EURECOM
Sophia Antipolis

Research activity report 2020

Data Science
Digital Security
Communication systems

EURECOM
CAMPUS SOPHIATECH

GRADUATE SCHOOL AND RESEARCH CENTER IN DIGITAL SCIENCE

Table des matières

DIGITAL SECURITY DEPARTMENT

Introduction

This report summarizes the main research activities carried out in 2020 in the Digital Security Department at EURECOM.

Despite the difficult times due to the pandemic, this was an exceptional year for the department.

Two new assistant professors joined the Cryptography and Privacy Enhancing Technologies team, expanding the activity of the group in the area of privacy preserving biometrics and cryptography. One assistant professor has instead left the Software and System Security team to pursue a career in the industry, but the position has already been filled with a new candidate who will join the department in 2021. In addition to the faculty members, during 2020 the department employed 30 PhD students and 10 Postdoc and Research Associates.

Overall, the research was funded by 4 European Projects, 11 National Projects, and 17 Industrial Contracts - which include fundings from DARPA and the US AirForce, as well as from different companies worldwide. The department also continued its ongoing training activity with UPNM and worked closely with Nozomi Networks, an American company leader in the OT/IoT market, to bootstrap its research department and identify long term research problems and future joint research projects with EURECOM.

This report discusses the research conducted in the Digital Security Department in 2020, by covering a total of 44 publications divided along three main thematic areas:

- System and Software Security

This area focuses on the security of systems, including networks and mobile devices. The work conducted this year in the area revolves mostly around malware and binary analysis, embedded system security, web security, and botnet detection.

- Cryptography and Privacy Enhancing Technologies

This newly renamed area has grown both in size and in scope in 2020. Its activities now span from privacy preserving neural networks to cryptography, and from crypto-biometrics to leak detection.

- Biometrics and Digital Media

This third area covers the application of image, video, and audio signals for security. In 2020, research in the area has focused primarily on biometrics, image forensics, and deepfake identification, as well as on privacy, anonymisation and pseudonymisation.

1. SYSTEM AND SOFTWARE SECURITY

Faculty members: Davide Balzarotti, Marc Dacier, Aurélien Francillon, Massimiliano Todisco, and Yanick Fratantonio

The research conducted along the axis of software and system security in 2020 can be summarized around the following four main topics:

1. Malware
2. Embedded Systems Security
3. Web Security
4. Botnets Detection

1.1 Malware

Research team
Simone Aonzo, Davide Balzarotti, Emanuele Cozzi, Alessandro Mantovani
Research Conducted in Collaboration with
CISCO, Norton LifeLock, University of Genoa, University of California - Santa Barbara,

In 2020 we continued our research to support data-driven malware analysis. In particular, we first looked at an open research problem in the malware analysis field: how to statically distinguish between packed and non-packed executables. This has an impact on antivirus software and malware analysis systems, which may need to apply different heuristics or to resort to more costly code emulation solutions to deal with the presence of potential packing routines. It can also affect the results of many research studies in which the authors adopt algorithms that are specifically designed for packed or non-packed binaries. Therefore, a wrong answer to the question "is this executable packed?" can make the difference between malware evasion and detection. It has long been known that packing and entropy are strongly correlated, often leading to the wrong assumption that a low entropy score implies that an executable is NOT packed. Exceptions to this rule exist, but they have always been considered as one-off cases, with a negligible impact on any large scale experiment. However, if such an assumption might have been acceptable in the past, our experiments show that this is not the case anymore as an increasing and remarkable number of packed malware samples implement proper schemes to keep their entropy low. In a paper published at the NDSS security conference [MA20], we empirically investigate and measure this problem by analyzing a dataset of 50K low-entropy Windows malware samples. Our tests show that, despite the fact that all samples had a low entropy value, over 30% of them adopted some form of runtime packing. We then extended our analysis beyond the pure entropy, by considering all static features that have been proposed so far to identify packed code. Again, our tests show that even a state of the art machine learning classifier is unable to conclude whether a low-entropy sample is packed or not by relying only on features extracted with static analysis.

