis Efimov.

The problem of behaviour prediction for linear parameter-varying (LPV) systems was considered in the interval framework in [41]. It was assumed that the system is subject to uncertain inputs and the vector of scheduling parameters is unmeasurable, but all uncertainties take values in a given admissible set. Then an interval predictor was designed and its stability was guaranteed applying Lyapunov function with a novel structure. The conditions of stability were formulated in the form of linear matrix inequalities. Efficiency of the theoretical results was demonstrated in the application to safe motion planning for autonomous vehicles (see videos of numeric experiments).

7.7. Finite-time/Fixed-time stability and stabilization

Participants: Andrey Polyakov, Rosane Ushirobira.

In [33], we studied the issues of prescribed-time stabilization of a chain of integrators of arbitrary length, that can be either pure (i.e. with no disturbance) or perturbed. The feedback law proposed by Song et al. was revisited, and it was shown that by an appropriately recast within the framework of time-varying homogeneity it can stabilize the system in a prescribed time. Characterizations (necessary and sufficient conditions) of finite-time and fixed-time stability of evolution inclusions in Banach spaces were presented in [44] in terms of Lyapunov functionals.

7.8. Robust stability analysis for Persidskii systems

Participant: Denis Efimov.

A class of generalized nonlinear Persidskii systems was considered in [36]. The conditions of input-to-state and integral input-to-state stability were established, which can be checked using linear matrix inequalities. The issues of discretization of this class of dynamics were analyzed using the Euler methods. The proposed theory was applied to a Lotka-Volterra model.

7.9. Design of a distributed finite-time observer

Participants: Denis Efimov, Rosane Ushirobira.

In [46], a distributed observer was presented to estimate the state of a linear time-invariant plant in finite-time in each observer node. The design was based on a decomposition into locally observable and unobservable substates and on properties of homogeneous systems. Each observer node can reconstruct in finite-time its locally observable substate with its measurements only. Then exploiting the coupling, a finite-time converging observer was constructed for the remaining states by adding the consensus terms.

7.10. Parameter estimation in finite-time without PE

Participant: Denis Efimov.

The problem of estimation in the linear regression model was studied in [24], [48], [25] under the hypothesis that the regressor may be excited on a limited initial interval of time only. Then the estimation solution was searched on a finite interval of time also based on the framework of finite-time or fixed-time converging dynamical systems. The robustness issue was analyzed and a short-time input-to-state stability property was introduced for fixed-time converging time-varying systems with a sufficient condition, which was formulated with the use of a Lyapunov function. Several estimation algorithms were proposed and compared with existing solutions.

7.11. Quadrotor control

Participant: Andrey Polyakov.

The problem of a state feedback design for control of a quadrotor system under state and time constraints was studied in [49]. Convex embedding approach and Implicit Lyapunov function were employed to design a finite-time controller. The feedback gains were calculated by a system of LMIs (see video of experiments).

7.12. Blimp robot control

Participant: Denis Efimov.

The papers [50], [26] presented the robust controller design for an indoor blimp robot to achieve application such as the surveillance. The commonly used 6 degrees of freedom dynamic model was simplified under reasonable assumptions and decoupled into two independent parts. The blimp simplified horizontal plane movement model was complemented with disturbance terms to ensure the modeling accuracy, then it was transformed to a simpler form for the ease of controller design. Next, the disturbance terms were evaluated by the designed real-time estimator, and the perturbation estimates were compensated in the conceived motion controller for cancellation of the influence of disturbances. The performance and robustness of the disturbance compensation-based controller were verified by both simulations and experiments on the developed blimp robot.

7.13. Output finite-time stability and stabilization

Participants: Denis Efimov, Andrey Polyakov.

The equivalent Lyapunov characterizations of output finite-time stability were presented in [51]. Another sufficient condition for output finite-time stability was presented in [52]. Based on this condition a scheme of adaptive finite-time control design was provided. The presented results were obtained with the use of homogeneity property.

8. Partnerships and Cooperations

8.1. Regional Initiatives

The team participates in CPER Data programs and projects:

- ControlHub, coordinator A. Polyakov, see the dedicated platform description above
- “ContrATech” subprogram of CPER ELSAT, coordinator J.-M. Foucaut (LMFL)

8.2. National Initiatives

8.2.1. ANR

- Digitslid, coordinator B. Brogliato (Inria, Grenoble)
- Finite4SoS, coordinator W. Perruquetti (École Centrale de Lille)
- WaQMoS, coordinator D. Efimov (Inria, Lille)
- TurboTouch, coordinator G. Casiez (Inria, Lille)

8.2.2. Inria project labs

The team participates in IPL COSY, coordinator E. Cinquemani (Inria, Grenoble).

