RESEARCH CENTRE

Inria Centre at Université Côte d'Azur

2024 ACTIVITY REPORT

Project-Team NEO

Network Engineering and Operations

DOMAIN

Networks, Systems and Services, Distributed Computing

THEME

Networks and Telecommunications

Inria

Contents

Project-Team NEO

Creation of the Project-Team: 2017 December 01

Keywords

Computer sciences and digital sciences

A1.1.11. – Quantum architectures
A1.2.4. – QoS, performance evaluation
A1.2.5. – Internet of things
A1.2.6. – Sensor networks
A1.3. – Distributed Systems
A1.3.1. – Web
A1.3.4. – Peer to peer
A1.5. – Complex systems
A1.5.1. – Systems of systems
A1.5.2. – Communicating systems
A3.1.2. – Data management, quering and storage
A3.3.3. – Big data analysis
A3.4. – Machine learning and statistics
A3.4.2. – Unsupervised learning
A3.4.3. – Reinforcement learning
A3.4.4. – Optimization and learning
A3.4.5. – Bayesian methods
A3.5. – Social networks
A3.5.2. – Recommendation systems
A4.1. – Threat analysis
A4.8. – Privacy-enhancing technologies
A5.9. – Signal processing
A5.9.2. – Estimation, modeling
A5.9.4. – Signal processing over graphs
A5.9.6. – Optimization tools
A6.1.1. – Continuous Modeling (PDE, ODE)
A6.1.2. – Stochastic Modeling
A6.2.2. – Numerical probability
A6.2.3. – Probabilistic methods
A6.2.6. – Optimization
A6.4.1. – Deterministic control
A6.4.2. – Stochastic control

A6.4.6. – Optimal control

A7.1. – Algorithms

A7.1.1. – Distributed algorithms

A7.1.2. – Parallel algorithms

A7.1.4. – Quantum algorithms

A8.1. – Discrete mathematics, combinatorics

A8.2.1. – Operations research

A8.6. – Information theory

A8.8. – Network science

A8.9. – Performance evaluation

A8.11. – Game Theory

A9.2. – Machine learning

A9.6. – Decision support

A9.7. – AI algorithmics

A9.9. – Distributed AI, Multi-agent

Other research topics and application domains

B2.3. – Epidemiology

B2.5.1. – Sensorimotor disabilities

B3.1. – Sustainable development

B3.1.1. – Resource management

B4.3.4. – Solar Energy

B4.4. – Energy delivery

B4.4.1. – Smart grids

B4.5.1. – Green computing

B6. – IT and telecom

B6.2. – Network technologies

B6.2.1. – Wired technologies

B6.2.2. – Radio technology

B6.3.3. – Network Management

B6.3.4. – Social Networks

B6.4. – Internet of things

B6.6. – Embedded systems

B8.1. – Smart building/home

B9.2.1. – Music, sound

B9.5.1. – Computer science

B9.5.2. – Mathematics

B9.6.3. – Economy, Finance

B9.6.4. – Management science

B9.6.5. – Sociology

1 Team members, visitors, external collaborators

Research Scientists

- Alain Jean-Marie [Team leader, INRIA, Senior Researcher]
- Sara Alouf [INRIA, Researcher]
- Eitan Altman [INRIA, Senior Researcher]
- Konstantin Avrachenkov [INRIA, Senior Researcher]
- Samir Medina Perlaza [INRIA, Researcher]
- Philippe Nain [INRIA, Emeritus]
- Giovanni Neglia [INRIA, Senior Researcher]

Post-Doctoral Fellows

- Khushboo Agarwal [INRIA, Post-Doctoral Fellow, from May 2024]
- Emmanouil Marios Athanasakos [INRIA, Post-Doctoral Fellow, until Jan 2024]
- Diego Goldsztajn [INRIA, Post-Doctoral Fellow]
- Ashok Krishnan Komalan Sindhu [INRIA, Post-Doctoral Fellow, until Jan 2024]
- Hariprasad Manjunath Hegde [INRIA, Post-Doctoral Fellow, until May 2024]
- Charlotte Rodriguez [INRIA, Post-Doctoral Fellow, until Nov 2024]

PhD Students

- Younes Ben Mazziane [Univ Côte d'Azur, until Oct 2024]
- Yaiza Bermudez [INRIA, from Nov 2024]
- José Francisco Daunas Torres [Univ Sheffield]
- Ibtihal El Mimouni [NSP-SmartProfile, CIFRE]
- Louis Hauseux [Univ Côte d'Azur]
- Caelin Kaplan [SAP Labs France, CIFRE, until Nov 2024]
- Ahmad Nasser [NOKIA, CIFRE]
- Angelo Rodio [INRIA, until Jul 2024]
- Julian Alfonso Santos Bustos [ORANGE, CIFRE]
- Xufeng Zhang [INRIA]
- Xinying Zou [INRIA]

Interns and Apprentices

- Albachiara Bellaroba [INRIA, Intern, from Mar 2024 until Sep 2024]
- Maysae Hmamouchi [Polytech Nice Sophia, Intern, until Feb 2024]

Administrative Assistant

- Jane Desplanques [INRIA]

Visiting Scientists

- Madhu Madhu [IIT Bombay, from Sep 2024]
- Daniel Richards Ravi Arputharaj [Freelance, from Sep 2024]
- Lucas Siviero Sibemberg [UFRGS, from Mar 2024 until Oct 2024]
- Jacopo Talpini [Univ Milan-Biccoca, until Apr 2024]

2 Overall objectives

NEO is an Inria project-team whose members are located in Sophia Antipolis (S. Alouf, K. Avrachenkov, G. Neglia, and S. M. Perlaza), in Avignon (E. Altman) and in Montpellier (A. Jean-Marie). E. Altman is also with the LINCS (Lab. for Information, Networking and Communication Sciences). S. M. Perlaza is also with the ECE department at Princeton Univ., N.J. USA; and the Mathematics Department of the Univ. de la Polynésie française (Laboratoire GAATI), Faaa, Tahiti.

The team is positioned at the intersection of Operations Research and Network Science. By using the tools of Stochastic Operations Research, we model situations arising in several application domains, involving networking in one way or the other. The aim is to understand the rules and the effects in order to influence and control them so as to engineer the creation and the evolution of complex networks.

3 Research program

The problems studied in NEO involve generally optimization, dynamic systems or randomness, and often all at the same time. The techniques we use to tackle these problems are those of Stochastic Operations Research, Applied Probabilities and Information Theory.

Stochastic Operations Research is a collection of modeling, optimization and numerical computation techniques, aimed at assessing the behavior of man-made systems driven by random phenomena, and at helping to make decisions in such a context.

The discipline is based on applied probability and focuses on effective computations and algorithms. Its core theory is that of Markov chains over discrete state spaces. This family of stochastic processes has, at the same time, a very large modeling capability and the potential of efficient solutions. By "solution" is meant the calculation of some *performance metric*, usually the distribution of some random variable of interest, or its average, variance, etc. This solution is obtained either through exact "analytic" formulas, or numerically through linear algebra methods. Even when not analytically or numerically tractable, Markovian models are always amenable to "Monte-Carlo" simulations with which the metrics can be statistically measured.

An example of this is the success of classical Queueing Theory, with its numerous analytical formulas. Another important derived theory is that of the Markov Decision Processes, which allows to formalize *optimal* decision problems in a random environment. This theory allows to characterize the optimal decisions, and provides algorithms for calculating them.

Strong trends of Operations Research are: a) an increasing importance of multi-criteria multi-agent optimization, and the correlated introduction of Game Theory in the standard methodology; b) an increasing concern of (deterministic) Operations Research with randomness and risk, and the consequent introduction of topics like Chance Constrained Programming and Stochastic Optimization. Data analysis is also more and more present in Operations Research: techniques from statistics, like filtering and estimation, or Artificial Intelligence like clustering, are coupled with modeling in Machine Learning techniques like Q-Learning.

4 Application domains

4.1 Network Science

Network Science is a multidisciplinary body of knowledge, principally concerned with the emergence of global properties in a network of individual agents. These global properties emerge from "local" properties of the network, namely, the way agents interact with each other. The central model of "networks" is the graph (of Graph Theory/Operations Research). Nodes represent the different entities managing information and taking decisions, whereas, links represent the fact that entities interact, or not. Links are usually equipped with a "weight" that measures the intensity of such interaction. Adding evolution rules to this quite elementary representation leads to dynamic network models, the properties of which Network Science tries to analyze.

A classical example of properties sought in networks is the famous "six degrees of separation" (or "small world") property: how and why does it happen so frequently? Another ubiquitous property of real-life networks is the Zipf or "scale-free" degree distribution. Some of these properties, when properly exploited, lead to successful business opportunities: just consider the PageRank algorithm of Google, which miraculously connects the relevance of some Web information with the relevance of the other information that points to it.

4.2 Network Engineering

In its primary acceptation, Network Science involves little or no engineering: phenomena are assumed to be "natural" and emerge without external interventions. However, the idea comes fast to intervene in order to modify the outcome of the phenomena. This is where NEO is positioned. Beyond the mostly descriptive approach of Network Science, we aim at using the techniques of Operations Research so as to engineer complex networks.

To quote two examples: controlling the spread of diseases through a "network" of people is of primarily interest for mankind. Similarly, controlling the spread of information or reputation through a social network is of great interest in the Internet. Precisely, given the impact of web visibility on business income, it is tempting (and quite common) to manipulate the graph of the web by adding links so as to drive the PageRank algorithm to a desired outcome.

Another interesting example is the engineering of community structures. Recently, thousands of papers have been written on the topic of community *detection* problem. In most of the works, the researchers propose methods, most of the time, heuristics, for detecting communities or dense subgraphs inside a large network. Much less effort has been put in the understanding of community formation process and even much less effort has been dedicated to the question of how one can influence the process of community formation, e.g. in order to increase overlap among communities and reverse the fragmentation of the society.

Our ambition for the medium term is to reach an understanding of the behavior of complex networks that will make us capable of influencing or producing a certain property in a given network. For this purpose, we will develop families of models to capture the essential structure, dynamics, and uncertainty of complex networks. The "solution" of these models will provide the correspondence between metrics of interest and model parameters, thus opening the way to the synthesis of effective control techniques.

In the process of tackling real, very large size networks, we increasingly deal with large graph data analysis and the development of decision techniques with low algorithmic complexity, apt at providing answers from large datasets in reasonable time.

5 Social and environmental responsibility

5.1 Impact of research results

Some of NEO's research is devoted to environmental issues, either related to water management, or to the exploitation of renewable resources, *e.g.* [24]. The involvement in the (ongoing) Chile-funded project MICCHI (Section 10.1.3) is aimed at connecting these theoretical results with actual drought problems.

Several research actions are directly aimed at reducing carbon/energy footprint in the IT sector. This is the case in the FedMalin Inria challenge (Section 10.4), and in the “Energy-Aware Federated Learning” project with Accenture (Section 9.1.1). The current output is the survey [63], and the team is in the final stages of research on algorithms aimed at reducing the carbon footprint of cross-silo federated learning.

6 Highlights of the year

6.1 Awards

Louis Hauseux won the second prize of the Student Research Competition (graduate category) at the SIGMETRICS/PERFORMANCE 2024 conference [41]; this prize led also to the Université Côte d’Azur (UniCA) "Prix d’excellence 2024”.

6.2 Courtesy Appointments

Samir Medina Perlaza was re-appointed “Visiting Research Collaborator” in the Department of Electrical and Computer Engineering at Princeton University for the academic year 2024-2025. He was also re-appointed "Associate Researcher" in the Laboratory of Algebraic Geometry and Applications to Information Theory (GAATI) at the Université de la Polynésie Française for the academic year 2024-2025.

7 New software, platforms, open data

7.1 New software

7.1.1 marmote

Name: MARkovian MOdeling: The Environment

Keyword: Markov model

Functional Description: marmote is a C++ library for modeling with Markov chains. It consists in a reduced set of high-level abstractions for constructing state spaces, transition structures and Markov chains (discrete-time and continuous-time). It provides the ability of constructing hierarchies of Markov models, from the most general to the particular, and equip each level with specifically optimized solution methods. The current release features the library marmoteMDP for modeling Markov Decision Processes and solving them.

This software was started within the ANR MARMOTE project: ANR-12-MONU-00019 under the name marmoteCore. Within the marmote project, the code conforms the latest C++ standards and the library is available on multiple platforms via a conda distribution.

Release Contributions: Version 1.2.0 was released in 2024. This version further improves the code portability and introduces structural changes improving code maintainability. From the functional point of view, it introduces a Python interface to the library: pyMarmote for the core, and pyMarmoteMDP for the library for manipulating Markov Decision Processes. A complete set of examples is available on the project’s website, both for the C++ and the Python languages. For the latter, a notebook-based tutorial demonstrates the principal functionalities of the library.

News of the Year: The software was awarded an Inria ADT (Action of Technological Development). Thanks to the software engineers of Inria’s SED (Service Experimentation Development), a Python interface was built for the core and for the application library devoted to Markov Decision Processes. Many improvements have been brought to internal self-testing and to programming examples, in particular via Python tutorial notebooks.

URL: https://marmote.gitlabpages.inria.fr/marmote/

Publications: hal-03770430v1, hal-03781620v1, hal-02395100v1, hal-04176076v1

Contact: Alain Jean-Marie

Participants: Alain Jean-Marie, Patrick Brown, Emmanuel Hyon

Partner: Université Paris Nanterre

7.2 Open data

Flamby.

Contributors: Giovanni Neglia and Othmane Marfoq were part of a large international consortium, coordinated by the company Owkin and including EPFL, FeML Inc., University of Southern California, École Polytechnique, Institut Polytechnique de Paris, University Hospital Bonn, Helmholtz Munich, University of California at Berkeley and other Inria teams (EPIONE and MAGNET).

Description: Federated Learning (FL) enables several clients holding sensitive data to collaboratively train machine learning models, without centralizing data. The cross-silo FL setting corresponds to the case of few (2–50) reliable clients, each holding medium to large datasets, and is typically found in applications such as healthcare, finance, or industry. While previous works have proposed representative datasets for cross-device FL, few realistic healthcare cross-silo FL datasets exist, thereby slowing algorithmic research in this critical application. FLamby (Federated Learning AMple Benchmark of Your cross-silo strategies) is a novel cross-silo dataset suite focused on healthcare to bridge the gap between theory and practice of cross-silo FL. FLamby encompasses 7 healthcare datasets with natural splits, covering multiple tasks, modalities, and data volumes, each accompanied with baseline training code. It also includes benchmark standard FL algorithms for all datasets.

Project link: https://github.com/owkin/flamby

Publications: Ogier du Terrail et al, FLamby: Datasets and Benchmarks for Cross-Silo Federated Learning in Realistic Healthcare Settings, Advances in Neural Information Processing Systems (NeurIPS), 2022 https://inria.hal.science/hal-03900026.

Contact: Jean Ogier Du Terrail jean.du-terrail@owkin.com

Release contributions: NEO members contributed to defining the benchmark, developing the code, and writing the corresponding paper.

Similarity caching traces.

Contributors: Abdelkarim Hafid (previously master intern in NEO), Caelin Kaplan, Giovanni Neglia, Anirudh Sabnis (previously visiting PhD student in NEO), Tareq Si Salem (previously PhD student in NEO).

Description: A similarity cache can return an item that is similar, but not necessarily identical, to the query element. Similarity caching systems can be effectively employed in several application areas, like multimedia retrieval, recommender systems, genome study, and machine learning training/serving. We produced and released four distinct traces to evaluate similarity caching policies. Each trace includes a catalog of items and a time series of requests, with both items and requests represented as real vectors. In three datasets, these representations were derived from machine learning models, while in the fourth, the items correspond to tiles in a 360° video, and each request represents the tiles a user is watching.