While remaining on the important subject of packing, we then look at its impact on machine learning classifiers. In fact, machine learning techniques are widely used in addition to signatures and heuristics to increase the detection rate of anti-malware software, as they automate the creation of detection models, making it possible to handle an ever-increasing number of new malware samples. Because of this, packing is a common technique used by malware authors to foil the analysis of anti-malware systems and evade detection. However, benign applications use packing and obfuscation as well, to protect intellectual property and prevent license abuse. Malware researchers have often assumed that packing would prevent machine learning techniques from building effective classifiers. However, both industry and academia have published results that show that machine-learning-based classifiers can achieve good detection rates, leading many experts to think that classifiers are simply detecting the fact that a sample is packed, as packing is more prevalent in malicious samples.

To better understand this phenomenon, in a second paper we published at the NDSS symposium [AG20], we studied how machine learning based on static analysis features operates on packed samples. We show that, different from what is commonly assumed, packers do preserve some information when packing programs that is "useful" for malware classification. However, this information does not necessarily capture the sample's behavior. Our work also demonstrates that the signals extracted from packed executables are not rich enough for machine-learning-based models to (1) generalize their knowledge to operate on unseen packers, and (2) be robust against adversarial examples. Finally, our paper shows that

a naive application of machine learning techniques results in a substantial number of false positives, which, in turn, might have resulted in incorrect labeling of ground-truth data used in past work.

In parallel to these studies on packing, in 2020 our group also continued to conduct research on IoT malware. As we have shown in our previous research, the recent emergence of IoT devices and the rise of large-scale IoT botnets has dramatically increased the volume and sophistication of Linux-based malware observed in the wild. The security community has put a lot of effort to document these threats but analysts mostly rely on manual work, which makes it difficult to scale and hard to regularly maintain. Moreover, the vast amount of code reuse that characterizes IoT malware calls for an automated approach to detect similarities and identify the phylogenetic tree of each family. In a paper presented at the ACSAC conference [CO20-2] we presented the largest measurement of IoT malware conducted to date. In our work, we systematically reconstructed – through the use of binary code similarity – the lineage of IoT malware families, and we tracked their relationships, evolution, and variants. We applied our technique on a dataset of more than 93k samples submitted to VirusTotal over a period of 3.5 years and discussed the findings and the tangled relationship among the existing families.