8.3. European Initiatives

8.3.1. FP7 & H2020 Projects

The team is involved in 1 EU project UCoCoS, coordinator W. Michiels (KUL, Belgium).

8.4. International Initiatives

8.4.1. Inria North European associate teams

- WeCare with Uppsala University (Sweden), coordinator R. Ushirobira
- RECoT with IBM Research (Ireland), coordinator A. Polyakov

8.4.2. Inria International Partners

- UNAM (Mexico), L. Fridman and J. Moreno
- ITMO University (Russia), A. Bobtsov and I. Furtat

8.5. International Research Visitors

8.5.1. Visits of International Scientists

- A. Aleksandrov, SPbSU (Russia), from Mar 2019 until Apr 2019
- S. Aranovskiy, École supérieure d’électricité, Jun 2019
- J. Epperlein, IBM Research (Ireland), from Dec 2019

- E. Fridman, Tel Aviv University (Israel), from Jul 2019 until Sep 2019, **Inria invited professor**
- A. Medvedev, Upsalla University (Sweden), from Nov 2019 until Dec 2019
- J. Moreno, UNAM (Mexico), from Dec 2019
- Y. Orlov, CICESE (Mexico), from Dec 2019
- M. Ruderman, University of Agder (Norway), from Dec 2019
- L. Tupak Aguilar Bustos, CICESE (Mexico), from Dec 2019
- J. Zhang, HDU (China), from Aug 2019
- K. Zimenko, ITMO University (Russia), from Sep 2019 until Nov 2019
- S. Zhuk, IBM Research (Ireland), from Dec 2019

8.5.1.1. Internships

- M. Ballesteros Escamilla, Cinvestav (Mexico), until Apr 2019
- D. Cruz Ortiz, Cinvestav (Mexico), until Apr 2019
- J. Franco Jaramillo, Technological Institute of La Laguna (Mexico), from Oct 2019 until Nov 2019
- J. Mendoza Avila, UNAM (Mexico), from Sep 2019 until Dec 2019

9. Dissemination

9.1. Promoting Scientific Activities

9.1.1. Scientific Events: Selection

9.1.1.1. Member of the Conference Program Committees

- J.-P. Richard, EUCA-IEEE ECC (Napoli, Italy)
- J.-P. Richard, IFAC TDS (Sinaia, Romania)
- J.-P. Richard, IEEE CODIT (Paris, France)
- D. Efimov, IFAC ALCOS (Winchester, UK)

9.1.1.2. Reviewer

The members of the team participate in reviewing for all major international conferences in the domain of control theory.

9.1.2. Journal

9.1.2.1. Member of the Editorial Boards

- R. Ushirobira, Guest editor, European Journal of Control
- A. Polyakov, Associate editor, Automation and Remote Control
- D. Efimov, Guest editor, International Journal of Control
- D. Efimov, Guest editor, European Journal of Control
- D. Efimov, Guest editor, International Journal on Robust and Nonlinear Control
- D. Efimov, Associate editor, IFAC Journal on Nonlinear Analysis: Hybrid Systems
- D. Efimov, Associate editor, Asian Journal of Control
- D. Efimov, Associate editor, IEEE Transactions on Automatic Control

9.1.2.2. Reviewer - Reviewing Activities

The members of the team participate in reviewing for all major international journals in the domain of control theory: IEEE Transactions on Automatic Control, Automatica, European Journal of Control, International Journal of Control, International Journal on Robust and Nonlinear Control, Asian Journal of Control, SIAM Journal on Control and Optimization etc.

9.1.3. Invited Talks

D. Efimov and A. Polyakov participated in organization of a workshop Finite-, fixed-, and prescribed-time stabilization and estimation for IEEE Conference on Decision and Control (CDC) at Nice.