Project link: https://sim-cache.gitlabpages.inria.fr/similarity-caching-traces/

Publications:
- A. Sabnis, T. Si Salem, G. Neglia, M. Garetto, E. Leonardi and R. K. Sitaraman, GRADES: Gradient Descent for Similarity Caching, IEEE INFOCOM 2021 https://inria.hal.science/hal-03484133

- A. Sabnis, T. Si Salem, G. Neglia, M. Garetto, E. Leonardi and R. K. Sitaraman, GRADES: Gradient Descent for Similarity Caching, IEEE/ACM Transactions on Networking, Vol 31, n 1, 2023 https://inria.hal.science/hal-03906099

Contact: Giovanni Neglia.

Release contributions: The NEO researchers were at the origin of this project and carried it out in collaboration with close partners. Three of the four traces were derived from existing datasets, where a machine learning model implicitly defines item embeddings within a space equipped with a similarity metric. The fourth trace is synthetic, generated using a Markovian model to simulate user navigation in a 360° video.

8 New results

8.1 Stochastic Modeling

Participants: Konstantin Avrachenkov, Alain Jean-Marie, Diego Goldsztajn, Samir Medina Perlaza.

8.1.1 Learning and Balancing Unknown Loads in Large-Scale Systems

In [21], Diego Goldsztajn in collaboration with Sem Borst (Technical University of Eindhoven, the Netherlands) and Johan van Leeuwaarden (Tilburg University, the Netherlands) considered a system of identical server pools where tasks with exponentially distributed service times arrive as a time-inhomogeneous Poisson process. An admission threshold is used in an inner control loop to assign incoming tasks to server pools, while in an outer control loop, a learning scheme adjusts this threshold over time to keep it aligned with the unknown offered load of the system. In a many-server regime, they prove that the learning scheme reaches an equilibrium along intervals of time when the normalized offered load per server pool is suitably bounded and that this results in a balanced distribution of the load. Furthermore, they established a similar result when tasks with Coxian distributed service times arrive at a constant rate and the threshold is adjusted using only the total number of tasks in the system. The novel proof technique developed in this paper, which differs from a traditional fluid limit analysis, allowed them to handle rapid variations of the first learning scheme, triggered by excursions of the occupancy process that have vanishing size. Moreover, our approach allowed them to characterize the asymptotic behavior of the system with Coxian distributed service times without relying on a fluid limit of a detailed state descriptor.

8.1.2 Server Saturation in Skewed Networks

In [40, 22], Diego Goldsztajn in collaboration with Sem Borst (Technical University of Eindhoven, the Netherlands) and Johan van Leeuwaarden (Tilburg University, the Netherlands) considered a model inspired by compatibility constraints that arise between tasks and servers in data centers, cloud computing systems and content delivery networks. The constraints are represented by a bipartite graph or network that interconnects dispatchers with compatible servers. Each dispatcher receives tasks over time and sends every task to a compatible server with the least number of tasks, or to a server with the least number of tasks among d compatible servers selected uniformly at random. They focus on networks where the neighborhood of at least one server is skewed in a limiting regime. This means that a diverging number of dispatchers are in the neighborhood which are each compatible with a uniformly bounded number of servers; thus, the degree of the central server approaches infinity while the degrees of many neighboring dispatchers remain bounded. They proved that each server with a skewed neighborhood saturates, in the sense that the mean number of tasks queueing in front of it in steady state approaches infinity. Paradoxically, this pathological behavior can even arise in random networks where nearly all the servers have at most one task in the limit.

8.1.3 The Granger–Johansen Representation Theorem for Integrated Time Series on Banach Space

In [23], Konstantin Avrachenkov in collaboration with Phil Howlett (University of South Australia, Australia), Brendan Beare (University of Sydney, Australia), Massimo Franchi (Sapienza Universita di Roma, Italy) and John Boland (University of South Australia, Australia) proved an extended Granger-Johansen Representation Theorem (GJRT) for finite—or infinite—order integrated autoregressive time series on Banach space. They assumed only that the resolvent of the autoregressive polynomial for the series is analytic on and inside the unit circle except for an isolated singularity at unity. If the singularity is a pole of finite order, the time series is integrated of the same order. If the singularity is an essential singularity, the time series is integrated of order infinity. When there is no deterministic forcing, the value of the series at each time is the sum of an almost surely convergent stochastic trend, a deterministic term depending on the initial conditions and a finite sum of embedded white noise terms in the prior observations. This is the extended GJRT. In each case the original series is the sum of two separate autoregressive time series on complementary subspaces—a singular component which is integrated of the same order as the original series and a regular component which is not integrated. The extended GJRT applies to all integrated autoregressive processes irrespective of the spatial dimension, the number of stochastic trends and cointegrating relations in the system and the order of integration.

8.1.4 An Information Theoretic Metric for Measurement Vulnerability to Data Integrity Attacks on Smart Grids

In [28], Samir Medina Perlaza, Xiuzhen Ye, Iñaki Esnaola, and Robert F. Harrison (University of Sheffield, UK) studied a novel metric that describes the vulnerability of the measurements in power systems to data integrity attacks. The new metric, coined vulnerability index (VuIx), leverages information theoretic measures to assess the attack effect in terms of the fundamental limits of the disruption and detection tradeoff. The result of computing the VuIx of the measurements in the system yields an ordering of their vulnerability based on the degree of exposure to data integrity attacks. This new framework is used to assess the measurement vulnerability of IEEE 9-bus and 30-bus test systems and it is observed that power injection measurements are significantly more vulnerable to data integrity attacks than power flow measurements. A detailed numerical evaluation of the VuIx values for IEEE test systems is provided.

8.1.5 Development of pyMarmote and pyMarmoteMDP

The development of the `Marmote` library has continued, led by Alain Jean-Marie and Emmanuel Hyon of Univ. Paris-Nanterre. The software was awarded an Inria ADT (Action of Technological Development), during which several software engineers of Inria's SED (Service Experimentation Development) helped improve the robustness and portability of the library. Most of all, they endowed the library with a Python interface, named `pyMarmote` for the core, and `pyMarmoteMDP` for the application library devoted to Markov Decision Processes. A tutorial providing an introduction to the capabilities of `Marmote` via its Python interfaces was presented by Alain Jean-Marie and Emmanuel Hyon in [65]. It is now part of the online documentation of the package.

8.2 Theory of Learning

Participants: Eitan Altman, Konstantin Avrachenkov, José Francisco Daunas Torres, Diego Goldsztajn, Alain Jean-Marie, Samir Medina Perlaza, Xinying Zou.

8.2.1 Recovering Static and Time-Varying Communities Using Persistent Edges

In [13], Konstantin Avrachenkov in collaboration with Maximilien Dreveton (EPFL, Switzerland) and Lasse Leskelä (Aalto University, Finland) developed spectral methods for recovering communities in temporal networks. In the case of fixed communities, spectral clustering on the simple time-aggregated graph (i.e., the weighted graph formed by the sum of the interactions over all temporal snapshots) does not always

produce satisfying results. To utilize information carried by temporal correlations, they proposed to employ different weights on freshly appearing and persistent edges. They showed that spectral clustering on such weighted graphs can be explained as a relaxation of the maximum likelihood estimator of an extension of the degree-corrected stochastic block model with Markov interactions. They also studied the setting of evolving communities, for which they used the prediction at time $t-1$ as an oracle for inferring the community labels at time t. They demonstrated the accuracy of the proposed methods on synthetic and real data sets.

8.2.2 Almost Exact Recovery in Noisy Semi-supervised Learning

Graph-based semi-supervised learning methods combine the graph structure and labeled data to classify unlabeled data. In [12], Konstantin Avrachenkov and Maximilien Dreveton (EPFL, Switzerland) studied the effect of a noisy oracle on classification. In particular, they derived the maximum a posteriori (MAP) estimator for clustering a degree corrected stochastic block model when a noisy oracle reveals a fraction of the labels. They then proposed an algorithm derived from a continuous relaxation of the MAP, and established its consistency. Numerical experiments show that the proposed approach achieves promising performance on synthetic and real data sets, even in the case of very noisy labeled data.

8.2.3 The Generalization Error of Machine Learning Algorithms

The generalization error is one of the standard metrics used to evaluate the generalization capabilities of machine learning algorithms. Selecting a good algorithm in terms of generalization can be framed as a problem of Empirical Risk Minimization (ERM). When this minimization problem is regularized by incorporating a penalty term based on relative entropy, the resulting unique solution is a Gibbs probability measure (a Gibbs algorithm), which possesses several notable properties. Interestingly, the worst-case data-generating probability measure under this framework is also a Gibbs probability measure. Exploiting this dual characterization of Gibbs measures has led to several significant results, as outlined below.

In [62], Samir Medina Perlaza and Xinying Zou introduced the method of gaps, a novel technique for deriving closed-form expressions in terms of information measures for the generalization error of machine learning algorithms. The method relies on two central observations: 1) the generalization error is an average of the variation of the expected empirical risk with respect to changes on the probability measure (used for expectation); and 2) these variations, also referred to as gaps, exhibit closed-form expressions in terms of information measures. The expectation of the empirical risk can be either with respect to a measure on the models (with a fixed dataset) or with respect to a measure on the datasets (with a fixed model), which results in two variants of the method of gaps. The first variant, which focuses on the gaps of the expected empirical risk with respect to a measure on the models, appears to be the most general, as no assumptions are made on the distribution of the datasets. The second variant develops under the assumption that datasets are made of independent and identically distributed data points. All existing exact expressions for the generalization error of machine learning algorithms can be obtained with the proposed method. Also, this method allows obtaining numerous new exact expressions, which improves the understanding of the generalization error; establishes connections with other areas in statistics, e.g., hypothesis testing; and potentially, might guide algorithm designs.

In [25], Samir Medina Perlaza, Alain Jean-Marie, and Gaetan Bisson (Univ. Polynésie française), Iñaki Esnaola (University of Sheffield, UK), and Stefano Rini (National Chiao Tung University, Taiwan) studied the ERM problem with relative entropy regularization (ERM-RER) under the assumption that the reference measure is a σ-finite measure, and not necessarily a probability measure. Under this assumption, which leads to a generalization of the ERM-RER problem allowing a larger degree of flexibility for incorporating prior knowledge, numerous relevant properties are stated. Among these properties, the solution to this problem, if it exists, is shown to be a unique probability measure, mutually absolutely continuous with the reference measure. Such a solution, which is a Gibbs probability measure, exhibits a probably-approximately-correct guarantee for the ERM problem independently of whether the latter possesses a solution. For a fixed dataset and under a specific condition, the empirical risk is shown to be a sub-Gaussian random variable when the models are sampled from the solution to the ERM-RER problem. The generalization capabilities of the solution to the ERM-RER problem (the Gibbs algorithm) are studied via the sensitivity of the expected empirical risk to deviations from such a solution towards alternative

probability measures. Finally, an interesting connection between sensitivity, generalization error, and lautum information is established.

In [57], José Francisco Daunas Torres and Samir Medina Perlaza, along with Iñaki Esnaola (University of Sheffield, UK) and H. Vincent Poor (Princeton University, USA), analyzed the effect of relative entropy asymmetry in the context of ERM-RER. Two regularizations are considered:

1. the relative entropy of the measure to be optimized with respect to a reference measure (Type-I ERM-RER); and

2. the relative entropy of the reference measure with respect to the measure to be optimized (Type-II ERM-RER).

The main result is the characterization of the solution to the Type-II ERM-RER problem and its key properties. By comparing the well-understood Type-I ERM-RER with Type-II ERM-RER, the effects of entropy asymmetry are highlighted. The analysis shows that in both cases, regularization by relative entropy forces the solution's support to collapse into the support of the reference measure, introducing a strong inductive bias that can overshadow the evidence provided by the training data. Finally, it is shown that Type-II regularization is equivalent to Type-I regularization with an appropriate transformation of the empirical risk function.

In [36], the same authors studied the solution to ERM with f-divergence regularization (ERM-fDR) under mild conditions on f. Under such conditions, the optimal measure is shown to be unique. Examples of the solution for particular choices of the function f are presented. Previously known solutions to common regularization choices are obtained by leveraging the flexibility of the family of f-divergences. These include the unique solutions to empirical risk minimization with relative entropy regularization (Type-I and Type-II). Furthermore, the analysis of the solution unveils the following properties of f-divergences when used in the ERM-fDR problem:

- f-divergence regularization forces the support of the solution to coincide with the support of the reference measure, introducing a strong inductive bias that dominates the evidence provided by the training data; and

- any f-divergence regularization is equivalent to a different f-divergence regularization with an appropriate transformation of the empirical risk function.

In [29, 49], Xinying Zou, Samir Medina Perlaza, Eitan Altman, along with Iñaki Esnaola (University of Sheffield, UK), introduced the worst-case probability measure over the data as a tool for characterizing the generalization capabilities of machine learning algorithms. More specifically, the worst-case probability measure is a Gibbs probability measure and the unique solution to the maximization of the expected loss under a relative entropy constraint with respect to a reference probability measure. Fundamental generalization metrics, such as the sensitivity of the expected loss, the sensitivity of the empirical risk, and the generalization gap, are shown to have closed-form expressions involving the worst-case data-generating probability measure. Existing results for the Gibbs algorithm, such as characterizing the generalization gap as a sum of mutual information and lautum information, up to a constant factor, are recovered. A novel parallel is established between the worst-case data-generating probability measure and the Gibbs algorithm. Specifically, the Gibbs probability measure is identified as a fundamental commonality of the model space and the data space for machine learning algorithms.

In [64], the same authors and H. Vincent Poor (Princeton University, USA), studied the worst-case data-generating (WCDG) probability measure as a tool for characterizing the generalization capabilities of machine learning algorithms. Such a WCDG probability measure is shown to be the unique solution to two different optimization problems:

1. the maximization of the expected loss over the set of probability measures on the datasets whose relative entropy with respect to a *reference measure* is not larger than a given threshold; and

2. the maximization of the expected loss with regularization by relative entropy with respect to the reference measure.

Such a reference measure can be interpreted as a prior on the datasets. The WCDG cumulants are finite and bounded in terms of the cumulants of the reference measure. To analyze the concentration of the expected empirical risk induced by the WCDG probability measure, the notion of (ϵ,δ)-robustness of models is introduced. Closed-form expressions are presented for the sensitivity of the expected loss for a fixed model. These results lead to novel expressions for the generalization error of arbitrary machine learning algorithms. These expressions can be broadly divided into two classes. The first one involves a WCDG probability measure, while the second one involves a Gibbs algorithm. This finding reveals that an exploration into the generalization error of the Gibbs algorithm facilitates the derivation of overarching insights applicable to any machine learning algorithm.

8.2.4 Density-based Clustering

Many of clustering algorithms are based on density estimates in R^d. Building geometric graphs on the dataset X is an elegant way of doing this. In fact, the connected components of a geometric graph match exactly with the high-density clusters of the 1-Nearest Neighbor density estimator. In [42], Louis Hauseux, Konstantin Avrachenkov and Josiane Zerubia (AYANA) showed that the natural way to generalize geometric graphs is to use hypergraphs with a more restrictive notion of connected component called K-Polyhedron. Herein, they proved that K-Polyhedra correspond to high-density clusters of K-Nearest Neighbors density estimator. Furthermore, the percolation phenomenon is omnipresent behind the family of clustering algorithms they studied in this paper.

Many of clustering algorithms for a point cloud $X_n \subset R^d$ in the Euclidean space are based on density estimates. In fact, the density function f of point generation contains the relevant information. It is quite natural to try to extract what Hartigan called 'high-density clusters'. One elegant solution to do this task consists in constructing a graph whose nodes are the points of the cloud and whose edges connect nearby points. One wants the connected components of this graph to reflect the high-density clusters. Some very classical algorithms such as (Robust) Single-Linkage or (H)DBSCAN work in this way. It is particularly helpful because its connected components correspond exactly to the high-density clusters of the density estimator of the 1-Nearest Neighbor. An example of the Single-Linkage will show us the mathematical phenomenon at the heart of these algorithms: the percolation. In [41], Louis Hauseux defined and measured the percolation rate to evaluate the performance of such algorithms. By way of example, the Robust Single-Linkage algorithm is analyzed and its percolation rate is calculated. This showed theoretically why it is actually preferable to use K-Nearest Neighbours rather than 1-NN. However, convergence in K towards a perfect estimator is very slow, so this analysis explains why in practice $K = 10$ often represents a good trade-off.