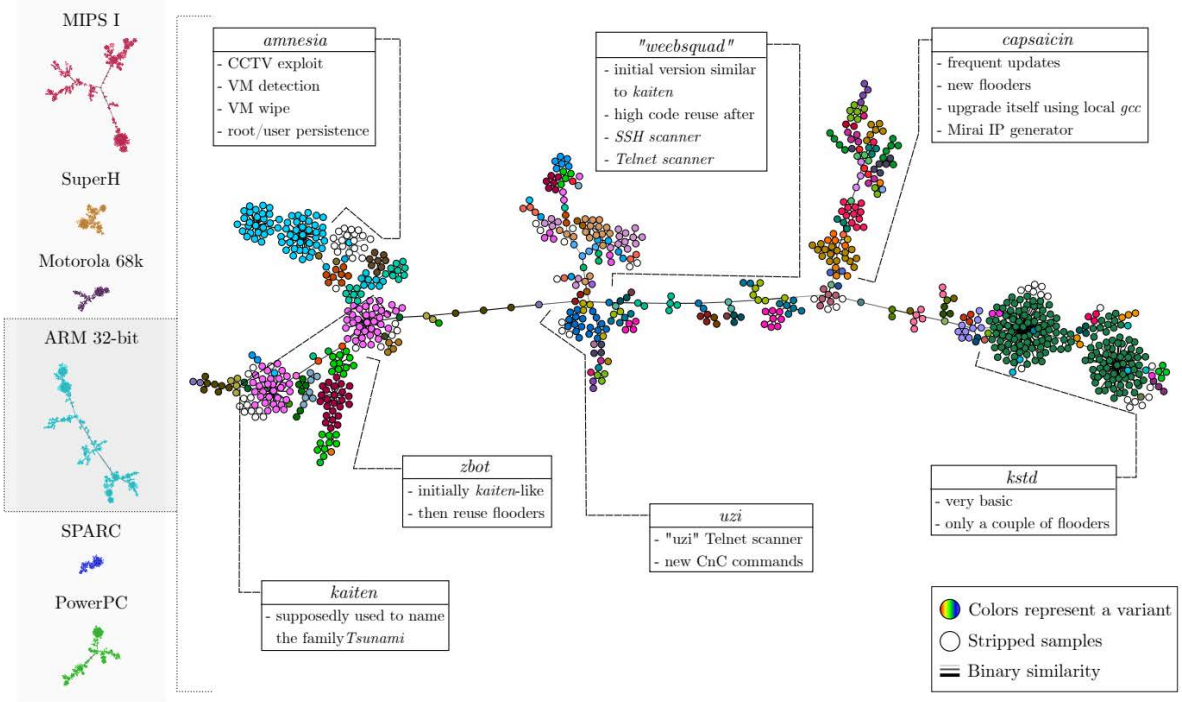


Fig.1 - Lineage graph of the Tsunami malware family (focus on the ARM 32-bit architecture)

1.2 Embedded System Security

Research team
Aurélien Francillon, Paul Olivier, Sebastian Pöplau, Giovanni Camurati, Nassim Corteggiani
Research Conducted in Collaboration with
Université catholique de Louvain, Siemens AG, Maxim Integrated

1.2.1 Embedded device security testing

The growing complexity of Systems-on-Chip challenges our ability to ensure their correct operation, on which we rely on for more and more sensitive activities. Many security vulnerabilities appear in subtle and unexpected ways in the interaction among blocks and across layers, where current verification tools fail at catching them or do not scale. For this reason, security evaluation still heavily relies on manual review. Students from the group won the Hack@DAC19 contest (in the student category), organized at the DAC conference. Inspired by this experience, we presented our reflections on this topic

from a software and system security perspective in an invited journal paper published at IEEE Design & Test [CO20-3]. The paper outlines an approach that extends the dynamic analysis of firmware to the hardware.

1.2.2 Embedded device security testing with Hardware Snapshots

Advanced dynamic analysis techniques such as fuzzing and Dynamic Symbolic Execution (DSE) are a cornerstone of software security testing and are becoming popular with embedded systems testing. Testing software in a virtual machine provides more visibility and control. VM snapshots also save testing time by facilitating crash reproduction, performing root cause analysis and avoiding re-executing programs from the start. However, because embedded systems are very diverse virtual machines that perfectly emulate them are often unavailable. Previous work therefore either attempts to model hardware or perform partial emulation (forwarding interaction to the real hardware), which leads to inaccurate or slow emulation. However, such limitations are unnecessary when the whole design is available, e.g., to the device manufacturer or on open hardware. In a paper published at DSN conference [CO20-1], we propose a novel approach, called HardSnap, for co-testing hardware and software with a high level of introspection. HardSnap aims at improving security testing of hardware/software co-designed systems, where embedded systems designers have access to the whole HW/SW stack. HardSnap is a virtual-machine-based solution that extends visibility and controllability to the hardware peripherals with a negligible overhead. HardSnap introduces the concept of a hardware snapshot that collects the hardware state (together with software state). In our prototype, Verilog hardware blocks are either simulated in software or synthesized to an FPGA. In both cases, HardSnap is able to generate HW/SW snapshots on demand. HardSnap is designed to support new peripherals automatically, to have high performance, and full controllability and visibility on software and hardware. We evaluated HardSnap on open-source peripherals and synthetic firmware to demonstrate improved ability to find and diagnose security issues