9.1.4. Research Administration

- J.-P. Richard, Director of the professional training "Researcher" for last year students at Centrale Lille
- R. Ushirobira, a member of the executive board of CIMPA
- R. Ushirobira, the vice-president of the Recruitment research committee of Inria Lille (PhD, postdoc, secondments, visitors)

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master: J.-P. Richard, Dynamical systems, 10h, M2, Université de Lille, France

Licence: R. Ushirobira, Basic courses in Linear algebra and Calculus, 75h, L3, Polytech Lille, France

Master: D. Efimov, Dynamical systems, 17h, M2, Université de Lille, France

Licence: D. Efimov, Estimation for engineers, 24h, L3, École Centrale de Lille, France

9.2.2. Supervision

PhD: Y. Wang, Development of a blimp robot for indoor operation, École Centrale de Lille, 15/03/2019

PhD: T. Kharkovskaya, Design of interval observers for uncertain distributed systems, École Centrale de Lille & ITMO University (Russia), 02/12/2019

9.2.3. Juries

- R. Ushirobira, a member of PhD thesis committee of D. Yamalova (Uppsala University, Sweden), June 2019

9.3. Popularization

9.3.1. Internal or external Inria responsibilities

At Inria (Lille), R. Ushirobira is responsible for a monthly organized talks "30 minutes de sciences".

9.3.2. Education

- R. Ushirobira, referent researcher for projects with middle-school students at Collège A. Daudet (Leers) within the framework of Math en Jeans

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] T. KHARKOVSKAIA.*Design of interval observers for uncertain distributed systems*, Ecole Centrale Lille, December 2019, https://tel.archives-ouvertes.fr/tel-02425063

Articles in International Peer-Reviewed Journal

[2] H. AHMED, R. USHIROBIRA, D. EFIMOV, L. FRIDMAN, Y. WANG.*Output Global Oscillatory Synchronization of Heterogeneous Systems*, in "International Journal of Control", 2019, forthcoming, https://hal.inria.fr/hal-02344850

[3] A. APARICIO, L. FRIDMAN, D. EFIMOV.*Stabilization of systems with switchings on the axis of their coordinates and its input-to-state properties*, in "Nonlinear Analysis: Hybrid Systems", 2019, forthcoming [*DOI :* 10.1016/J.NAHS.2018.10.008], https://hal.inria.fr/hal-01908289

[4] S. ARANOVSKIY, R. USHIROBIRA, D. EFIMOV, G. CASIEZ.*A switched dynamic model for pointing tasks with a computer mouse*, in "Asian Journal of Control", 2019, forthcoming [*DOI :* 10.1002/ASJC.2063], https://hal-centralesupelec.archives-ouvertes.fr/hal-02083760

[5] S. ARANOVSKIY, R. USHIROBIRA, D. EFIMOV, G. CASIEZ.*An adaptive FIR filter for trajectory prediction and latency reduction in direct Human-Computer interactions*, in "Control Engineering Practice", 2019, forthcoming [*DOI :* 10.1016/J.CONENGPRAC.2019.07.011], https://hal.inria.fr/hal-02189181

[6] 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

[7] C. CHENAVIER.*Topological rewriting systems applied to standard bases and syntactic algebras*, in "Journal of Algebra", 2020, forthcoming, https://hal.archives-ouvertes.fr/hal-02355020

[8] C. CHOVET, M. FEINGESICHT, B. PLUMJEAU, M. LIPPERT, L. KEIRSBULCK, F. KERHERVÉ, A. POLYAKOV, J.-P. RICHARD, W. ABASSI, J.-M. FOUCAUT.*Sliding mode control applied to a square-back Ahmed body*, in "European Journal of Mechanics - B/Fluids", July 2019 [*DOI :* 10.1016/J.EUROMECHFLU.2019.07.010], https://hal.archives-ouvertes.fr/hal-02422123

[9] B.-H. DU, A. POLYAKOV, G. ZHENG, Q. QUAN.*Quadrotor trajectory tracking by using fixed-time differentiator*, in "International Journal of Control", 2019, vol. 2019 - Issue 12 [*DOI :* 10.1080/00207179.2018.1462534], https://hal.inria.fr/hal-01947365

[10] D. EFIMOV, N. BARABANOV, R. ORTEGA.*Robustness of linear time-varying systems with relaxed excitation*, in "International Journal of Adaptive Control and Signal Processing", April 2019 [*DOI :* 10.1002/ACS.2997], https://hal.inria.fr/hal-02082769

[11] D. EFIMOV, A. POLYAKOV, A. ALEKSANDROV.*Discretization of Homogeneous Systems Using Euler Method with a State-Dependent Step*, in "Automatica", 2019, vol. 109 [*DOI :* 10.1016/J.AUTOMATICA.2019.108546], https://hal.inria.fr/hal-02190206