In [43], Louis Hauseux, Konstantin Avrachenkov and Josiane Zerubia (AYANA) firstly studied how to measure the theoretical performance of hierarchical clustering algorithms which depend on a scale parameter; and secondly, improved the State-of-the-Art of the family of hierarchical clustering with respect to the introduced new measure. Single-Linkage is perhaps the simplest and the most famous algorithm belonging to this family of hierarchical clustering. Nowadays, the State-of-the-Art clustering algorithm, HDBSCAN, works in a similar way to Single-Linkage (with some refinements analyzed in this work). Specifically, they explained why the percolation phenomenon is omnipresent behind this family of clustering algorithms. Previously, they defined an index named "percolation rate" to measure the theoretical capability of algorithms to identify different high-density levels. In [43], they showed that using hypergraphs is a natural way to generalize Single-Linkage with higher-order interactions. New high-order connected components on hypergraphs, 'K-polyhedra', have much better percolation rates than the classic 'robustification' of Single-Linkage (used e.g. by HDBSCAN), the K-Robust Single-Linkage components. They investigate in detail the important cases of R^2 and R^3 for $K \in \{1,2,3\}$.

8.2.5 Reinforcement Learning for Networks and Weakly Coupled MDPs

The Whittle index policy is a heuristic that has shown remarkably good performance (with guaranteed asymptotic optimality) when applied to the class of problems known as Restless Multi-Armed Bandit Problems (RMABPs). In [27], Konstantin Avrachenkov in collaboration with Francisco Robledo and Urtzi Ayesta (IRIT) and Vivek Borkar (IITB, India) presented QWI and QWINN, two reinforcement learning algorithms, respectively tabular and deep, to learn the Whittle index for the total discounted criterion.

The key feature is the use of two time-scales, a faster one to update the state-action Q-values, and a relatively slower one to update the Whittle indices. As the main theoretical result they showed that QWI, which is a tabular implementation, converges to the real Whittle indices. They then presented QWINN, an adaptation of QWI algorithm using neural networks to compute the Q-values on the faster time-scale, which is able to extrapolate information from one state to another and scales naturally to large state-space environments. For QWINN, they showed that all local minima of the Bellman error are locally stable equilibria, which is the first result of its kind for DQN-based schemes. Numerical computations showed that QWI and QWINN converge faster than the standard Q-learning algorithm, neural-network based approximate Q-learning and other state of the art algorithms.

In [47], Konstantin Avrachenkov in collaboration with Francisco Robledo and Urtzi Ayesta (IRIT) introduced the Lagrange Policy for Continuous Actions (LPCA), a reinforcement learning algorithm specifically designed for weakly coupled Markov Decision Process (MDP) problems with continuous action spaces. LPCA addresses the challenge of resource constraints dependent on continuous actions by introducing a Lagrange relaxation of the weakly coupled MDP problem within a neural network framework for Q-value computation. This approach effectively decouples the MDP, enabling efficient policy learning in resource-constrained environments. They presented two variations of LPCA: LPCA-DE, which utilizes differential evolution for global optimization, and LPCA-Greedy, a method that incrementally and greadily selects actions based on Q-value gradients. Comparative analysis against other state-of-the-art techniques across various settings highlight LPCA's robustness and efficiency in managing resource allocation while maximizing rewards.

Influence maximization (IM) has been widely studied in recent decades, aiming to maximize the spread of influence over networks. Despite many works for static networks, fewer research studies have been dedicated to the IM problem for dynamic networks, which creates many challenges. An IM method for such an environment, should consider its dynamics and perform well under different network structures. To fulfill this objective, more computations are required. Hence, an IM approach should be efficient enough to be applicable for the ever-changing structure of a network. In [17], Konstantin Avrachenkov and Kishor Patil in collaboration with Haleh Dizaji (University of Klagenfurt, Austria) proposed an IM method for dynamic networks which uses a deep Q-learning (DQL) approach. To learn dynamic features from the network and retain previously learned information, incremental and transfer learning methods have been applied. Experiments substantiate the good performance of the DQL methods and their superiority over compared methods on larger sizes of tested synthetic and real-world networks. These experiments illustrate better performance for incremental and transfer learning methods on real-world networks.

8.3 Distributed Learning

Participants: Caelin Kaplan, Othmane Marfoq, Giovanni Neglia, Angelo Rodio, Charlotte Rodriguez.

Federated Learning (FL) enables multiple clients, such as mobile phones and IoT devices, to collaboratively train a global machine learning model while keeping their data localized. During the last year, much activity has been devoted to FL, which has also been the subject of Angelo Rodio's PhD thesis [54].

FL algorithms, such as FedAvg, are negatively affected by data heterogeneity and partial client participation. To mitigate the latter problem, global variance reduction methods, like FedVARP, leverage stale model updates for non-participating clients. These methods are effective under homogeneous client participation. Yet, in [48], Angelo Rodio and Giovanni Neglia show that, when some clients participate much less than others, aggregating updates with different levels of staleness can detrimentally affect the training process. They introduce FedStale, a novel algorithm that updates the global model in each round through a convex combination of "fresh" updates from participating clients and "stale" updates from non-participating ones. By adjusting the weight in the convex combination, FedStale interpolates between FedAvg, which only uses fresh updates, and FedVARP, which treats fresh and stale updates equally. Their analysis of FedStale convergence yields novel findings: i) it integrates and extends previous FedAvg and FedVARP analyses to heterogeneous client participation; ii) it underscores how the least participating client influences convergence error; iii) it provides practical guidelines to best exploit stale

updates, showing that their usefulness diminishes as data heterogeneity decreases and participation heterogeneity increases. Extensive experiments featuring diverse levels of client data and participation heterogeneity not only confirm these findings but also show that FedStale outperforms both FedAvg and FedVARP in many settings.

In [34], Giovanni Neglia in collaboration with Fabio Busacca and Sergio Palazzo (Univ. of Catania, Italy), Stefano Mangione and Ilenia Tinnirello (Univ. of Palermo, Italy), and Francesco Restuccia (Northeastern Univ., USA) study FL application to spectrum sensing in Long Range Wide Area Networks (LoRaWANs). Effective spectrum sensing is quintessential to decrease spectrum congestion across time, space and frequency in Internet of Things (IoT) networks. To circumvent the severe bandwidth constraints of IoT networks, federated machine learning (FML) can be used, but it is still unclear whether FML can be successfully performed in resource-constrained bandwidth-limited IoT networks. In this paper, they demonstrate that FML can tolerate losses up to a certain percentage and still converge. Then, they leverage this key result to design FedLoRa, an optimization framework for LoRaWANs that is (i) fast, as it reduces the FML round time in comparison with other resource allocation schemes; (ii) energy-efficient, as the time reduction does not imply a higher energy consumption. The key idea is to balance the network load over the available spreading factors, and to exploit sequential polling of nodes to maximize the number of simultaneous non-interfering transmissions, leading to a shorter FML round time. As the problem is NP-Hard, they provide an approximation algorithm. They evaluate the performance of FedLoRa through experimental evaluation on the Colosseum channel emulator, as well as with real-world data collection with off-the-shelf LoRaWAN devices in an 5 km×5 km urban setting in Portland, Maine. Their results show that FedLoRa reduces the round time by up to about 35%, as compared to the baselines.

In [35], Giovanni Neglia, in collaboration with Marina Costantini (EURECOM) and Thrasyvoulos Spyropoulos (EURECOM and Technical University of Crete, Greece), present FedDec, an algorithm that interleaves peer-to-peer communication and parameter averaging between the local gradient updates of FL. They analyze the convergence of FedDec and show that inter-agent communication alleviates the negative impact of infrequent communication rounds with the server by reducing the dependence on the number of local updates H from $O(H^2)$ to $O(H)$. Furthermore, their analysis reveals that the term improved in the bound vanishes quickly the more connected the network is. They confirm the predictions of their theory in numerical simulations, showing that FedDec converges faster than FedAvg and that the gains are greater as either H or the connectivity of the network increase.

In [45], Giovanni Neglia, in collaboration with Batiste Le Bars and Kevin Scaman (ARGO), Aurélien Bellet (PREMEDICAL), and Marc Tommasi (MAGNET), presents a new generalization error analysis for Decentralized Stochastic Gradient Descent (D-SGD) based on algorithmic stability. Their results challenge recent works that suggested increased instability due to decentralization and a detrimental impact of poorly-connected communication graphs on generalization. Contrary to these suggestions, they demonstrate that for convex, strongly convex, and non-convex functions, D-SGD can recover generalization bounds analogous to those of classical SGD, indicating that the choice of graph does not adversely affect generalization. They argue that this finding stems from a worst-case analysis and provide a refined optimization-dependent generalization bound for general convex functions. This new bound reveals that the choice of graph can, in fact, improve the worst-case bound in certain regimes, and, surprisingly, a poorly-connected graph can even be beneficial for generalization.

Collaborative learning, like in FL, introduces a bias-variance trade-off when client local data distributions differ. A key challenge is for each agent to identify clients with similar distributions while learning the model, a problem that remains largely unresolved. In [39], Giovanni Neglia, in collaboration with Franco Galante and Emilio Leonardi (Polytecnic Univ. of Turin, Italy), addresses this issue by focusing on a simplified version of the overarching problem, where each agent collects samples from a real-valued distribution over time to estimate its mean. Existing algorithms face impractical space and time complexities (quadratic in the number of agents A). To address scalability challenges, they propose a framework where agents self-organize into a graph, allowing each agent to communicate with only a selected number of peers r. They introduce two collaborative mean estimation algorithms: one draws inspiration from belief propagation, while the other employs a consensus-based approach, with complexities of $O(r|A|\log|A|)$ and $O(r|A|)$, respectively. They establish conditions under which both algorithms yield asymptotically optimal estimates and offer a theoretical characterization of their performance.

As Internet of Things (IoT) technology advances, end devices like sensors and smartphones are progressively equipped with AI models tailored to their local memory and computational constraints. Local

inference reduces communication costs and latency; however, these smaller models typically underperform compared to more sophisticated models deployed on edge servers or in the cloud. Cooperative Inference Systems (CISs) address this performance trade-off by enabling smaller devices to offload part of their inference tasks to more capable devices. These systems often deploy hierarchical models that share numerous parameters, exemplified by deep neural networks that utilize strategies like early exits or ordered dropout. FL may be employed to jointly train the models within a CIS. In an ongoing work [61], Caelin Kaplan, Angelo Rodio, and Giovanni Neglia, in collaboration with Tareq Si Salem (Huawei Paris Research Center, previously PhD student in NEO) and Chuan Xu (COATI), propose a novel FL approach designed explicitly for use in CISs that accounts for these variations in serving rates. Their framework not only offers rigorous theoretical guarantees but also surpasses state-of-the-art training algorithms for CISs, especially in scenarios where inference request rates or data availability are uneven among clients.

Recent studies reveal that the training phase of FL is vulnerable to reconstruction attacks, such as attribute inference attacks (AIA), where adversaries exploit exchanged messages and auxiliary public information to uncover sensitive attributes of targeted clients. While these attacks have been extensively studied in the context of classification tasks, their impact on regression tasks remains largely unexplored. In [37], Giovanni Neglia, in collaboration with Francesco Diana, Chuan Xu and Frédéric Giroire (COATI), Othmane Marfoq (META, USA, previously PhD student in NEO) and Eoin Thomas (Amadeus, France), proposes novel model-based AIAs specifically designed for regression tasks in FL environments. Their approach considers scenarios where adversaries can either eavesdrop on exchanged messages or directly interfere with the training process. They benchmark their proposed attacks against state-of-the-art methods using real-world datasets. The results demonstrate a significant increase in reconstruction accuracy, particularly in heterogeneous client datasets, a common scenario in FL. The efficacy of their model-based AIAs makes them better candidates for empirically quantifying privacy leakage for federated regression tasks.

In the realm of privacy-preserving machine learning, empirical privacy defenses have been proposed to achieve satisfactory levels of training data privacy without a significant drop in model utility. Most existing defenses against membership inference attacks assume access to reference data, defined as an additional dataset coming from the same (or a similar) underlying distribution as training data. Despite the common use of reference data, previous works are notably reticent about defining and evaluating reference data privacy. As gains in model utility and/or training data privacy may come at the expense of reference data privacy, it is essential that all three aspects are duly considered. In [44], Caelin Kaplan and Giovanni Neglia, in collaboration with Chuan Xu (COATI), Othmane Marfoq (META, previously a PhD student in NEO), and Anderson Santana de Oliveira (SAP Labs, France), conduct the first comprehensive analysis of empirical privacy defenses. They examine the availability of reference data and its privacy treatment in previous works, demonstrating its necessity for fairly comparing defenses. They propose a baseline defense that enables the utility-privacy tradeoff with respect to both training and reference data to be easily understood. Their method is formulated as an empirical risk minimization with a constraint on the generalization error, which, in practice, can be evaluated as a weighted empirical risk minimization (WERM) over the training and reference datasets. Although conceived as a simple baseline, their experiments show that WERM surprisingly outperforms the most well-studied and current state-of-the-art empirical privacy defenses using reference data for nearly all relative privacy levels of reference and training data. Their investigation also reveals that existing methods are unable to trade off reference data privacy for model utility and/or training data privacy, thus failing to operate outside of the high reference data privacy case. Overall, their work highlights the need for a proper evaluation of the triad "model utility / training data privacy / reference data privacy" when comparing privacy defenses.

In an ongoing work [63], Charlotte Rodriguez and Giovanni Neglia, in collaboration with Laura Degioanni, Laetitia Kameni, and Richard Vidal (Accenture Labs, France), address the challenge of evaluating the energy consumption of Machine Learning (ML) tasks. They note the absence of a universal tool suitable for all use cases and the lack of consensus on evaluation methods. To tackle this, they conduct a systematic literature review of tools and methods for assessing ML energy consumption during both training and inference, regardless of their original design purpose. Additionally, they develop an experimental protocol to compare a selection of these tools and methods, both qualitatively and quantitatively, across various ML tasks differing in nature (e.g., vision, language) and computational complexity. Their systematic literature review serves as a comprehensive guide for understanding the array of tools and methods used in evaluating ML energy consumption, catering to use cases ranging from basic energy

monitoring to consumption optimization. They provide two open-source repositories for further exploration: one containing tools to replicate or extend their review, and another housing the experimental protocol, allowing users to augment it with new ML computing tasks and additional energy evaluation tools.

8.4 Game Theory and Applications

Participants: Khushboo Agarwal, Eitan Altman, Emmanouil Marios Athanasakos, Alain Jean-Marie, Samir Medina Perlaza.

8.4.1 Games among selfish and team stations in polling systems

In [30], Khushboo Agarwal and Eitan Altman focus on a particular polling system known as the cyclic Bernoulli polling (CBP) system, where a server moves cyclically between the stations and serves the queue at a station with a certain probability when polled. Each station follows either a gated or partially exhaustive service discipline. In the steady state of such a system, they study a new game-theoretic aspect, where, the stations strategically choose the probability of accepting or rejecting the service from the server when polled. They examine three variants of non-cooperative games among stations: (i) each station selfishly minimizes its expected waiting time, (ii) a team game where each station minimizes the expected workload of the system, and (iii) stations act with partial cooperation, incurring an additional linear cost. They begin by presenting a new result for the CBP system regarding the continuity of expected waiting times in relation to the probabilities selected by the stations. For each game, they then investigate the existence and uniqueness of the Nash equilibrium (NE). In some cases, the NE is explicitly derived, while in others, characterizing the NE remains challenging due to the complex dependence of waiting times on the non-trivial buffer occupancy equations. Nonetheless, they analyze the NE and its properties through numerical experiments. Notably, in many instances, stations opt to accept service with a probability less than 1—a trend observed even among selfish stations.