1.2.3 Screaming Channels improvement

Last year we introduced a new class of side-channel vulnerability, called Screaming Channels. These leaks might appear if the sensitive leaks from the processor are unintentionally broadcasted by a radio transmitter placed on the same chip. Previous work focuses on identifying the root causes, and on mounting an attack at a distance considerably larger than the one achievable with conventional electromagnetic side channels, which was demonstrated in the low-noise environment of an anechoic chamber. However, a detailed understanding of the leak, attacks that take full advantage of the novel vector, and security evaluations in more practical scenarios are still missing. In a paper published in IACR Transactions on Cryptographic Hardware and Embedded Systems (CHES 2020) [CA20], we conduct a thorough experimental analysis of the peculiar properties of Screaming Channels. For example, we learn about the coexistence of intended and unintended data, the role of distance and other parameters on the strength of the leak, the distortion of the leak model, and the portability of the profiles. With such insights, we build better attacks. We profile a device connected via cable with 10000·500 traces. Then, 5 months later, we attack a different instance at 15 m in an office environment. We recover the AES-128 key with 5000·1000 traces and key enumeration up to 223. Leveraging spatial diversity, we mount some attacks in the presence of obstacles. As a first example of application to a real system, we show a proof-of-concept attack against the authentication method of Google Eddystone beacons. On the one side, this work lowers the bar for more realistic attacks, highlighting the importance of the novel attack vector. On the other side, it provides a broader security evaluation of the leaks, helping the defender and radio designers to evaluate risk, and the need of countermeasures. The vulnerability was disclosed to Google, who made an Honorable Mention in their Bughunter Hall of Fame.

1.3 Web Security

Research team
Davide Balzarotti, Matteo dell'Amico
Research Conducted in Collaboration with
Norton LifeLock, University of California - Santa Barbara

In 2020 we published two papers related to the privacy of web users. In the first, we looked at how from one side web pages have evolved into very complex dynamic applications, which are often very opaque and difficult for non-experts to understand. However, at the same time, security researchers push for more transparent web applications, which can help users in taking important security-related decisions about which information to disclose, which link to visit, and which online service to trust. In a paper published at the Web Conference (previously known as WWW) [SR20-2], we looked at one of the most simple but also most representative aspects that captures the struggle between these opposite demands: a mouse

click. In particular, we presented the first comprehensive study of the possible security and privacy implications that clicks can have from a user perspective, by analyzing the disconnect that exists between what is shown to users and what actually happens after each click. We started by identifying and classifying possible problems. We then implemented a crawler that performed nearly 2.5M clicks looking for signs of misbehavior and we analyzed all the interactions created as a result of those clicks. Our experiments show that the vast majority of domains are putting users at risk by either obscuring the real target of links or by not providing sufficient information for users to make an informed decision.

The second work was an extension of a paper we published in 2019 called "BakingTimer: Privacy Analysis of Server-Side Request Processing Time". The paper looks at how cookies, originally introduced as a way to provide state awareness to websites, are now one of the backbones of the current web. While cookies are often deliberately used for web tracking, user analytics, and for online advertisement (with the subsequent privacy loss for the end users), they are not the only technique capable of retrieving the users' browsing history. In fact, history sniffing techniques are capable of tracking the users' browsing history without relying on any specific code in a third-party website, but only on code executed within the visited site. In the BakingTimer paper we proposed a new history sniffing technique based on timing the execution of server-side request processing code. The extended version published on the ACM Digital Threats journal [SR20-1], additionally includes a comprehensive analysis of existing countermeasures, starting from its evolution and adoption, and finishing with a large-scale experiment to assess the impact on the presented technique.

1.4 Botnets detection and mitigation

Research team
Marc Dacier, Massimiliano Todisco, Elisa Chiapponi, Antonino Vitale
Research Conducted in Collaboration with
Amadeus

In this research activity, we have partnered with Amadeus to help them address a very severe and expensive problem they are facing: armies of bots scraping their websites.