[12] D. EFIMOV, J. SCHIFFER.*On boundedness of solutions of state periodic systems: a multivariable cell structure approach*, in "IEEE Transactions on Automatic Control", October 2019, vol. 10, 4094-4104 [*DOI :* 10.1109/TAC.2019.2892395], https://hal.inria.fr/hal-01966708

[13] N. ESPITIA, A. POLYAKOV, D. EFIMOV, W. PERRUQUETTI.*Boundary time-varying feedbacks for fixed-time stabilization of constant-parameter reaction-diffusion systems*, in "Automatica", May 2019, vol. 103, p. 398-407 [*DOI :* 10.1016/J.AUTOMATICA.2019.02.013], https://hal.inria.fr/hal-01962662

[14] L. ETIENNE, L. HETEL, D. EFIMOV.*Observer analysis and synthesis for perturbed Lipschitz systems under noisy time-varying measurements*, in "Automatica", 2019, vol. 106, p. 406-410 [*DOI :* 10.1016/J.AUTOMATICA.2019.04.003], https://hal.archives-ouvertes.fr/hal-02304891

[15] T. KHARKOVSKAIA, D. EFIMOV, E. FRIDMAN, A. POLYAKOV, J.-P. RICHARD.*Interval observer design and control of uncertain non-homogeneous heat equations*, in "Automatica", 2019, forthcoming, https://hal.inria.fr/hal-02283008

[16] F. LOPEZ-RAMIREZ, D. EFIMOV, A. POLYAKOV, W. PERRUQUETTI.*Conditions for Fixed-Time Stability and Stabilization of Continuous Autonomous Systems*, in "Systems and Control Letters", July 2019, vol. 129, p. 26-35 [*DOI :* 10.1016/J.SYSCONLE.2019.05.003], https://hal.inria.fr/hal-02142027

[17] 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

[18] J. G. RUEDA-ESCOBEDO, R. USHIROBIRA, D. EFIMOV, J. A. MORENO.*Gramian-based uniform convergent observer for stable LTV systems with delayed measurements*, in "International Journal of Control", 2019 [*DOI :* 10.1080/00207179.2019.1569256], https://hal.inria.fr/hal-01889193

[19] H. RÍOS, L. HETEL, D. EFIMOV.*Robust output-feedback control for uncertain linear sampled-data systems: A 2D impulsive system approach*, in "Nonlinear Analysis: Hybrid Systems", May 2019, vol. 32, p. 177-201 [*DOI :* 10.1016/J.NAHS.2018.11.005], https://hal.archives-ouvertes.fr/hal-01954885

[20] T. SANCHEZ, D. EFIMOV, A. POLYAKOV, J. A. MORENO, W. PERRUQUETTI.*On homogeneity of discrete-time systems: stability and convergence rates*, in "International Journal of Robust and Nonlinear Control", 2019, vol. 29, n^{o} 8, p. 2406-2421 [*DOI :* 10.1002/RNC.4497], https://hal.inria.fr/hal-01996519

[21] T. SANCHEZ, A. POLYAKOV, E. FRIDMAN, L. HETEL.*A Switching Controller for a class of MIMO Bilinear Systems with Time-Delay*, in "IEEE Transactions on Automatic Control", September 2019, https://hal.inria.fr/hal-02285126

[22] T. SANCHEZ, A. POLYAKOV, J.-P. RICHARD.*A Sliding Mode Controller for a Model of Flow Separation in Boundary Layers*, in "International Journal of Robust and Nonlinear Control", 2019, forthcoming, https://hal.inria.fr/hal-02373078

[23] J. SCHIFFER, D. EFIMOV, R. ORTEGA.*Global Synchronization Analysis of Droop-Controlled Microgrids: A Multivariable Cell Structure Approach*, in "Automatica", November 2019, vol. 109, 108550 [*DOI :* 10.1016/J.AUTOMATICA.2019.108550], https://hal.inria.fr/hal-02177492

[24] J. WANG, D. EFIMOV, S. ARANOVSKIY, A. A. BOBTSOV.*Fixed-time estimation of parameters for non-persistent excitation*, in "European Journal of Control", August 2019 [*DOI :* 10.1016/J.EJCON.2019.07.005], https://hal.inria.fr/hal-02196637