8.4.2 Leveraging Noisy Observations in Zero-Sum Games

In [55, 31], Emmanouil Marios Athanasakos and Samir Medina Perlaza studied an instance of zero-sum games in which one player (the leader) commits to its opponent (the follower) to choose its actions by sampling a given probability measure (strategy). The actions of the leader are observed by the follower as the output of an arbitrary channel. In response to that, the follower chooses its action based on its current information, that is, the leader's commitment and the corresponding noisy observation of its action. Within this context, the equilibrium of the game with noisy action observability is shown to always exist and the necessary conditions for its uniqueness are identified. Interestingly, the noisy observations have important impact on the cardinality of the follower's set of best responses. Under particular conditions, such a set of best responses is proved to be a singleton almost surely. The proposed model captures any channel noise with a density with respect to the Lebesgue measure. As an example, the case in which the channel is described by a Gaussian probability measure is investigated.

8.4.3 Constrained Correlated Equilibria

In [33], Eitan Altman, in collaboration with Omar Boufous, Mikaël Touati, and Mustapha Bouhtou (Orange Labs), and Rachid El-Azouzi (Avignon Univ.), introduces constrained correlated equilibrium, a solution concept combining correlation and coupled constraints in finite non-cooperative games. They study the conditions for equilibrium in the general case of an arbitrary correlation device and coupled constraints in the extended game. In the particular case of constraints induced by a feasible set of probability distributions over action profiles, they first show that canonical correlation devices are sufficient to characterize the set of constrained correlated equilibrium distributions and provide conditions for their existence. Second, they demonstrate that constrained correlated equilibria of the mixed extension of the game do not lead to additional equilibrium distributions. Third, they show that the constrained correlated

equilibrium distributions may not belong to the polytope of correlated equilibrium distributions. Finally, they illustrate these results through numerical examples.

8.4.4 Positional and conformist effects in voluntary public good provision

The literature featuring game–theoretical models aimed at explaining the effect of the status concerns on the voluntary provision of a public good is generally focused on snob agents, driven by a desire for exclusiveness. However, the social context literature highlights that status concerns can give rise to a desire, in some individuals to be different from the "common herd", and in some others to conform with other people. In [15], Francisco Cabo (Univ. Valladolid), Mabel Tidball (INRAE) and Alain Jean-Marie analyze a two-player public good game under two different settings: The standard case with two positional players, versus the case in which the positional player faces a conformist player. Giving entrance to conformism has two main implications. Strong status concerns by both players can lead to a (socially) virtuous cycle in which the conformist player wishes to imitate the contributing behavior of the positional player, and the latter wishes to increase contribution to distinguish herself from the former. Then, the contribution to the public good can be higher than in the case with only snob agents. This higher contribution can increase social welfare, but only if endowments are not too large and the status concern of the positional player is not excessively high.

8.4.5 Equilibrium bids for reverse auctions with budget constraints

Alain Jean-Marie and Mabel Tidball (INRAE) consider in [60] a relatively new class of auctions: the budget-constrained reverse auction. In this auction, offers are accepted until some predefined budget is exhausted. Payments are discriminatory: bidders receive their own bid as payment when they win the bid. Bidders have a private value for their item which is uniformly distributed between 0 and 1; they are risk-neutral and items are indivisible. For small budget and an arbitrary number of bidders, as well as in the case of two bidders, an equilibrium bidding function is identified. When possible, this equilibrium bidding function is compared with that of target-constrained reverse auctions.

8.5 Applications in Telecommunications

Participants: Sara Alouf, Eitan Altman, Emmanouil Marios Athanasakos, Younes Ben Mazziane, Hariprasad Manjunath Hegde, Samir Medina Perlaza, Philippe Nain, Giovanni Neglia.

8.5.1 Local Approximation of Secrecy Capacity

In [32], Emmanouil Marios Athanasakos, Hariprasad Manjunath Hegde, and Nicholas Kalouptsidis studied the wiretap channel using Euclidean Information Theory (EIT). The focus is on a scenario in which a small amount of information shall be efficiently transmitted subject to compression and secrecy constraints. The information-theoretic problem is transformed into a linear algebra problem and obtain the perturbed probability distributions such that secrecy is achievable. An approximate estimate of the secrecy capacity is obtained by solving a generalized eigenvalue problem.

8.5.2 Caching

Probabilistic analysis of caches has been the topic of Younes Ben Mazziane's PhD Thesis [51].

In [18], Giovanni Neglia, in collaboration with Francescomaria Faticanti (École Normale Supérieure de Lyon, previously postdoc in NEO), investigates "optimistic" online caching policies, distinguished by their use of future request predictions derived, for example, from machine learning models. Traditional online optimistic policies, grounded in the Follow-The-Regularized-Leader (FTRL) algorithm, incur a higher computational cost compared to classic policies like Least Frequently Used (LFU) and Least Recently Used (LRU). This is due to each cache state update necessitating the resolution of a constrained optimization problem. To address this problem, they introduce and analyze the "batched" version of

two distinct FTRL-based optimistic policies. In this approach, the cache updates occur less frequently, thereby amortizing the update cost over time or over multiple requests. Rather than updating the cache with each new request, the system accumulates a batch of requests before modifying the cache content. First, they present a batched version of the Optimistic Bipartite Caching (OBC) algorithm, which works for single requests, then introduce a new optimistic batched caching policy, the Per-Coordinate Optimistic Caching (PCOC) algorithm, derived from the per-coordinate-based FTRL. They demonstrate that these online algorithms maintain "vanishing regret" in the batched case, meaning their average performance approaches over time that of an optimal static file allocation, regardless of the sequence of file requests. They then compare the performance of these two strategies with each other and against optimistic versions of LFU and LRU. Their experimental results indicate that this batched optimistic approach outperforms traditional caching policies on both stationary and real-world file request traces.

Commonly used caching policies, such as LRU or LFU, exhibit optimal performance only under specific traffic patterns. Even advanced machine learning-based methods, which detect patterns in historical request data, struggle when future requests deviate from past trends. Recently, a new class of policies has emerged that is robust to varying traffic patterns. These algorithms address an online optimization problem, enabling continuous adaptation to the context. They offer theoretical guarantees on the regret metric. However, the high computational complexity of these solutions hinders their practical adoption. In an ongoing work [56], Giovanni Neglia in collaboration with Damiano Carra (Univ. of Verona, Italy) introduce a new variant of the gradient-based online caching policy that achieves groundbreaking logarithmic computational complexity relative to catalog size, while also providing regret guarantees. This advancement allows them to test the policy on large-scale, real-world traces featuring millions of requests and items—a significant achievement, as such scales have been beyond the reach of existing policies with regret guarantees. Their experimental results demonstrate for the first time that the regret guarantees of gradient-based caching policies offer substantial benefits in practical scenarios.

In [14], Younes Ben Mazziane, Sara Alouf, and Giovanni Neglia, in collaboration with Daniel S. Menasche (Federal Univ. of Rio de Janeiro, Brazil), address the challenge of similarity caching, which allows requests for an item to be served by a similar item—a concept applicable in recommendation systems, multimedia retrieval, and machine learning. They note that while many similarity caching policies, such as SIM-LRU and its generalization RND-LRU, have been proposed, their performance analysis, particularly regarding hit ratio, remains limited. To address this, they extend the popular time-to-live (TTL) approximation from classic caching to similarity caching. Specifically, they propose the RND-TTL approximation method to estimate the hit ratio of the RND-LRU similarity caching policy. This involves introducing the RND-TTL cache model and tuning its parameters to emulate RND-LRU's behavior. The parameter tuning requires solving a fixed-point system of equations, for which they provide both an algorithm for numerical resolution and sufficient conditions for convergence. Their approach is evaluated using both synthetic and real-world traces, demonstrating its effectiveness in approximating the hit ratio of RND-LRU.

8.5.3 Design and cross-layer optimization of low cost RIS-assisted communication systems

In [16], Eitan Altman, in collaboration with Antoine Dejonghe and Zwi Altman (Orange Labs) and Francesco De Pellegrini (Avignon Univ.), introduces a new cross-layer low-complexity scheme for the online optimization of RIS-assisted communication systems in Orthogonal Frequency-Division Multiple Access (OFDMA) systems. This scheme jointly combines base station (BS) and reconfigurable intelligent surface (RIS) configuration with fair user equipment (UE) scheduling, which is critical for high-performance deployments. A RIS beam synthesis method is proposed for RIS configuration. The proposed solution embeds two nested control loops: i) a fast control loop operating at the OFDMA slot scale, consisting of a standard UE proportional fair scheduler; ii) a slow control loop operating at the OFDMA frame scale, which adapts the RIS configuration to the UEs' spatial distribution and maximizes the UEs' aggregated performance. The slow control loop is based on an online stochastic approximation algorithm, with convergence to the optimal rest point proved. In a reference scenario, the proposed scheduler achieves a gain of 47% in mean spectral efficiency for NLOS UEs over a baseline scheme.

8.5.4 Optimal Flow Admission Control in Edge Computing via Safe Reinforcement Learning

With the uptake of intelligent data-driven applications, edge computing infrastructures necessitate a new generation of admission control algorithms to maximize system performance under limited and highly heterogeneous resources. In [38], Eitan Altman, in collaboration with Andrea Fox, Francesco de Pellegrini (Avignon Univ.), and Francesco Bronzino (ENS Lyon), studies how to optimally select information flows which belong to different classes and dispatch them to multiple edge servers where applications perform flow analytic tasks. The optimal policy is obtained via the theory of constrained Markov decision processes to take into account the demand of each edge application for specific classes of flows, the constraints on computing capacity of edge servers and the constraints on access network capacity. They develop DRCPO, a specialized primal-dual Safe Reinforcement Learning (SRL) method which solves the resulting optimal admission control problem by reward decomposition. DRCPO operates optimal decentralized control and mitigates effectively state-space explosion while preserving optimality. Compared to existing Deep Reinforcement Learning (DRL) solutions, extensive results show that it achieves 15% higher reward on a wide variety of environments, while requiring on average only 50% learning episodes to converge. Finally, they further improve the system performance by matching DRCPO with load-balancing in order to dispatch optimally information flows to the available edge servers.

8.5.5 Learning optimal edge processing with offloading and energy harvesting

In [19], Eitan Altman, in collaboration with Andrea Fox and Francesco De Pellegrini (Avignon Univ.), introduces a Markovian model for systems where multiple battery-operated devices perform data processing and energy harvesting in parallel. The study focuses on modern portable devices capable of executing sophisticated AI models on sensed data, with data-dependent complexity and significant energy costs. The model addresses scenarios where part of the computational burden is offloaded to an edge server that polls devices at a given rate. The authors derive structural properties of an optimal policy for a single device-server system, leading to the development of a new model-free reinforcement learning method specialized for monotone policies, termed Ordered Q-Learning. This method provides a fast procedure to learn the optimal policy, independent of devices' battery capacities, the cost and value of data batch processing, and the dynamics of the energy harvesting process. Additionally, the server's polling strategy is optimized by combining this policy improvement technique with stochastic approximation methods. Extensive numerical results offer insights into system properties and demonstrate that the proposed learning algorithms outperform existing baselines.

8.5.6 Covert Communications

In [46], Philippe Nain, in collaboration with Amir Reza Ramtin and Don Towsley (Univ. Massachusetts), James Z. Hare and Lance Kaplan (DEVCOM Army Research Laboratory), and Venugopal Veeravalli (Univ. of Illinois), investigates the damage that an adversary can effect while remaining covert in the presence of the Cumulative Sum (CuSum) procedure. An adversary is covert if the time to detection is on the same order as the time to false alarm. Damage is given as an increasing function of the Kullback–Leibler divergence of the adversarial actions and the normal distribution prior to the adversarial attack. By analyzing the problem with a focus on the growth function g(n), which measures the cumulative expected log-likelihood ratio after n time slots following the change, conditions are established under which the adversary remains covert. An analysis of the impact of different adversarial strategies on damage is also provided.

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

NEO has contracts with Accenture (see §9.1.1), Nokia (see §9.1.2), NSP SmartProfile (see §9.1.3), Orange Labs (see §9.1.4), QITI (see §9.1.5), and SAP (see §9.1.6).

9.1.1 Accenture “Plan de Relance” (PLR) contract on the topic “Energy-Aware Federated Learning” (Oct 2022 – September 2024)

Participants: Giovanni Neglia, Charlotte Rodriguez.

- Contractor: Accenture Labs
- Collaborators: Laura Degioanni, Laetitia Kameni, Richard Vidal

Deep neural networks have enabled impressive accuracy improvements across many machine learning tasks. Often the highest scores are obtained by the most computationally-hungry models. As a result, training a state-of-the-art model now requires substantial computational resources which demand considerable energy, along with the associated economic and environmental costs. Research and development of new models multiply these costs by thousands of times due to the need to try different model architectures and different hyper-parameters. In this project, we investigate a more algorithmic/system-level approach to reduce energy consumption for distributed ML training over the Internet. The postdoc of C. Rodriguez is funded by this project. Publications in 2024: [63].

9.1.2 Contracts with Nokia

Inria Challenge LearnNet

Participants: Ahmad Nasser, Giovanni Neglia.

Project Acronym: LearnNet

Project Title: Learning Networks

Duration: January 2024 - December 2027

Abstract: While machine learning is revolutionizing entire sectors of the digital economy and scientific research, its robust deployment in digital infrastructures raises many questions. The challenge Learning Networks (LearnNet) explores new avenues of research at the intersection of the fields of networks and learning. This challenge has two complementary objectives: rethinking the design of network protocols to serve machine learning applications, and exploring how learning can improve network management. Thus the LearnNet challenge studies the growing entanglement between the challenges of large-scale learning and network design.

LearnNet is a research project that spans 8 Inria research teams. The teams from Nokia are AIRL and NSSR.

Inria Challenge SmartNet

Participant: Sara Alouf.

Project Acronym: SmartNet

Project Title: AI Methods for Smart Network Management

Duration: January 2024 - December 2027

Abstract: The advent of virtualization, combined with the power of AI, has brought new opportunities in network management. To effectively address the challenges that come with this paradigm shift, the SmartNet project is dedicated to exploring the transformative potential of AI methods in enabling smart network management. The project strategically focuses on two key areas: slice provisioning and causal analysis of network malfunctions. The project is dedicated to the development of cutting-edge methods to respond effectively to the growing complexity of networks, particularly in multi-domain scenarios.

SmartNet is a research project that spans 5 Inria research teams. The teams from Nokia are MLS and NSSR.

9.1.3 Cifre contract with NSP-SmartProfile "Automated and responsible recommendation systems for digital marketing" (August 2022 – July 2025)

Participants: Konstantin Avrachenkov, Ibtihal El Mimouni.

- Contractor: NSP-SmartProfile
- Collaborators: Hervé Baile, Julien Musso

SmartProfile is a marketing platform that allows to collect, to enhance and to analyze marketing data. Digital marketing campaigns continue to expand across all digital channels and media. The 'mass marketing' strategies implemented by most companies show limits in terms of performance and acceptance by clients, as well as in terms of their impact on the environment. In opposite to these practices, we believe that current technologies, particularly in terms of Artificial Intelligence (AI), should make marketing interactions more efficient and virtuous. Through this research project, we want to create an alternative solution to mass marketing by switching to an intelligent, automated and eco-responsible system, which will support the heterogeneity of data and the diversity of sectors, and whose purpose is to recommend the best content by determining the most relevant target and taking into account the communication constraints. This contract complements the Cifre thesis of Ibtihal El Mimouni.

9.1.4 Cifre contract with Orange Labs "Analytical modeling of large-scale wireless networks integrating RIS" (March 2023 – March 2026)

Participants: Eitan Altman, Julian Alfonso Santos Bustos.

- Contractor: Orange Labs
- Collaborators: Jean-Marc Kelif

A Reconfigurable Intelligent Surface (RIS) is a programmable surface structure that allows one to control the reflection of electromagnetic (EM) waves by changing the electric and magnetic properties of the surface. In the absence of RIS, short wavelentghs signals as in 5G, are subject to a huge attenuation when there is no direct line of sight channel. Within our collaboration, we shall evaluate and optimize the position of RIS.

This contract complements the Cifre thesis of J. Santos.

9.1.5 QITI 3IA Start-It-Up contract on the topic "Reinforcement Learning for Conversational Recommender Systems (RLCRS)" (April 2023 – March 2024)

Participants: Konstantin Avrachenkov, Hariprasad Manjunath Hegde.