The Internet has discovered the existence of botnets and the nuisance they can cause in February 2000 with the early DDoS attacks against Yahoo!, Amazon.com, CNN.com, and other major Web sites. They have continuously evolved from relatively rudimentary pieces of software to very sophisticated components such as the numerous ``all in one sneaker bots" (eg., aiobot.com) that automate the buying process of luxury goods in high demands. To increase their resilience, the conjecture is that the bots take advantage of proxying services publicly available on the web, for a fee. A 2019 Imperva report describes very clearly how the airline industry is particularly impacted by these armies of bots. In 2017, according to that report, the proportion of bad bots traffic to airline websites was 43.9 percent. Almost a third of these bad bots were sophisticated ones, referred to as Advanced Persistent Bots (APBs). APBs can mimic human behavior, load JavaScript and external assets, tamper with cookies, perform browser automation, give the impression that the mouse is moving on the screen, etc.

Almost all these bots are used to gather free information from the airlines' sites about flights and ticket prices. The conjecture is that unauthorized business intelligence companies, online travel agencies, and data aggregators are beyond such bots activities because a large part of their business relies on web scraping. The damages to the business model, as well as the costs incurred to support these bot requests, explain the explosion of a whole new ecosystem of anti-bots techniques. An arms race exists between bot makers and anti-bot providers: as soon as a family of bots is blocked, their operators replace them with a new one which defeats the protection for a varying amount of time.

Rather than trying to detect the bots themselves, we have instead looked into how we could impact the business model of those using these bots. To do so, we have built a platform where identified bots are provided inaccurate information, at a cheap cost for the data service provider, so that the attackers are the ones consuming needlessly their resources without any hope for a return on their investment. The first results of this joint work have been published in [CH20-1].

While carrying this work, we have confirmed that the bots take advantage of proxy services. Thanks to these services, the bots use temporary IP addresses that are owned and used by legit users. There are, supposedly, tens of millions of such IPs made available to bots. Would the targeted websites decide to block each IP which is considered to behave like a bot, they would quickly deny access to millions of IPs, some of them belonging to potential customers. Clearly, at first sight, an IP

blocking solution does not appear to be a viable approach due to the, supposedly, sheer volume of IPs, available all over the world. However, the data at our disposal seemed to tell another story. Using a real-world setup, we have had access to the logs of close to 20 heavily targeted websites and have carried out an experiment over a two months period. Based on the gathered empirical pieces of evidence, we have proposed mathematical models that indicate that the amount of IPs is likely 2 to 3 orders of magnitude smaller than the one claimed by the proxy services. This finding suggests that an IP reputation-based blocking strategy could be effective, contrary to what operators of these websites think today. These results have been published in [CH20-2].

Last but not least, we have leveraged the TTCN3 environment to build a distributed infrastructure to carry out in 2021 a large experiment to validate the claims made in [CH20-2]. TTCN-3 (Testing and Test Control Notation version 3) is a strongly typed testing language used in conformance testing of communicating systems. TTCN-3 has been defined by ETSI and standardized by ITU-T. We will use it to test several proxy services and uncover networking fingerprints that would enable us, at the web server site, to recognize requests that came through such proxies, thus most likely originating from bots.

1.5 Other Topics

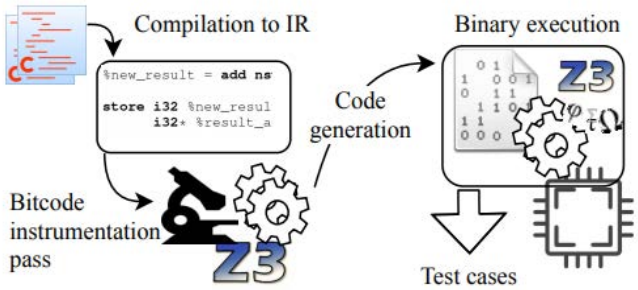
Research team
Emna Abid, Davide Balzarotti, Marc Dacier, Savino Dambra, Dhia Farrah, Aurélien Francillon, Yanick Fratantonio, Sebastian Poeplau, Andrea Possemato
Research Conducted in Collaboration with
IDEMIA, Norton LifeLock