[25] J. WANG, D. EFIMOV, A. A. BOBTSOV.*On robust parameter estimation in finite-time without persistence of excitation*, in "IEEE Transactions on Automatic Control", April 2020, forthcoming, https://hal.inria.fr/hal-02196992

[26] Y. WANG, G. ZHENG, D. EFIMOV, W. PERRUQUETTI.*Disturbance Compensation Based Controller for an Indoor Blimp Robot*, in "Robotics and Autonomous Systems", 2019, forthcoming, https://hal.inria.fr/hal-02406793

[27] K. ZIMENKO, D. EFIMOV, A. POLYAKOV, A. KREMLEV.*Independent of delay stabilization using implicit Lyapunov function method*, in "Automatica", March 2019, vol. 101, p. 103-110 [*DOI :* 10.1016/J.AUTOMATICA.2018.11.051], https://hal.inria.fr/hal-02418531

International Conferences with Proceedings

[28] M. BALLESTEROS, A. POLYAKOV, D. EFIMOV, I. CHAIREZ, A. POZNYAK.*Differential Neural Network Identification for Homogeneous Dynamical Systems* ⋆, in "NOLCOS 2019 - 11th IFAC Symposium on Nonlinear Control Systems", Vienna, Austria, September 2019, https://hal.inria.fr/hal-02278726

[29] N. F. BARROSO, R. USHIROBIRA, D. EFIMOV.*Universal formula for robust stabilization of affine nonlinear multistable systems*, in "Conference on Decision and Control", Nice, France, December 2019, https://hal.inria.fr/hal-02340333

[30] F. J. BEJARANO, R. USHIROBIRA, G. ZHENG.*On the observability and detectability of non-commensurate time-delay linear systems*, in "2019 18th European Control Conference (ECC)", Naples, Italy, IEEE, June 2019, p. 2921-2925 [*DOI :* 10.23919/ECC.2019.8796285], https://hal.inria.fr/hal-02406860

[31] Y. BRAIDIZ, D. EFIMOV, A. POLYAKOV, W. PERRUQUETTI.*Robust Finite-time stability of homogeneous systems with respect tomultiplicative disturbances*, in "ECC 2019 - European Control Conference", Naples, Italy, June 2019, https://hal.inria.fr/hal-02084980

[32] Y. BRAIDIZ, A. POLYAKOV, D. EFIMOV, W. PERRUQUETTI.*On finite-time stability of homogeneous systems with multiplicative bounded function* *, in "ECC 2019 - European Control Conference", Nalpes, Italy, June 2019, https://hal.inria.fr/hal-02137648

[33] Y. CHITOUR, R. USHIROBIRA.*Stabilization of a perturbed chain of integrators in prescribed time*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://hal.inria.fr/hal-02418557

[34] A. DOS REIS DE SOUZA, D. EFIMOV, A. POLYAKOV, J.-L. GOUZÉ.*Robust Control of a Competitive Environment in the Chemostat using Discontinuous Control Laws*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://hal.inria.fr/hal-02308076

[35] A. DOS REIS DE SOUZA, J.-L. GOUZÉ, A. POLYAKOV, D. EFIMOV.*On Adaptive Estimation of Bacterial Growth in the Competitive Chemostat*, in "NOLCOS 2019 - 11th IFAC Symposium Nonlinear Control Systems", Vienne, Austria, September 2019, https://hal.inria.fr/hal-02416265

[36] D. EFIMOV, A. ALEKSANDROV.*Robust stability analysis and implementation of Persidskii systems*, in "Proc. IEEE CDC", Nice, France, December 2019, https://hal.inria.fr/hal-02418534

[37] D. EFIMOV, E. FRIDMAN.*A note on converse Lyapunov-Krasovskii theorems for nonlinear neutral systems in Sobolev spaces*, in "Proc. IFAC Nolcos", Vienna, Austria, September 2019, https://hal.inria.fr/hal-02418539

[38] D. EFIMOV, E. FRIDMAN, J.-P. RICHARD.*On robust stability of sine-Gordon equation*, in "Proc. IEEE CDC", Nice, France, December 2019, https://hal.inria.fr/hal-02418533

[39] N. ESPITIA, A. POLYAKOV, D. EFIMOV, W. PERRUQUETTI.*Some characterizations of boundary time-varying feedbacks for fixed-time stabilization of reaction-diffusion systems*, in "CPDE'19 - 3rd IFAC/IEEE CSS Workshop on Control of Systems Governed by Partial Differential Equations", Oaxaca, Mexico, May 2019, https://hal.inria.fr/hal-02146572