- Contractor: QITI
- Collaborators: Christophe Bremard, Guillaume Dion

Qiti is a start-up created in Nice in 2021 which, among other things, develops a Conversational Recommender Systems (CRS) for insurance holders and insurers. The CRS should reduce the load on the workers of the insurers and to simplify the process of insurance establishment and modification. The goal of the present cooperation is to test and to improve various Reinforcement Learning schemes for CRS. The post-doc of H. Manjunath is funded by this contract.

9.1.6 Cifre contract with SAP "Privacy and fairness for ML" (December 2021 – December 2024)

Participants: Caelin Kaplan, Giovanni Neglia.

- Contractor: SAP Labs France
- Collaborators: Anderson Santana de Oliveira

There are increasing concerns among scholars and the public about bias, discrimination, and fairness in AI and machine learning. Decision support systems may present biases, leading to unfair treatment of some categories of individuals, for instance, systematically assigning high risk of recidivism in a criminal offense analysis system. Essentially, the analysis of whether an algorithm's output is fair (e.g. does not disadvantages a group with respect to others) depends on substantial contextual information that often requires human intervention. There are though several metrics for fairness that have been developed, which may rely on collecting additional sensitive attributes (e.g., ethnicity) before imposing strong privacy guarantees to be used in any situation. It is known that differential privacy has the effect of hiding outliers from the data analysis, perhaps compounding existing bias in certain situations. This project encompasses the search for a mitigating strategy. This contract complements the Cifre thesis of C. Kaplan. Publications in 2024: [44, 53, 61].

10 Partnerships and cooperations

10.1 International initiatives

10.1.1 Inria associate team not involved in an IIL or an international program

LION

Participants: Khushboo Agarwal, Eitan Altman, Konstantin Avrachenkov, Madhu Madhu, Samir Medina Perlaza.

Title: Learning In Operations and Networks

Duration: 2022 -> 2024

Coordinator: Kavitha Veeraruna (vkavitha@iitb.ac.in)

Partners:

- Indian Institute of Technology Bombay Bombay (Inde)

Inria contact: Eitan Altman

Summary: Artificial Intelligence has affected all walks of life. We propose to study its application in various domains like

1. Learning and Control in Healthcare: Our aim is to use novel AI methodologies, to predict the results of possible actions of involved decision-makers, using the available data.
2. Dual Learning Algorithms in wireless networks: We aim to develop learning algorithms for beam alignment in 5G Wireless networks to maintain high rates. We propose to use AoI as a metric.
3. Distributed and reinforcement learning: We will develop and analyze Deep Q Network (DQN) based learning algorithms and analyze their performance.

Publications in 2024: [27, 30].

10.1.2 STIC/MATH/CLIMAT AmSud projects

GSA

Participants: Konstantin Avrachenkov, Alain Jean-Marie, Hariprasad Manjunath Hegde, Lucas Siviero Sibemberg.

Title: Graph spectra and its applications

Program: MATH-AmSud

Duration: January 1, 2023 – December 31, 2024

Local supervisor: Konstantin Avrachenkov

Partners:

- Trevisan (Brasil)
- Pastine (Argentina)
- Universidad de Antofagasta (Chile)
- A. M. França (Brasil)

Inria contact: Konstantin Avrachenkov

Summary: The present project proposes the establishment of a network of collaboration among Argentina, Chile, Brazil, and France, using the strength of 5 mathematics groups of 5 different institutions. The research topic of the proposal is Algebraic Graph Theory, an important and modern area of discrete mathematics. The proposal is structured in such a way that the training of highly qualified human resources and research activities are intertwined, this will ensure the generation of new knowledge in a relevant scientific area and leave permanent ties of collaboration between the different research groups beyond the completion of the project.

10.1.3 Participation in other International Programs

MICCHI

Participant: Alain Jean-Marie.

Title: Mecanismos e Incentivos Contra la Crisis HIdrica

Funding: Chile's Agencia Nacional de Investigación y Desarollo (ANID)

Partner Institutions:

- Universidad de O'Higgins, Chile (coordinator)
- Universidad de Chile, Chile
- Universidad de Barcelona, Spain
- INRAe, France

Date/Duration: January 2024 - December 2025

Additionnal info/keywords: The water crisis caused by global warming is one of the most important problems affecting agricultural regions such as the sixth chilean Region of O'Higgins. The main objective of this project is to investigate different mechanisms for the allocation of water resources in times of scarcity.

10.2 International research visitors

10.2.1 Visits of international scientists

Other international visits to the team: research stay

Victor Bucarey

Status researcher

Institution of origin: Univ. O'Higgins (Rancagua)

Country: Chile

Dates: December 5-11, 2024

Context of the visit: Project MICCHI

Mobility program/type of mobility: research stay

Andrea Canales

Status researcher

Institution of origin: Univ. O'Higgins (Rancagua)

Country: Chile

Dates: December 5-11, 2024

Context of the visit: Project MICCHI

Mobility program/type of mobility: research stay

Damiano Carra

Status: Associate Professor

Institution of origin: University of Verona

Country: Italy

Dates: September 30 to October 4, 2024

Context of the visit: collaboration on low-regret online learning

Mobility program/type of mobility: research stay

Cecilia Gonzalez-Tokman

Status: Associate Professor

Institution of origin: University of Queensland

Country: Australia

Dates: June 24, 2024 to June 28, 2024

Context of the visit: collaboration on dynamic systems, in particular, on ergodicity of dynamic systems

Mobility program/type of mobility: research stay

Eduard Gorbunov

Status: Research Scientist

Institution of origin: Mohamed bin Zayed University of Artificial Intelligence (MBZUAI)

Country: United Arab Emirates

Dates: November 25 to November 29, 2024

Context of the visit: collaboration on Scalable Federated Learning algorithms

Mobility program/type of mobility: research stay

Madhu Madhu

Status: PhD student

Institution of origin: IIT Bombay

Country: India

Dates: September 15 to December 15, 2024

Context of the visit: collaboration on Partially Observed Markov Decision Processes

Mobility program/type of mobility: research stay

Maksim Mironov

Status: Postdoc

Institution of origin: -

Country: Portugal

Dates: January 22 to January 26, 2024

Context of the visit: collaboration on Ising-type models for graph clustering

Mobility program/type of mobility: research stay

Fernando Ordóñez

Status researcher

Institution of origin: Univ. de Chile (Santiago de Chile)

Country: Chile

Dates: December 5-11, 2024

Context of the visit: Project MICCHI

Mobility program/type of mobility: research stay

Daniel Ratton Figueiredo

Status: Associate Professor

Institution of origin: UFRJ, on sabbatical leave at EPFL

Country: Brazil

Dates: September 13 to September 18, 2024

Context of the visit: collaboration on Complex Networks and Federated Learning

Mobility program/type of mobility: research stay

Daniel Richards Ravi Arputharaj

Institution of origin: KTH

Country: Sweden

Dates: September 1, 2024 to February 28, 2025

Context of the visit: working on semi-distributed federated learning

Mobility program/type of mobility: research stay

Jan Schuurmans

Status: Researcher and CEO

Institution of origin: DotX Control Solutions

Country: The Netherlands

Dates: August 26 to September 14, 2024

Context of the visit: project TESTBED2

Mobility program/type of mobility: research stay

Vsevolod Shneer

Status: Associate Professor

Institution of origin: Heriot Watt University

Country: United Kingdom

Dates: April 13 to April 19, 2024

Context of the visit: collaboration on statistical mechanic methods for graph clustering

Mobility program/type of mobility: research stay

Lucas Siviero Sibemberg

Status: PhD student

Institution of origin: Universidade Federal do Rio Grande do Sul (Porto Alegre)

Country: Brazil

Dates: March 13 to October 31, 2024

Context of the visit: MATH AmSUD exchange program (project GSA)

Mobility program/type of mobility: research stay

Rajesh Sundaresan

Status: Full Professor

Institution of origin: Indian Institute of Science (IISc)

Country: India

Dates: May 26 to June 2, 2024

Context of the visit: collaboration on stochastic perturbation of dynamic systems

Mobility program/type of mobility: research stay

Jacopo Talpini

Status: PhD student

Institution of origin: University of Milano - Bicocca

Country: Italy

Dates: November 1, 2023 to April 30, 2024

Context of the visit: collaboration on federated learning

Mobility program/type of mobility: research stay

Pieter Van Klaveren

Status: Researcher Engineer

Institution of origin: DotX Control Solutions

Country: The Netherlands

Dates: August 26 to September 14, 2024

Context of the visit: project TESTBED2

Mobility program/type of mobility: research stay

Kavitha Veeraruna

Status: Professor

Institution of origin: IIT Bombay

Country: India

Dates: November 19 to November 29, 2024

Context of the visit: collaboration on stochastic processes, queueing theory and game theory (associated team LION)

Mobility program/type of mobility: research stay

Other international visits to the team: lecture

Erik Larsson

Status: Full Professor

Institution of origin: Linköping University

Country: Sweden

Dates: July 2 to July 4, 2024

Context of the visit: Seminar: Topics in wireless federated learning: streaming data and over-the-air-aggregation

Mobility program/type of mobility: lecture, jury member of PhD defense

Fionn McInerney

Status: Postdoc

Institution of origin: TU Wien

Country: Austria

Dates: January 10 to January 12, 2024

Context of the visit: Seminar: The Complexity of Optimizing Atomic Congestion

Mobility program/type of mobility: lecture

Stefano Rini

Status: Associate Professor

Institution of origin: National Yang Ming Chiao Tung University (NYCU)

Country: Taiwan

Dates: April 24 to April 26, 2024

Context of the visit: Seminar: Communication-Efficient Federated Learning: Challenges, Techniques, and Insights

Mobility program/type of mobility: lecture

Uri Yechiali

Status: Professor Emeritus

Institution of origin: Tel Aviv University

Country: Israel

Dates: April 11 to April 14, 2024

Context of the visit: Seminar: Explicit solutions for continuous-time QBD processes by using relations between matrix geometric analysis and the probability generating functions method

Mobility program/type of mobility: lecture

Other international visits to the team: internship

Albachiara Bellaroba

Status: intern (master/eng)

Institution of origin: University of Rome La Sapienza

Country: Italy

Dates: March 1 to September 18, 2024

Context of the visit: working on online learning

Mobility program/type of mobility: internship

10.2.2 Visits to international teams

Research stays abroad

Konstantin Avrachenkov

Visited institution: Indian Institute of Technology (Bombay)

Country: India

Dates: Februrary 24 to March 10

Context of the visit: Project Cefipra LION

Mobility program/type of mobility: research stay

Visited institution: Eindhoven University of Technology and Leiden University

Country: the Netherlands

Dates: March 20 to March 29

Context of the visit: collaboration with Profs. Nelly Litvak and Floske Spieksma

Mobility program/type of mobility: research stay

Visited institution: Umea University

Country: Sweden

Dates: October 10 to October 18

Context of the visit: collaboration with Prof. L. Freidovich

Mobility program/type of mobility: research stay

José Francisco Daunas Torres

Visited institution: Princeton University

Country: USA

Dates: April 1 to June 30

Context of the visit: Project TESTBED2

Mobility program/type of mobility: research stay

Alain Jean-Marie

Visited institution: Univ. O'Higgins (Rancagua)

Country: Chile

Dates: January 15 to January 19

Context of the visit: Project MICCHI

Mobility program/type of mobility: research stay

Angelo Rodio

Visited institution: Linköping University University

Country: Sweden

Dates: April 3 to April 6

Context of the visit: collaboration with Profs. Erik G. Larsson and Zheng Chen on decentralized learning

Mobility program/type of mobility: research stay

Xinying Zou

Visited institution: Southeast University (Nanjing)

Country: China

Dates: November 15 to December 13

Context of the visit: Project TESTBED2

Mobility program/type of mobility: research stay

10.3 European initiatives

10.3.1 Horizon Europe

dAIEDGE

Participants: Sara Alouf, Alain Jean-Marie, Giovanni Neglia.

dAIEDGE project on cordis.europa.eu

Title: A network of excellence for distributed, trustworthy, efficient and scalable AI at the Edge

Duration: From September 1, 2023 to August 31, 2026

Partners:

- INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET AUTOMATIQUE (INRIA), France
- NVISO SA (NVISO), Switzerland
- UBOTICA TECHNOLOGIES LIMITED, Ireland
- CSEM CENTRE SUISSE D'ELECTRONIQUE ET DE MICROTECHNIQUE SA - RECHERCHE ET DEVELOPPEMENT (CSEM), Switzerland
- VARJO TECHNOLOGIES OY, Finland

- FRAUNHOFER GESELLSCHAFT ZUR FORDERUNG DER ANGEWANDTEN FORSCHUNG EV (Fraunhofer), Germany
- THALES SIX GTS FRANCE SAS (THALES SIX GTS France), France
- COMMISSARIAT A L ENERGIE ATOMIQUE ET AUX ENERGIES ALTERNATIVES (CEA), France
- INTERUNIVERSITAIR MICRO-ELECTRONICA CENTRUM (IMEC), Belgium
- SOFIA UNIVERSITY ST KLIMENT OHRIDSKI (UNISOFIA), Bulgaria
- IDRYMA TECHNOLOGIAS KAI EREVNAS (FOUNDATION FOR RESEARCH AND TECHNOLOGYHELLAS), Greece
- FUNDACION INSTITUTO INTERNACIONAL DE INVESTIGACION EN INTELIGENCIA ARTIFICIAL Y CIENCIAS DE LA COMPUTACION, Spain
- BONSEYES COMMUNITY ASSOCIATION, Switzerland
- SINTEF AS (SINTEF), Norway
- DEUTSCHES FORSCHUNGSZENTRUM FUR KUNSTLICHE INTELLIGENZ GMBH (DFKI), Germany
- DEUTSCHES ZENTRUM FUR LUFT - UND RAUMFAHRT EV (DLR), Germany
- FUNDACION CENTRO DE TECNOLOGIAS DE INTERACCION VISUAL Y COMUNICACIONES VICOMTECH (VICOM), Spain
- FUNDINGBOX ACCELERATOR SP ZOO (FBA), Poland
- BLEKINGE TEKNISKA HOGSKOLA (BTH), Sweden
- EIDGENOESSISCHE TECHNISCHE HOCHSCHULE ZUERICH (ETH Zürich), Switzerland
- SYNOPSYS INTERNATIONAL LIMITED (SYNOPSYS), Ireland
- UNIVERSIDAD DE CASTILLA - LA MANCHA (UCLM), Spain
- SAFRAN ELECTRONICS & DEFENSE, France
- VERSES GLOBAL BV, Netherlands
- HAUTE ECOLE SPECIALISEE DE SUISSE OCCIDENTALE (HES-SO), Switzerland
- UNIVERSITY OF GLASGOW, United Kingdom
- STMICROELECTRONICS SRL, Italy
- Aegis Rider AG (Aegis Rider), Switzerland
- CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS (CNRS), France
- KATHOLIEKE UNIVERSITEIT LEUVEN (KU Leuven), Belgium
- UNIVERSITA DEGLI STUDI DI MODENA E REGGIO EMILIA (UNIMORE), Italy
- THE UNIVERSITY OF EDINBURGH (UEDIN), United Kingdom
- HIPERT SRL, Italy
- UNIVERSIDAD DE SALAMANCA (USAL), Spain
- SORBONNE UNIVERSITE, France
- CENTRE D'EXCELLENCE EN TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION (CETIC), Belgium

Inria contact: Giovanni Neglia

Coordinator: Alain Pagani (DFKI)

Summary: The proposal focuses on the Next Generation AI topic of the call HORIZON-CL4-2022-HUMAN-02-02. The vision of dAIEDGE Network of Excellence (NoE) is to strengthen and support the development of the dynamic European edge and distributed Artificial Intelligence (AI) ecosystem as an essential ingredient in the growth and competitiveness of European industrial sectors. The dAIEDGE Network aims to reinforce the research and innovation value chains to accelerate the

digital and green transitions through advanced edge AI technologies, applications, and innovations, building on Europe's existing assets and industrial strengths. In parallel, it will fortify the edge AI research and industrial communities through technological developments beyond state of the art and become a dependable and strategic pillar for the European AI Lighthouse. This will be achieved by mobilizing and connecting the European AI and edge AI constituency, the relevant stakeholders, European partnerships, and projects, to provide roadmaps, guidelines and trends supporting the next-generation edge AI technologies. The key aim is to support and ensure rapid development, market uptake and open strategic sovereignty for Europe in the critical technologies for distributed edge AI (hardware, software, frameworks, tools). The dAIEDGE NoE will play a catalyst role in building a solid edge AI virtual network of research facilities and laboratories to benefit the European research and industrial community. The NoE multidisciplinary concept provide an arena for matchmaking, exchanging ideas, tools, and services, by bringing together the leading research centres, AI-on-demand platforms, digital innovation hubs, AI projects and initiatives. The ultimate goal for the dAIEDGE NoE is to support Europe to become a global centre of excellence with unique human-centred edge AI competence addressing the social and economic challenges and the needs of the citizens and society.