[40] N. ESPITIA, A. POLYAKOV, E. FRIDMAN.*On local finite-time stabilization of the Viscous Burgers equation via boundary switched linear feedback*, in "CDC 2019 - 58thIEEE Conference on Decision and Control,", Nice, France, December 2019, https://hal.inria.fr/hal-02343238

[41] E. LEURENT, D. EFIMOV, T. RAÏSSI, W. PERRUQUETTI.*Interval Prediction for Continuous-Time Systems with Parametric Uncertainties*, in "58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://arxiv.org/abs/1904.04727 - Webpage: https://eleurent.github.io/interval-prediction/, https://hal.inria.fr/hal-02383571

[42] A. MERCADO-URIBE, J. A. MORENO, A. POLYAKOV, D. EFIMOV.*Integral Control Design using the Implicit Lyapunov Function Approach*, in "58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://hal.inria.fr/hal-02371919

[43] 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

[44] A. POLYAKOV.*Characterization of Finite/Fixed-time Stability of Evolution Inclusions*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://hal.inria.fr/hal-02278740

[45] A. POLYAKOV, S. ZHUK.*Generalized Lyapunov Exponents of Homogeneous Systems*, in "CDC 2019 - 58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://hal.inria.fr/hal-02285137

[46] H. SILM, R. USHIROBIRA, D. EFIMOV, W. MICHIELS, J.-P. RICHARD.*Design of a distributed finite-time observer using observability decompositions*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, IEEE, 2019, p. 1816-1821, https://hal.archives-ouvertes.fr/hal-02279201

[47] R. USHIROBIRA, D. EFIMOV, P.-A. BLIMAN.*Estimating the infection rate of a SIR epidemic model via differential elimination*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, June 2019, https://hal.inria.fr/hal-02418553

[48] J. WANG, D. EFIMOV, A. A. BOBTSOV.*Finite-time parameter estimation without persistence of excitation*, in "ECC 2019 - European Control Conference", Naples, Italy, June 2019, https://hal.inria.fr/hal-02084983

[49] S. WANG, A. POLYAKOV, G. ZHENG.*Quadrotor Control Design under Time and State Constraints: Implicit Lyapunov Function Approach*, in "ECC'19 - European Control Conference", Naples, Italy, June 2019, https://hal.inria.fr/hal-02093641

[50] Y. WANG, G. ZHENG, D. EFIMOV, W. PERRUQUETTI.*Disturbance Compensation Based Control for an Indoor Blimp Robot*, in "2019 International Conference on Robotics and Automation (ICRA)", Montreal,

Canada, IEEE, May 2019, p. 2040-2046 [*DOI :* 10.1109/ICRA.2019.8793535], https://hal.inria.fr/hal-02406847

[51] K. ZIMENKO, D. EFIMOV, A. POLYAKOV, A. KREMLEV.*On Notions of Output Finite-Time Stability*, in "Proc. European Control Conference", Naples, Italy, June 2019, https://hal.inria.fr/hal-02084981

[52] K. ZIMENKO, D. EFIMOV, A. POLYAKOV.*On Condition for Output Finite-Time Stability and Adaptive Finite-Time Control Scheme*, in "58th IEEE Conference on Decision and Control", Nice, France, December 2019, https://hal.inria.fr/hal-02371955

[53] N. DE FIGUEIREIDO BARROSO, R. USHIROBIRA, D. EFIMOV, A. L. FRADKOV.*On robust stability of multistable passive systems*, in "ECC 2019 - 17th European Control Conference", Naples, Italy, June 2019, https://hal.inria.fr/hal-02418552

Conferences without Proceedings

[54] C. CHENAVIER, M. LUCAS.*The Diamond Lemma for non-terminating rewriting systems using deterministic reduction strategies*, in "IWC 2019 - 8th International Workshop on Confluence", Dortmund, Germany, June 2019, p. 1-5, https://hal.archives-ouvertes.fr/hal-02385139

[55] D. DILEEP, M. FUSCO, J. VERHAEGH, L. HETEL, J.-P. RICHARD, W. MICHIELS.*Achieving an $\mathcal{L}_2$ string stable one vehicle look-ahead platoon with heterogeneity in time-delays*, in "ECC 2019 - 18th European Control Conference", Naples, Italy, IEEE, June 2019, p. 1220-1226 [*DOI :* 10.23919/ECC.2019.8796064], https://hal.archives-ouvertes.fr/hal-02337916