Publications in 2024: [18, 34, 35, 44, 45, 48, 39, 58, 61].

10.3.2 H2020 projects

TESTBED2

Participants: José Francisco Daunas Torres, Samir Medina Perlaza, Xinying Zou.

Project Title: Testing and Evaluating Sophisticated information and communication Technologies for enaBling scalablE smart griD Deployment

Program: H2020-MSCA-RISE

Coordinator: University of Durham, UK

Duration: February 2020 - June 2025

Local supervisor: Samir Medina Perlaza

Others Partners: The University of Durham (UDUR); University of Tuebingen (EKUT); Heriot-Watt University (HWU); University of Klagenfurt (AAU); University of Northumbria at Newcastle (UNN); DotX Control Solutions (DotX); BEIA Consult International (BEIA); DEPSys (DEPS); Hellenic Telecommunications Organization S.A (OTE); Princeton University (PU); University of California, Santa Barbara (UC); University of Nebraska-Lincoln (UNL); Institute of Electrical Engineering of the Chinese Academy of Sciences (CAS); China Electric Power Research Institute (EPRI); Southeast University (SEU); and Jinan University (JNU)

Abstract: TESTBED2 is a major interdisciplinary project that combines wisdoms in three academic disciplines - Electronic & Electrical Engineering, Computing Sciences and Macroeconomics, for developing new techniques to improve the scalability of smart grid services, particularly considering the joint evolution of decarbonized power, heat and transport systems. Moreover, new experimental testbeds will be created to evaluate scalable smart grid solutions. Overall, the main objective of this project is to coordinate the action of 12 Universities and 5 enterprises (3 SMEs and 2 large enterprises) with complementary expertise to develop and test various promising strategies for ensuring the scalability of smart grid services, thereby facilitating successful deployment and full roll-out of smart grid technologies.

Publications in 2024: [28, 29, 36, 57, 62, 64].

10.4 National initiatives

NF-FOUNDS PC9 PEPR 5G

Participants: Khushboo Agarwal, Eitan Altman, Samir Medina Perlaza.

Project Acronym: NF-FOUNDS

Project Title: Networks of the Future - Foundations of Future Communications Networks

Program: ANR-22-PEFT-0010

Coordinator: CEA (Dmitri Kténas), CNRS (Serge Verdeyme), IMT (Daniel Koffman)

Duration: 2023 - 2030

Other Partners: EURECOM

Abstract: The 5G network and the networks of the future represent a key issue for French and European industry, society and digital sovereignty. This is why the French government has decided to launch a dedicated national strategy. One of this strategy's priority ambitions is to produce significant public research efforts so the national scientific community contributes fully to making progress that clearly responds to the challenges of 5G and the networks of the future. In this context, the CNRS, the CEA and the Institut Mines-Télécom (IMT) are co-leading the '5G' acceleration PEPR to support upstream research into the development of advanced technologies for 5G and the networks of the future. NEO is involved in the theme "Networks and Telecommunications" and more specifically in the targeted projet 9 (PC9) Foundations of Future Communications Networks (FOUNDS).

Publications in 2024: [30].

ANR PARFAIT

Participants: Eitan Altman, Ashok Krishnan Komalan Sindhu, Samir Medina Perlaza, Xinying Zou.

Project Acronym: PARFAIT

Project Title: Planning And leaRning For AI-Edge compuTing

Coordinator: Avignon Univ.

Duration: October 2021 - September 2025

Other Partners: Conservatoire National des Arts et Métiers (CNAM), Univ. Savoie Mont Blanc (USMB)

Abstract: The PARFAIT project develops theoretical foundations for distributed and scalable resource allocation schemes on edge computing infrastructures tailored for AI-based processing tasks. Algorithmic solutions will be developed based on the theory of constrained, delayed, and distributed Markov decision processes to account for edge service orchestration actions and quantify the effect of orchestration policies. Furthermore, using both game and team formulations, the project will pave the way for a theory of decentralized orchestration, a missing building block necessary to match the application quest for data proximity and the synchronization problems that arise when multiple edge orchestrators cooperate under local or partial system view. Finally, to achieve efficient online edge service orchestration, such solutions will be empowered with reinforcement learning techniques to define a suit of orchestration algorithms able to at once adapt over time to the applications' load and cope with the uncertain information available from AI-based applications' footprints.

Publications in 2024: [29, 38, 49, 62, 64].

Inria Challenge FedMalin

Participants: Othmane Marfoq, Giovanni Neglia, Angelo Rodio.

Project Acronym: FedMalin

Project Title: FEDerated MAchine Learning over the INternet

Coordinator: Giovanni Neglia and Aurélien Bellet (Inria team PREMEDICAL)

Duration: November 2022 - November 2026

Abstract: In many use-cases of Machine Learning (ML), data is naturally decentralized: medical data is collected and stored by different hospitals, crowdsensed data is generated by personal devices, etc. Federated Learning (FL) has recently emerged as a novel paradigm where a set of entities with local datasets collaboratively train ML models while keeping their data decentralized.

FedMalin is a research project that spans 10 Inria research teams and aims to push FL research and concrete use-cases through a multidisciplinary consortium involving expertise in ML, distributed systems, privacy and security, networks, and medicine. We propose to address a number of challenges that arise when FL is deployed over the Internet, including privacy and fairness, energy consumption, personalization, and location/time dependencies. FedMalin will also contribute to the development of open-source tools for FL experimentation and real-world deployments, and use them for concrete applications in medicine and crowdsensing. The FedMalin Inria Challenge is supported by Groupe La Poste, sponsor of the Inria Foundation.

Publications in 2024: [34, 35, 45, 48, 58].

DIAMOND

Participants: Yaiza Bermudez, Samir Medina Perlaza.

Project Acronym: DIAMOND

Project Title: Data-Injection Attacks in Supervised Machine Learning Systems

Funding Agency: Agence de l'Innovation de Défense (AID)

Principal Investigators: Samir Medina Perlaza and Iñaki Esnaola (University of Sheffield)

Duration: November 2024 - October 2027

Note: The project funds the PhD scholarship of Yaiza Bermudez.

11 Dissemination

11.1 Promoting scientific activities

11.1.1 Scientific events: organisation

Steering committee chair, steering committee member

- Eitan Altman is
  - Member of the steering committee and founder of "Workshop on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)";

- Member of the steering committee and co-founder of "Workshop on Networking Games Control and Optimization (NetGcoop)";
- Member of the steering committee of "International Conference on Performance Evaluation Methodologies and Tools (ValueTools)".

General chair, scientific chair

- Sara Alouf is General Co-Chair of the 36th Intl. Teletraffic Congress (ITC 36), to be held in Trondheim, Norway, June 2-6, 2025.

Member of the organizing committees

- Sara Alouf was Tutorials Chair for ACM SIGMETRICS/ IFIP PERFORMANCE 2024, held in Venice, Italy, June 10-14, 2024.

11.1.2 Scientific events: selection

Member of the conference program committees

- ACM International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MobiHoc 2024), October 14-17, 2024, Athens, Greece (Sara Alouf);
- ACM SIGMETRICS/ IFIP Performance 2024, *Winter* TPC, June 10-14, 2024, Venice, Italy (Konstantin Avrachenkov);
- ACM SIGMETRICS 2025, *Summer* and *Fall* TPCs, June 9-13, 2025, Stony Brook, New York, USA (Konstantin Avrachenkov, Diego Goldsztajn);
- European Wireless 2024, September 9-11, 2024, Brno, Czech Republic (Samir Medina Perlaza);
- IEEE International Conference on Computer Communications (INFOCOM 2025), May 19-22, 2025, London, UK (Sara Alouf, Konstantin Avrachenkov, Giovanni Neglia);
- IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN 2024), May 5-8, 2024, Stockholm, Sweden (Samir Medina Perlaza);
- IEEE Virtual Conference on Communications (VCC 2024), December 3-5, 2024, Virtual Conference (Samir Medina Perlaza);
- International Symposium on Wireless Communication Systems (ISWCS 2024), July 14-17, 2024, Rio de Janeiro, Brazil (Samir Medina Perlaza);
- 28th International Conference on Analytical & Stochastic Modelling Techniques & Applications (ASMTA 2024), June 14, 2024, Venice, Italy (Konstantin Avrachenkov);
- 27th International Workshop on Smart Antennas, March 17-19, 2024. Dresden, Germany (Samir Medina Perlaza);
- 20th European Performance Engineering Workshop (EPEW 2024), June 14, 2024, Venice, Italy (Alain Jean-Marie);
- 19th Workshop on Modelling and Mining Networks (WAW 2024), June 3-6, 2024, Warsaw, Poland (Konstantin Avrachenkov);
- 15th Conference on Decision and Game Theory for Security (GameSec 2024), October 16-18, 2024, New York, USA (Konstantin Avrachenkov);
- SIAM Conference on Data Mining (SDM 2024), April 18-20, 2024, Houston, United States (Konstantin Avrachenkov);

- 17th International Conference on Performance Evaluation Methodologies and Tools (ValueTools 2024), December 12-13, 2024, Milan, Italy (Konstantin Avrachenkov);
- 13th International Conference on Complex Networks and their Applications, December 10-12, 2024, Istanbul, Turkey (Konstantin Avrachenkov);
- 4th French Regional Conference on Complex Systems (FRCCS 2024), May 29-31, 2024, Montpellier, France (Konstantin Avrachenkov);
- Reinforcement Learning for Stochastic Networks Conference, June 17-21, 2024, Toulouse (Konstantin Avrachenkov);
- IEEE International Symposium on Information Theory (ISIT 2025), June 22-27, 2025, Ann Harbor, Michigan, USA (Samir Medina Perlaza);
- IEEE International Conference in Communications (ICC 2024), June 9-13, 2024, Denver, CO, USA (Samir Medina Perlaza);
- Workshop on Smart Antennas (WSA 2024), March 17-19, 2024, Dresden, Germany (Samir Medina Perlaza);
- 22nd International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt 2024), October 21-24, 2024, Seoul, South Korea (Konstantin Avrachenkov);
- 32nd International Symposium on the Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS 2024), October 21-23, 2024, Krakow, Poland (Alain Jean-Marie);
- 25th Conf. of the Société Française de Recherche Opérationnelle et d'Aide à la Décision (ROADEF 2024), March 4-7, 2024, Amiens, France (Alain Jean-Marie);
- 26th Workshop on MAthematical performance Modeling and Analysis (MAMA 2024), June 14, 2024, Venice, Italy (Alain Jean-Marie, Philippe Nain);
- 27th European Conference on Artificial Intelligence (ECAI 2024), October 19-24, 2024, Santiago de Compostela, Spain (Giovanni Neglia);
- 38th Annual Conference on Neural Information Processing Systems (NeurIPS 2024), December 10-15, 2024, Vancouver (Giovanni Neglia).

Reviewer NEO members have performed reviews for IEEE International Conference on Computer Communications (INFOCOM 2025), European Wireless (EW 2024), IEEE Conference on Decision and Control (CDC 2024).

11.1.3 Journal

Member of the editorial boards

- ACM Transactions on Modeling and Performance Evaluation of Computing Systems (ACM ToMPECS) (Konstantin Avrachenkov, since 2016);
- AIMS (American Institute of Mathematical Sciences) Journal of Dynamics and Games (JDG) (Eitan Altman, since 2015);
- Birkhauser Journal on Dynamic Games and Applications (DGAA) (Eitan Altman, since 2012);
- CUP Probability in the Engineering and Informational Sciences (Konstantin Avrachenkov, since 2018);
- Elsevier Computer Communications (Sara Alouf, since 2021; Giovanni Neglia, since 2014);
- Elsevier Performance Evaluation (Konstantin Avrachenkov, since 2009; Philippe Nain, member of Advisory Board since 2018).

- IEEE/ACM Transactions on Networking (Sara Alouf, since 2024, Eitan Altman, editor-at-large since 2013);
- IEEE Network Magazine (Konstantin Avrachenkov, since 2020);
- IEEE Transactions on Automatic Control (Konstantin Avrachenkov, since 2024);
- Polynesian Journal of Mathematics (Samir Medina Perlaza, since 2024);
- Springer Iran Journal of Computer Science (Eitan Altman, advisory board member);
- Taylor & Francis Stochastic Models (Konstantin Avrachenkov, since 2019).

Sara Alouf has led the edition of a special issue of Performance Evaluation on extended papers from the 35th International Teletraffic Congress 2023 [50].

Reviewer - reviewing activities NEO members regularly perform reviews for journals such as Dynamic Games and Applications, IEEE/ACM Transactions on Networking, IEEE Transactions on Automatic Control, IEEE Transactions on Information Theory, IEEE Transactions on Wireless Communications, IEEE Transactions on Communications, IEEE Transactions on Network and Service Management, IEEE Transactions on Network Science and Engineering, Performance Evaluation, Elsevier Computer Communications, Elsevier Computer Networks.

11.1.4 Invited talks

- Sara Alouf delivered
  - a lecture "Performance modeling of similarity caching" at the MAKI Scientific Workshop 2024, Darmstadt;
  - a seminar "TTL Model for an LRU-Based Similarity Caching Policy" at IISc CNI Seminar Series, January 2024, Online.
- Konstantin Avrachenkov delivered a lecture "A Journey Through the History of Markov Chains" at 13ème Atelier en Évaluation des Performances, December 2-4, 2024, Toulouse.
- Diego Goldsztajn delivered a seminar "Asymptotically Optimal Policies for Weakly Coupled Markov Decision Processes" at LAAS Solace Seminar Series, October 2024, Toulouse.
- Louis Hauseux delivered a seminar "Density-based clustering: (Hyper-)Graphs & Percolation" at Università di Genova, April 2024, Genova, Italy.
- Alain Jean-Marie delivered a plenary talk "Challenges in Dynamic Stackelberg Games" at the 2024 IWOBIP, Rancagua, Chile.
- Samir Medina Perlaza delivered
  - a talk "From Statistical Empirical Risk Minimization with f-Divergence Regularization to Generalization Theory", at the Institut Henri Poincaré, GDR-IASIS, 5 July 2024, Paris, France;
  - an invited talk "An Exact Characterization of the Generalization Error of Machine Learning Algorithms" at the Polytech Annecy-Chambéry / LISTIC, Université Savoie Mont Blanc, Annecy Le Vieux, 21 February 2024. Host: Prof. Kave Salamantian.
- Giovanni Neglia delivered
  - an invited talk on "Personalized Federated Learning" at EPFL Control Seminar Series, February 23, 2024, Lausanne, Switzerland;
  - an invited talk on "Collaborative Inference Systems" at 3rd Workshop on Principles of Distributed Learning 2024, June 21, 2024, Nantes;

  - an invited talk on "Scalable Decentralized Algorithms for Online Personalized Mean Estimation" at the Workshop on distributed computing, optimization & learning 2024, May 22-23, 2024, Paris;
  - an invited talk "Scalable Decentralized Algorithms for Online Personalized Mean Estimation" at 13ème Atelier en Évaluation des Performances, December 2-4, 2024, Toulouse.

11.1.5 Tutorials

- Alain Jean-Marie delivered a tutorial on "An introduction to pyMarmote and pyMarmoteMDP for Markovian modeling" at the SIGMETRICS/PERFORMANCE 2024 conference [65].
- Samir Medina Perlaza delivered a tutorial on "Characterizing the Generalization Error of Machine Learning Algorithms Via Information Measures" in the IEEE Information Theory Workshop 2024.

11.1.6 Leadership within the scientific community

- Eitan Altman
  - is Fellow Member of IEEE;
  - is Member of WG 7.3 of IFIP on Computer System Modeling;
  - is the elected Vice Chairman of WG 6.3 of IFIP on Performance of Communications Systems.
- Konstantin Avrachenkov is a member of Conseil Scientifique & Pédagogique EUR DS4H Univ. Côte d'Azur.
- Samir Medina Perlaza
  - is a member of the Digital Presence Committee of the IEEE Information Theory Society;
  - is a workpackage leader of the PEPR - Réseaux du Futur – A project funded by the French National Agency for Research (ANR) via the project n°ANR-22-PEFT-0010 of the France 2030 program;
  - is the organizer of the PC9 Seminar on Wireless Communications, a national online seminar part of the PEPR – Réseaux du Futur.
- Giovanni Neglia
  - holds a Chair by the Interdisciplinary Institute for Artificial Intelligence 3IA Côte d'Azur, in the theme "Core Elements of AI".
  - participated in the panel on "Challenges and opportunities of Distributed Learning" organized by the NetworkingChannel, December 11, 2024, online.

11.1.7 Research administration

- Sara Alouf
  - is a member of the Colloquium Jacques Morgenstern Committee of Inria center at Université Côte d'Azur, since March 2023;
  - is a member of NICE, the Invited Researchers Committee of Inria center at Université Côte d'Azur, since June 2020;
  - is member of CLF, the training committee of Inria center at Université Côte d'Azur, November 2014 - 2024;
  - is vice-head of project-team Neo since January 2017;
  - is a member of the Selection Committee of the dAIEGDE European project.
- Konstantin Avrachenkov is a representant (suppléant, collège A) before the Center Committee of Inria center at Université Côte d'Azur.

- Louis Hauseux is
  - a representant (titulaire, collège C) before the Center Committee of Inria center at Université Côte d'Azur;
  - a representant member of the catering commission of Inria center at Université Côte d'Azur.
- Alain Jean-Marie
  - was the scientific coordinator of Inria activities in Montpellier (2008-2023); as part of this duty, he represented Inria at: the Scientific Council of the Doctoral School "Sciences and Agrosciences" of the Univ. of Avignon; at the Regional Conference of Research Organisms (CODOR); at the board of the Labex NUMEV;
  - is Head of project-team NEO since January 2017.
- Philippe Nain chaired the Inria ad-hoc commission in charge of evaluating applications for the 2025 bonus campaign (RIPEC-C3).
- Samir Medina Perlaza
  - is a representant (suppléant, collège A) before the Center Committee of Inria center at Université Côte d'Azur;
  - is a Member of the Bureau of the Réseaux, Information et Société Numérique Excellence Academy of Université Côte d'Azur;
  - represents Inria at the Conseil du Département Disciplinaire Informatique of Université Côte d'Azur.
- Giovanni Neglia
  - was a member of the competitive exam jury for a researcher position (CRCN/ISFP) at Inria Paris Center in 2024;
  - is an elected member of Inria evaluation committee, since September 2023;
  - is a member of the steering commitee of UniCA Graduate School of Digital Systems for Humans (DS4H) since September 2022.

11.2 Teaching - Supervision - Juries

11.2.1 Teaching

PhD

- G. Neglia, "Networking Challenges with Federated Learning", 1.5h for Lipari School on Advanced Networking Systems, Lipari, Italy 7-13/7/2024.

Master

- Sara Alouf, "Performance Evaluation of Networks", 21H, M2 Ubinet, UniCA, France.
- Louis Hauseux, Konstantin Avrachenkov, "Statistical Analysis of Networks", 24H, M2 Data Science and Artificial Intelligence, UniCA, France.
- Louis Hauseux, "Statistical Inference", 30H, M1 Data Science and Artificial Intelligence, UniCA, France.
- Giovanni Neglia, Samir Medina Perlaza, "Machine Learning: Theory and Algorithms", 21H, M2 Ubinet, UniCA, France.
- Giovanni Neglia and Angelo Rodio together with Chuan Xu (COATIteam) "Federated Learning & Data Privacy", 24H, M2 Data Science and Artificial Intelligence, UniCA, France.
- Xufeng Zhang, "Mathématiques pour l'IA", 15H, Bac+5 Formations expertes, CentraleDigitalLab, LaPlateforme_ & Centrale Méditerranée, France.

Bachelor

- Younes Ben Mazziane, "Introduction à la progammation par Python", 24H, BAT3, Polytech Nice Sophia, UniCA, France.
- Younes Ben Mazziane, "Programmation avancée en Python", 40H, BAT3, Polytech Nice Sophia, UniCA, France.
- Xufeng Zhang, "Introduction to Research", 20H, Bac+3, Computer Science, Polytech Nice Sophia, UniCA, France.

11.2.2 Supervision

PhD defended

- Younes Ben Mazziane, "Probabilistic analysis for caching", UniCA, defended May 13, 2024, advisors: Sara Alouf and Giovanni Neglia [51].
- Olha Chuchuk, "Data access optimisation at CERN and in the Worldwide LHC Computing Grid (WLCG)", UniCA, defended February 2024, advisor: Giovanni Neglia [52].
- Caelin Kaplan, "Inherent Trade-offs in Privacy-Preserving Machine Learning", UniCA, defended November 22, 2024, advisors: Giovanni Neglia and Alain Jean-Marie [53].
- Angelo Rodio, "Client heterogeneity in federated learning systems", UniCA, defended July 3, 2024, advisors: Giovanni Neglia and Alain Jean-Marie [54].

PhD in progress

- Yaiza Bermudez, "Data Integrity in Distributed Learning Systems", UniCA, since Nov. 1, 2024, advisor: Samir Medina Perlaza.
- José Francisco Daunas Torres, "Empirical Risk Minimization with Statistical Regularizations", Univ. of Sheffield, since Oct. 1, 2020, co-advisor: Samir Medina Perlaza.
- Ibtihal El Mimouni, "Systèmes de recommandation automatisés et responsables pour le marketing digital", UniCA, Cifre thesis with NSP SmartProfile, since Oct. 1, 2022, advisor: Konstantin Avrachenkov.
- Louis Hauseux, "Classifiers on Random Graphs with applications to Social Networks and Image Processing", UniCA, since Oct. 1, 2023, advisors: Konstantin Avrachenkov and Josiane Zerubia (AYANA).
- Ahmad Nasser, "Distributed training of heterogeneous architectures", UniCA, Cifre thesis with Nokia, since Apr. 1, 2024, advisor: Giovanni Neglia.
- Julian Alfonso Santos Bustos, "Modélisation analytique de réseaux sans fils grande échelle intégrant les RIS - Optimisation de l'allocation dynamique des ressources", UniCA, Cifre thesis with Orange, since Sep. 1, 2023, advisor: Eitan Altman.
- Xufeng Zhang, "Incentives for Federated Learning", UniCA, since Dec. 1, 2023, advisors: Sara Alouf and Giovanni Neglia.
- Xinying Zou, "Generalization Capabilities of Machine Learning Algorithms", UniCA, since Dec. 1, 2022, advisors: Eitan Altman and Samir Medina Perlaza.

11.2.3 Juries

PhD

- Sannara Ek, "Personalized federated learning for sensor-based human activity recognition in pervasive heterogeneous environments", Université Grenoble Alpes, November 26, 2024 (Giovanni Neglia, jury member).
- Mathieu Even, "Towards Decentralization, Asynchrony, Privacy and Personalization in Federated Learning", June 27, 2024 (Giovanni Neglia, reviewer).
- Franco Galante, "Modeling Social Interactions in Complex Systems", Politecnico di Torino, June 25, 2024 (Sara Alouf, reviewer).
- Filippo Galli 2024, "Data Utility and User Privacy in Differentially Private Machine Learning", Scuola Normale Superiore di Pisa, Italy, October 16, 2024 (Giovanni Neglia, reviewer).
- Martijn Gösgens, "Detecting Small and Large Communities in Networks", Eindhoven University of Technology, June 11, 2024 (Konstantin Avrachenkov, reviewer).
- Hugo Jaquard, "Parallel Transport for Monte-Carlo Estimation on Graph Bundles", Université Grenoble Alpes, November 18, 2024 (Konstantin Avrachenkov, reviewer).
- Eunjeong Jeong, "Communication-efficient Decentralized Learning for Intelligent Networked Systems", September 27, 2024 (Giovanni Neglia, jury member).
- Ali Khalesi, "Multi-User Linearly-Decomposable Distributed Computing", Eurecom / Université Sorbonne, July 5, 2024 (Konstantin Avrachenkov, jury member).
- Li Zou, "Prediction and Intervention Strategy Design on Temporal Networks", Delft University of Technology, May 24, 2024 (Konstantin Avrachenkov, reviewer).

HDR

- Ana Bušić, “Decision and Control in Networks with Stochastic Demand and Supply”, PSL Univ., May 21, 2024 (Alain Jean-Marie, reviewer);
- Benjamin Legros, “Efficiency Enhancement in Service and Transportation through Queueing Models”, PSL Univ., April 12, 2024 (Alain Jean-Marie, jury member).
- Emanuele Natale, “On Random Subset Sum and Some Applications”, Univ. Côte d’Azur, March 20, 2024 (Giovanni Neglia, jury president).

12 Scientific production

12.1 Major publications

[1] K. Avrachenkov, A. Bobu and M. Dreveton. ‘Higher-Order Spectral Clustering for Geometric Graphs’. In: *Journal of Fourier Analysis and Applications* 27 (15th Mar. 2021). DOI: 10.1007/s00041-021-09825-2. URL: https://inria.hal.science/hal-03169834.

[2] K. Avrachenkov and M. Dreveton. *Statistical Analysis of Networks*. Now Publishers, 6th Oct. 2022. DOI: 10.1561/9781638280514. URL: https://inria.hal.science/hal-03932416.

[3] K. E. Avrachenkov and V. Borkar. ‘Whittle index based Q-learning for restless bandits with average reward’. In: *Automatica* 139 (May 2022), p. 110186. DOI: 10.1016/j.automatica.2022.110186. URL: https://inria.hal.science/hal-03582664.

[4] Y. Ben Mazziane, S. Alouf, G. Neglia and D. S. Menasche. ‘TTL model for an LRU-based similarity caching policy’. In: *Computer Networks* 241 (Mar. 2024), p. 110206. DOI: 10.1016/j.comnet.2024.110206. URL: https://inria.hal.science/hal-04746044.

[5] V. Bucarey López, E. Della Vecchia, A. Jean-Marie and F. Ordoñez. 'Stationary Strong Stackelberg Equilibrium in Discounted Stochastic Games'. In: *IEEE Transactions on Automatic Control* 68.9 (2023), pp. 5271–5286. DOI: 10.1109/TAC.2022.3220512. URL: https://inria.hal.science/hal-03934114.

[6] M. Datar, E. Altman and H. Le Cadre. 'Strategic Resource Pricing and Allocation in a 5G Network Slicing Stackelberg Game'. In: *IEEE Transactions on Network and Service Management* 20.1 (2023), pp. 502–520. DOI: 10.1109/TNSM.2022.3216588. URL: https://inria.hal.science/hal-03824540.

[7] O. Marfoq, G. Neglia, A. Bellet, L. Kameni and R. Vidal. 'Federated Multi-Task Learning under a Mixture of Distributions'. In: NeurIPS 2021 - 35th Conference on Neural Information Processing Systems. Sydney / Virtual, Australia, 6th Dec. 2021. URL: https://hal.science/hal-03406994.

[8] S. M. Perlaza, G. Bisson, I. Esnaola, A. Jean-Marie and S. Rini. 'Empirical Risk Minimization with Relative Entropy Regularization'. In: *IEEE Transactions on Information Theory* 70.7 (2024), pp. 5122–5161. DOI: 10.1109/TIT.2024.3365728. URL: https://hal.science/hal-03849748.

[9] T. Si Salem, G. Neglia and S. Ioannidis. 'No-regret Caching via Online Mirror Descent'. In: *ACM Transactions on Modeling and Performance Evaluation of Computing Systems* 8.4 (11th Aug. 2023), pp. 1–32. DOI: 10.1145/3605209. URL: https://hal.science/hal-04181387.

[10] G. Vardoyan, P. Nain, S. Guha and D. Towsley. 'On the Capacity Region of Bipartite and Tripartite Entanglement Switching'. In: *ACM Transactions on Modeling and Performance Evaluation of Computing Systems* 8.1-2 (June 2023), pp. 1–18. DOI: 10.1145/3571809. URL: https://inria.hal.science/hal-04018593.

[11] X. Zou, S. M. Perlaza, I. Esnaola, E. Altman and H. V. Poor. 'The Worst-Case Data-Generating Probability Measure in Statistical Learning'. In: *IEEE Journal on Selected Areas in Information Theory* 5 (2024), pp. 175–189. DOI: 10.1109/JSAIT.2024.3383281. URL: https://inria.hal.science/hal-04442591.

12.2 Publications of the year

International journals

[12] K. Avrachenkov and M. Dreveton. 'Almost exact recovery in noisy semi-supervised learning'. In: *Probability in the Engineering and Informational Sciences* (11th Nov. 2024), pp. 1–22. DOI: 10.1017/s0269964824000135. URL: https://inria.hal.science/hal-04776811 (cit. on p. 10).

[13] K. Avrachenkov, M. Dreveton and L. Leskelä. 'Recovering Static and Time-Varying Communities Using Persistent Edges'. In: *IEEE Transactions on Network Science and Engineering* 11.2 (23rd Feb. 2024), pp. 2087–2099. DOI: 10.1109/tnse.2023.3337281. URL: https://inria.hal.science/hal-04755767 (cit. on p. 9).

[14] Y. Ben Mazziane, S. Alouf, G. Neglia and D. S. Menasche. 'TTL model for an LRU-based similarity caching policy'. In: *Computer Networks* 241 (Mar. 2024), p. 110206. DOI: 10.1016/j.comnet.2024.110206. URL: https://inria.hal.science/hal-04746044 (cit. on p. 18).

[15] F. Cabo, A. Jean-Marie and M. Tidball. 'Positional and conformist effects in voluntary public good provision'. In: *Journal of Public Economic Theory* 26.2 (17th Mar. 2024). DOI: 10.1111/jpet.12685. URL: https://hal.umontpellier.fr/hal-04754332 (cit. on p. 17).

[16] A. Dejonghe, Z. Altman, F. De Pellegrini and E. Altman. 'Design and cross-layer optimization of low cost RIS-assisted communication systems'. In: *IEEE Transactions on Wireless Communications* 23.10 (30th May 2024). DOI: 10.1109/TWC.2024.3404231. URL: https://hal.science/hal-04601576 (cit. on p. 18).

[17] S. H. S. Dizaji, K. Patil and K. Avrachenkov. 'Influence Maximization in Dynamic Networks Using Reinforcement Learning'. In: *SN Computer Science* 5.1 (8th Jan. 2024), p. 169. DOI: 10.1007/s42979-023-02453-1. URL: https://inria.hal.science/hal-04755717 (cit. on p. 13).

[18] F. Faticanti and G. Neglia. 'Optimistic online caching for batched requests'. In: *Computer Networks* 244 (May 2024), p. 110341. DOI: 10.1016/j.comnet.2024.110341. URL: https://hal.science/hal-04763270 (cit. on pp. 17, 32).

[19] A. Fox, F. D. Pellegrini and E. Altman. 'Learning optimal edge processing with offloading and energy harvesting'. In: *Computer Communications* 225 (Sept. 2024), pp. 324–338. DOI: 10.1016/j.comcom.2024.07.009. URL: https://inria.hal.science/hal-04814405 (cit. on p. 19).