[56] O. LAKHAL, A. MELINGUI, G. DHERBOMEZ, R. MERZOUKI.*Control of a Hyper-Redundant Robot for Quality Inspection in Additive Manufacturing for Construction*, in "RoboSoft 2019 - IEEE International Conference on Soft Robotics", Seoul, South Korea, April 2019, p. 1-6, https://hal.archives-ouvertes.fr/hal-02058069

Scientific Books (or Scientific Book chapters)

[57] *Normal forms of matrix words for stability analysis of discrete-time switched linear systems*, May 2020, https://hal.archives-ouvertes.fr/hal-02069712

[58] A. POLYAKOV.*Generalized Homogeneity in Systems and Control*, Springer, April 2020, https://hal.inria.fr/hal-02426338

Patents and standards

[59] A. ZOLGHADRI, J. CIESLAK, A. GHEORGHE, D. EFIMOV, G. PHILIPPE, R. DAYRE.*Method and device for detecting a failure on an aircraft*, February 2019, n^{o} US010209162B2, https://hal.archives-ouvertes.fr/hal-02056326

Other Publications

[60] C. CHENAVIER.*A constructive version of Warfield's Theorem*, June 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02120656

[61] C. CHENAVIER, R. USHIROBIRA, G. VALMORBIDA.*A geometric stabilization of planar switched systems*, November 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-02366928

[62] D. EFIMOV, A. POLYAKOV, A. ALEKSANDROV.*Proofs for "Discretization of Homogeneous Systems Using Euler Method with a State-Dependent Step"*, June 2019, working paper or preprint, https://hal.inria.fr/hal-02164050

[63] A. POLYAKOV.*On Expansion of Regularity of Nonlinear Evolution Equations by Means of Dilation Symmetry*, May 2019, working paper or preprint, https://hal.inria.fr/hal-02093984

[64] A. POLYAKOV.*On global existence of strong and classical solutions of Navier-Stokes equations*, April 2019, working paper or preprint, https://hal.inria.fr/hal-02096215

[65] T. SANCHEZ, D. EFIMOV, A. POLYAKOV, J. A. MORENO.*Homogeneous discrete-time approximation*, January 2019, working paper or preprint, https://hal.archives-ouvertes.fr/hal-01990132

[66] J. WANG, D. EFIMOV, A. A. BOBTSOV.*On robust parameter estimation in finite-time without persistence of excitation*, January 2019, working paper or preprint, https://hal.inria.fr/hal-01998043

[67] S. ZHUK, A. POLYAKOV.*On Practical Fixed-Time Convergence for Differential Riccati Equations*, December 2019, working paper or preprint, https://hal.inria.fr/hal-02390409

References in notes

[68] H. AHMED, R. USHIROBIRA, D. EFIMOV, D. TRAN, J.-C. MASSABUAU.*Velocity estimation of valve movement in oysters for water quality surveillance*, in "IFAC-PapersOnLine", 2015, vol. 48, p. 333–338

[69] S. ARANOVSKIY, R. USHIROBIRA, D. EFIMOV, G. CASIEZ.*Frequency Domain Forecasting Approach for Latency Reduction in Direct Human-Computer Interaction*, in "Proc. 56th IEEE Conference on Decision and Control (CDC)", Melbourne, 2017

[70] E. BERNUAU, D. EFIMOV, W. PERRUQUETTI, E. MOULAY.*Robust finite-time output feedback stabilization of the double integrator*, in "International Journal of Control", 2015, vol. 88, n^{o} 3, p. 451–460

[71] E. BERNUAU, D. EFIMOV, W. PERRUQUETTI, A. POLYAKOV.*On Homogeneity and Its Application in Sliding Mode*, in "Int. J. Franklin Institute", 2014, vol. 351, n^{o} 4, p. 1866–1901

[72] E. BERNUAU, A. POLYAKOV, D. EFIMOV, W. PERRUQUETTI.*Verification of ISS, iISS and IOSS properties applying weighted homogeneity*, in "Systems & Control Letters", 2013, vol. 62, p. 1159–1167

[73] D. EFIMOV, W. PERRUQUETTI.*On conditions of oscillations and multi-homogeneity*, in "Mathematics of Control, Signals, and Systems", 2015, vol. 28, n^{o} 1, p. 1–37, http://dx.doi.org/10.1007/s00498-015-0157-y

[74] D. EFIMOV, W. PERRUQUETTI, J.-P. RICHARD.*Development of Homogeneity Concept For Time-Delay Systems*, in "SIAM Journal on Optimization and Control", 2014, vol. 52, n^{o} 3, p. 1403–1808

[75] D. EFIMOV, A. POLYAKOV, A. LEVANT, W. PERRUQUETTI.*Realization and Discretization of Asymptotically Stable Homogeneous Systems*, in "IEEE Trans. Automatic Control", 2017, vol. 62, n^{o} 11, p. 5962–5969

[76] D. EFIMOV, A. POLYAKOV, W. PERRUQUETTI, J.-P. RICHARD.*Weighted Homogeneity for Time-Delay Systems: Finite-Time and Independent of Delay Stability*, in "IEEE Trans. Automatic Control", 2016, vol. 61, n^o 1, p. 210–215

[77] N. ESPITIA, A. POLYAKOV, D. EFIMOV, W. PERRUQUETTI.*Boundary time-varying feedbacks for fixed-time stabilization of constant-parameter reaction-diffusion systems*, in "Automatica", 2019, vol. 103, p. 398–407, https://doi.org/10.1016/j.automatica.2019.02.013

[78] M. FEINGESICHT, A. POLYAKOV, F. KERHERVÉ, J.-P. RICHARD.*SISO model-based control of separated flows: Sliding mode and optimal control approaches*, in "International Journal of Robust and Nonlinear Control", 2017, vol. 27, n^o 18, p. 5008-5027

[79] E. LEURENT, D. EFIMOV, T. RAÏSSI, W. PERRUQUETTI.*Interval Prediction for Continuous-Time Systems with Parametric Uncertainties*, in "Proc. 58th IEEE Conference on Decision and Control (CDC)", Nice, 2019

[80] F. LOPEZ-RAMIREZ, A. POLYAKOV, D. EFIMOV, W. PERRUQUETTI.*Finite-time and Fixed-time Observer Design: Implicit Lyapunov function approach*, in "Automatica", 2018, vol. 87, n^o 1, p. 52-60

[81] A. POLYAKOV, J.-M. CORON, L. ROSIER.*On Homogeneous Finite-Time Control for Linear Evolution Equation in Hilbert Space*, in "IEEE Transactions on Automatic Control", 2018

[82] A. POLYAKOV, D. EFIMOV, B. BROGLIATO.*Consistent Discretization of Finite-time and Fixed-time Stable Systems*, in "SIAM Journal on Optimization and Control", 2019, vol. 57, n^o 1, p. 78–103

[83] A. POLYAKOV, D. EFIMOV, E. FRIDMAN, W. PERRUQUETTI.*On Homogeneous Distributed Parameter Systems*, in "IEEE Trans. Automatic Control", 2016, vol. 61, n^o 11, p. 3657–3662

[84] A. POLYAKOV, D. EFIMOV, W. PERRUQUETTI.*Finite-time and fixed-time stabilization: Implicit Lyapunov function approach*, in "Automatica", 2015, vol. 51, p. 332-340

[85] A. POLYAKOV.*Nonlinear feedback design for fixed-time stabilization of linear control systems*, in "IEEE Transactions on Automatic Control", 2012, vol. 57(8), p. 2106-2110

[86] H. RÍOS, D. EFIMOV, J. A. MORENO, W. PERRUQUETTI, J. G. RUEDA-ESCOBEDO.*Time-Varying Parameter Identification Algorithms: Finite and Fixed-Time Convergence*, in "IEEE Transactions on Automatic Control", 2017, vol. 62, n^o 7, p. 3671–3678, https://dx.doi.org/10.1109/TAC.2017.2673413

[87] R. USHIROBIRA, D. EFIMOV, P.-A. BLIMAN.*Estimating the infection rate of a SIR epidemic model via differential elimination*, in "Proceedings of ECC", Naples, 2019

[88] R. USHIROBIRA, D. EFIMOV, G. CASIEZ, N. ROUSSEL, W. PERRUQUETTI.*A forecasting algorithm for latency compensation in indirect human-computer interactions*, in "Proceedings of ECC", Alborg, 2016, p. 1081–1086

[89] K. ZIMENKO, D. EFIMOV, A. POLYAKOV, W. PERRUQUETTI.*A note on delay robustness for homogeneous systems with negative degree*, in "Automatica", 2017, vol. 79, n^o 5, p. 178–184