[20] D. Goldsztajn. 'Fluid Limits and Optimal Task Assignment Policies for Locally Pooled Service Systems'. In: *ACM SIGMETRICS Performance Evaluation Review* 52.3 (Dec. 2024), pp. 7–10. DOI: 10.1145/3712170.3712174. URL: https://inria.hal.science/hal-04929573.

[21] D. Goldsztajn, S. Borst and J. van Leeuwaarden. 'Learning and Balancing Unknown Loads in Large-Scale Systems'. In: *Mathematics of Operations Research* (3rd May 2024). DOI: 10.1287/moor.2021.0212. URL: https://inria.hal.science/hal-04830673 (cit. on p. 8).

[22] D. Goldsztajn, S. Borst and J. van Leeuwaarden. 'Server Saturation in Skewed Networks'. In: *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 8.2 (29th May 2024), pp. 1–37. DOI: 10.1145/3656018. URL: https://inria.hal.science/hal-04830708 (cit. on p. 8).

[23] P. Howlett, B. K. Beare, M. Franchi, J. Boland and K. Avrachenkov. 'The Granger–Johansen representation theorem for integrated time series on Banach space'. In: *Journal of Time Series Analysis* (18th Aug. 2024). DOI: 10.1111/jtsa.12766. URL: https://inria.hal.science/hal-04755738 (cit. on p. 9).

[24] A. Jean-Marie and M. Tidball. 'Multiple Exploitation of a Renewable Resource — Optimal Solution of a Deterministic Singular Control Model'. In: *International Game Theory Review* 26.02 (22nd Mar. 2024). DOI: 10.1142/S0219198924400073. URL: https://inria.hal.science/hal-04716327 (cit. on p. 5).

[25] S. M. Perlaza, G. Bisson, I. Esnaola, A. Jean-Marie and S. Rini. 'Empirical Risk Minimization with Relative Entropy Regularization'. In: *IEEE Transactions on Information Theory* 70.7 (2024), pp. 5122–5161. DOI: 10.1109/TIT.2024.3365728. URL: https://hal.science/hal-03849748 (cit. on p. 10).

[26] C. Rajput and H. Manjunath. 'Metallic means and Pythagorean triples'. In: *Notes on Number Theory and Discrete Mathematics* 30.1 (31st Mar. 2024), pp. 184–194. DOI: 10.7546/nntdm.2024.30.1.184-194. URL: https://inria.hal.science/hal-04811668.

[27] F. Robledo Relaño, V. Borkar, U. Ayesta and K. Avrachenkov. 'Tabular and Deep Learning for the Whittle Index'. In: *ACM Transactions on Modeling and Performance Evaluation of Computing Systems* (3rd June 2024). DOI: 10.1145/3670686. URL: https://hal.science/hal-04599291 (cit. on pp. 12, 23).

[28] X. Ye, I. Esnaola, S. M. Perlaza and R. F. Harrison. 'An Information Theoretic Metric for Measurement Vulnerability to Data Integrity Attacks on Smart Grids'. In: *IET Smart Grid* 7.5 (1st Oct. 2024), pp. 583–592. DOI: 10.1049/stg2.12163. URL: https://hal.science/hal-03849768 (cit. on pp. 9, 32).

[29] X. Zou, S. M. Perlaza, I. Esnaola, E. Altman and H. V. Poor. 'The Worst-Case Data-Generating Probability Measure in Statistical Learning'. In: *IEEE Journal on Selected Areas in Information Theory* 5 (2024), pp. 175–189. DOI: 10.1109/JSAIT.2024.3383281. URL: https://inria.hal.science/hal-04442591 (cit. on pp. 11, 32, 33).

International peer-reviewed conferences

[30] K. Agarwal and E. Altman. 'Games among selfish and team stations in polling systems'. In: VALUETOOLS 2024 - 17th EAI International Conference on Performance Evaluation Methodologies and Tools. Milan, Italy, 12th Dec. 2024. URL: https://inria.hal.science/hal-04829591 (cit. on pp. 16, 23, 33).

[31] E. M. Athanasakos and S. M. Perlaza. 'Leveraging Noisy Observations in Zero-Sum Games'. In: IEEE Information Theory Workshop (ITW). Shenzhen, China, 24th Nov. 2024. URL: https://inria.hal.science/hal-04432819 (cit. on p. 16).

[32] E.-M. Athanasakos, N. Kalouptsidis and H. Manjunath. 'Local Approximation of Secrecy Capacity'. In: 2024 32nd European Signal Processing Conference (EUSIPCO). Lyon, France, 26th Aug. 2024, pp. 745–749. DOI: 10.23919/EUSIPCO63174.2024.10715107. URL: https://inria.hal.science/hal-04811691 (cit. on p. 17).

[33] O. Boufous, R. El-Azouzi, M. Touati, E. Altman and M. Bouhtou. 'Constrained Correlated Equilibria'. In: 60th Annual Allerton Conference on Communication, Control, and Computing. Urbana, United States, 24th Sept. 2024, pp. 1–8. DOI: 10.1109/Allerton63246.2024.10735321. URL: https://inria.hal.science/hal-04814763 (cit. on p. 16).

[34] F. Busacca, S. Mangione, G. Neglia, I. Tinnirello, S. Palazzo and F. Restuccia. 'FedLoRa: IoT Spectrum Sensing Through Fast and Energy-Efficient Federated Learning in LoRa Networks'. In: MASS 2024 - IEEE 21st International Conference on Mobile Ad-Hoc and Smart Systems. Seoul, South Korea, 23rd Sept. 2024, pp. 295–303. DOI: 10.1109/MASS62177.2024.00047. URL: https://inria.hal.science/hal-04762819 (cit. on pp. 14, 32, 34).

[35] M. Costantini, G. Neglia and T. Spyropoulos. 'FedDec: Peer-to-peer Aided Federated Learning'. In: SPAWC 2024 - IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications. Lucca, Italy, 10th Sept. 2024, pp. 426–430. DOI: 10.1109/SPAWC60668.2024.10694344. URL: https://inria.hal.science/hal-04762825 (cit. on pp. 14, 32, 34).

[36] F. Daunas, I. Esnaola, S. M. Perlaza and H. V. Poor. 'Equivalence of the Empirical Risk Minimization to Regularization on the Family of f-Divergences'. In: Proceedings of the IEEE International Symposium on Information Theory. Athens, Greece, 7th July 2024, pp. 759–764. DOI: 10.1109/ISIT57864.2024.10619260. URL: https://hal.science/hal-04431558 (cit. on pp. 11, 32).

[37] F. Diana, O. Marfoq, C. Xu, G. Neglia, F. Giroire and E. Thomas. 'Attribute Inference Attacks for Federated Regression Tasks'. In: 39th Annual AAAI Conference on Artificial Intelligence - AAAI 2025. Philadelphia (Pennsylvania), United States, 25th Feb. 2025. URL: https://inria.hal.science/hal-04878082 (cit. on p. 15).

[38] A. Fox, F. de Pellegrini, F. Faticanti, E. Altman and F. Bronzino. 'Optimal Flow Admission Control in Edge Computing via Safe Reinforcement Learning'. In: WiOpt 2024 - 22nd International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks. Seoul, South Korea, 21st Oct. 2024. URL: https://hal.science/hal-04786022 (cit. on pp. 19, 33).

[39] F. Galante, G. Neglia and E. Leonardi. 'Scalable Decentralized Algorithms for Online Personalized Mean Estimation'. In: 39th Annual AAAI Conference on Artificial Intelligence - AAAI 2025. Philadelphia (Pennsylvania), United States, 25th Feb. 2025. URL: https://hal.science/hal-04878986 (cit. on pp. 14, 32).

[40] D. Goldsztajn, S. Borst and J. van Leeuwaarden. 'Server Saturation in Skewed Networks'. In: SIGMETRICS/PERFORMANCE 2024 - Joint International Conference on Measurement and Modeling of Computer Systems. Venice, Italy, 10th June 2024, pp. 97–98. DOI: 10.1145/3652963.3655080. URL: https://hal.science/hal-04830700 (cit. on p. 8).

[41] L. Hauseux. 'How can we theoretically measure the performance of density-based clustering algorithms?' In: ACM SIGMETRICS 2024 Student Research Competition. Venice, Italy, 10th June 2024. URL: https://hal.science/hal-04674019 (cit. on pp. 6, 12).

[42] L. Hauseux, K. Avrachenkov and J. Zerubia. 'Benefits of hypergraphs for density-based clustering'. In: EUSIPCO 2024 - 32nd IEEE European Signal Processing Conference. Lyon, France, 26th Aug. 2024, pp. 2302–2306. DOI: 10.23919/EUSIPCO63174.2024.10715271. URL: https://hal.science/hal-04617936 (cit. on p. 12).

[43] L. Hauseux, K. Avrachenkov and J. Zerubia. 'Hypergraphs, percolation, and hierarchical clustering'. In: Complex Networks 2024 - 13th International Conference on Complex Networks and their Applications. Istanbul, Turkey, 10th Dec. 2024. URL: https://inria.hal.science/hal-04723052 (cit. on p. 12).

[44] C. G. Kaplan, C. Xu, O. Marfoq, G. Neglia and A. Santana de Oliveira. 'A Cautionary Tale: On the Role of Reference Data in Empirical Privacy Defenses'. In: PETS 2024 - 24th Privacy Enhancing Technologies Symposium. Vol. 2024. 1. Bristol, United Kingdom, 15th July 2024, pp. 525–548. DOI: 10.56553/popets-2024-0031. URL: https://inria.hal.science/hal-04765695 (cit. on pp. 15, 22, 32).

[45] B. Le Bars, A. Bellet, M. Tommasi, K. Scaman and G. Neglia. 'Improved Stability and Generalization Guarantees of the Decentralized SGD Algorithm'. In: ICML 2024 - The Forty-first International Conference on Machine Learning. Vienne, Austria, 21st July 2024. URL: https://hal.science/hal-04611418 (cit. on pp. 14, 32, 34).

[46] A. R. Ramtin, J. Z. Hare, L. Kaplan, P. Nain, V. Veeravalli and D. Towsley. 'Quickest Change Detection in the Presence of Covert Adversaries'. In: 4th International Workshop on the Internet of Things for Adversarial Environments, Washington DC, United States, 28th Oct. 2024. URL: https://inria.hal.science/hal-04744798 (cit. on p. 19).

[47] F. Robledo, U. Ayesta and K. Avrachenkov. 'Deep reinforcement learning for weakly coupled MDP's with continuous actions'. In: ASMTA 2024 - Analytical and Stochastic Modelling Techniques and Applications. Vol. 14826. Lecture Notes in Computer Science. Venise, Italy, 13th Sept. 2025, pp. 67–80. DOI: 10.1007/978-3-031-70753-7_5. URL: https://hal.science/hal-04594762 (cit. on p. 13).

[48] A. Rodio and G. Neglia. 'FedStale: leveraging Stale Updates in Federated Learning'. In: ECAI 2024 - 27TH European Conference on Artificial Intelligence. Vol. 392. Santiago De Compostela, Spain, 19th Oct. 2024, pp. 3071–3078. DOI: 10.3233/FAIA240849. URL: https://inria.hal.science/hal-04762831 (cit. on pp. 13, 32, 34).

[49] X. Zou, S. M. Perlaza, I. Esnaola and E. Altman. 'Generalization Analysis of Machine Learning Algorithms via the Worst-Case Data-Generating Probability Measure'. In: AAAI 2024 - Conference on Artificial Intelligence. Vol. 38. 15. Vancouver, Canada, 24th Mar. 2024, pp. 17271–17279. DOI: 10.1609/aaai.v38i15.29674. URL: https://inria.hal.science/hal-04353957 (cit. on pp. 11, 33).

Edition (books, proceedings, special issue of a journal)

[50] S. Alouf, O. Hohlfeld and Z. Jiang, eds. *Preface: Special issue on ITC 2023* 167 (Mar. 2025). DOI: 10.1016/j.peva.2024.102462. URL: https://inria.hal.science/hal-04731314 (cit. on p. 37).

Doctoral dissertations and habilitation theses

[51] Y. Ben Mazziane. 'Probabilistic analysis for caching'. Université Côte d'Azur, 13th May 2024. URL: https://theses.hal.science/tel-04681458 (cit. on pp. 17, 40).

[52] O. Chuchuk. 'Data access optimisation at CERN and in the Worldwide LHC Computing Grid (WLCG)'. Université Côte d'Azur, 19th Feb. 2024. URL: https://theses.hal.science/tel-04620247 (cit. on p. 40).

[53] C. Kaplan. 'Inherent trade-offs in privacy-preserving machine learning'. Université Côte d'Azur, 22nd Nov. 2024. URL: https://theses.hal.science/tel-04874804 (cit. on pp. 22, 40).

[54] A. Rodio. 'Client heterogeneity in federated learning systems'. Université Côte d'Azur, 3rd July 2024. URL: https://theses.hal.science/tel-04685040 (cit. on pp. 13, 40).

Reports & preprints

[55] E.-M. Athanasakos and S. M. Perlaza. *Leveraging Noisy Observations in Zero-Sum Games*. RR-9538. Inria Sophia Antipolis - Méditerranée, Université Côte d'Azur, 29th Jan. 2024, p. 24. URL: https://inria.hal.science/hal-04424590 (cit. on p. 16).

[56] D. Carra and G. Neglia. *An Online Gradient-Based Caching Policy with Logarithmic Complexity and Regret Guarantees*. 2024. URL: https://inria.hal.science/hal-04762853 (cit. on p. 18).

[57] F. Daunas, I. Esnaola, S. M. Perlaza and H. V. Poor. *Asymmetry of the Relative Entropy in the Regularization of Empirical Risk Minimization*. 3rd Oct. 2024. URL: https://inria.hal.science/hal-04719783 (cit. on pp. 11, 32).

[58] F. Diana, O. Marfoq, C. Xu, G. Neglia, F. Giroire and E. Thomas. *Attribute Inference Attacks for Federated Regression Tasks*. 2024. DOI: 10.48550/arXiv.2411.12697. URL: https://inria.hal.science/hal-04792394 (cit. on pp. 32, 34).

[59] F. Galante, G. Neglia and E. Leonardi. *Scalable Decentralized Algorithms for Online Personalized Mean Estimation*. 2024. URL: https://inria.hal.science/hal-04762852.

[60] A. Jean-Marie and M. Tidball. *Equilibrium bids for reverse auctions when the budget is announced - Some preliminary results*. RR-9545. Inria; CEE-M. Centre d'Economie de l'Environnement - Montpellier, Mar. 2024, p. 19. URL: https://inria.hal.science/hal-04520388 (cit. on p. 17).

[61] C. Kaplan, A. Rodio, T. S. Salem, C. Xu and G. Neglia. *Federated Learning for Collaborative Inference Systems: The Case of Early Exit Networks*. 2024. URL: https://inria.hal.science/hal-04762846 (cit. on pp. 15, 22, 32).

[62] S. M. Perlaza and X. Zou. *The Generalization Error of Machine Learning Algorithms*. 18th Nov. 2024. URL: https://inria.hal.science/hal-04789606 (cit. on pp. 10, 32, 33).

[63] C. Rodriguez, L. Degioanni, L. Kameni, R. Vidal and G. Neglia. *Evaluating the Energy Consumption of Machine Learning: Systematic Literature Review and Experiments*. 2024. URL: https://inria.hal.science/hal-04762843 (cit. on pp. 6, 15, 20).

[64] X. Zou, S. M. Perlaza, I. Esnaola, E. Altman and H. V. Poor. *An Exact Characterization of the Generalization Error of Machine Learning Algorithms*. RR-9539. INRIA, 28th Mar. 2024, p. 80. URL: https://inria.hal.science/hal-04442566 (cit. on pp. 11, 32, 33).

Educational activities

[65] A. Jean-Marie and E. Hyon. 'An introduction to pyMarmote and pyMarmoteMDP for Markovian modeling - A tutorial'. Doctoral. Italy, 10th June 2024. URL: https://inria.hal.science/hal-04747584 (cit. on pp. 9, 38).