1.5.1 Cyber insurance

Cyber attacks have increased in number and complexity in recent years, and companies and organizations have accordingly raised their investments in more robust infrastructure to preserve their data, assets and reputation. However, the full protection against these countless and constantly evolving threats is unattainable by the sole use of preventive measures. Therefore, to handle residual risks and contain business losses in case of an incident, firms are increasingly adopting a cyber insurance as part of their corporate risk management strategy. As a result, the cyber insurance sector – which offers to transfer the financial risks related to network and computer incidents to a third party – is rapidly growing, with recent claims that already reached a $100M dollars. However, while other insurance sectors rely on consolidated methodologies to accurately predict risks, the many peculiarities of the cyber domain resulted in carriers to often resort to qualitative approaches based on experts opinions. I a paper published at the IEEE Symposium on Security & Privacy [DA20], we looked at past research conducted in the area of cyber insurance and classifies previous studies in four different areas, focused respectively on studying the economical aspects, the mathematical models, the risk management methodologies, and the predictions of cyber events. We then identify, for each insurance phase, a group of practical research problems where security experts can help develop new data-driven methodologies and automated tools to replace the existing qualitative approaches.

1.5.2 Compiling Symbolic execution

A major impediment to practical symbolic execution is speed, especially when compared to near-native speed solutions like fuzz testing. In SymCC [PS20] We propose a compilation-based approach to symbolic execution that performs better than state-of-the-art implementations by orders of magnitude. SYMCC is an LLVM-based C and C++ compiler that builds concolic execution right into the binary. It can be used by software developers as a drop-in replacement for clang and clang++, and we show how to add support for other languages with little effort. In comparison with KLEE, SYMCC is faster by up to three orders of magnitude and an average factor of 12. It also outperforms QSYM, a system that recently showed great performance improvements over other implementations, by up to two orders of magnitude and an average factor of 10. Using it on real-



world software, we found that our approach consistently achieves higher coverage, and we discovered two vulnerabilities in the heavily tested OpenJPEG project, which have been confirmed by the project maintainers and assigned CVE identifiers.

SymCC received the "Distinguished paper award at Usenix Security 2020" and is available in open source project, and already has a number of active users and contributors: https://github.com/eurecom-s3/symcc

1.5.3 HTTS usage in mobile applications

Nowadays, virtually all mobile apps rely on communicating with a network backend. Given the sensitive nature of the data exchanged between apps and their backends, securing these network communications is of growing importance. In recent years, Google has developed a number of security mechanisms for Android apps, ranging from multiple KeyStores to the recent introduction of the new Network Security Policy, an XML-based configuration file that allows apps to define their network security posture. In a paper published at Usenix Security 2020 [PA20], we performed the first comprehensive study on these new network defense mechanisms. In particular, our work discussed the attacks they are defending from, and the relevant threat models. It also discusses the first large-scale analysis on this aspect.

During June and July 2019, we crawled 125,419 applications and we found how only 16,332 apps adopt this new security feature. We then focus on these apps, and we uncover how developers adopt weak and potentially vulnerable network security configurations. We note that, in November 2019, Google then made the default policy stricter, which would help the adoption. We thus opted to re-crawl the same dataset (from April to June 2020) and we repeated the experiments: while more apps do adopt this new security mechanism, a significant portion of them still do not take full advantage of it (e.g., by allowing usage of insecure protocols). We then set out to explore the root cause of these weaknesses (i.e., the why). Our analysis showed that app developers often copy-paste vulnerable policies from popular developer websites (e.g., StackOverflow). We also found that several popular ad libraries require apps to weaken their security policy, the key problem lying in the vast complexity of the ad ecosystem. As a last contribution, our work proposed a new extension of the Network Security Policy, so to allow app developers to embed problematic ad libraries without the need to weaken the security of their entire app.

1.5.4 MITM in zero conf protocols

In 2020, we have carried out an in depth evaluation of a number of so-called man in the middle attacks made possible thanks to the growing usage of a family of protocols named "zero conf protocols".

Zero conf protocols have been designed and implemented since 1999. They provide plug and play mechanisms to set up networks without having to configure DNS or DHCP servers. Almost every device (PCs, printers, scanners, etc.) nowadays "speaks" one of these protocols, sometimes without its owner being even aware of it. The booming IoT ecosystem, in particular, relies heavily on them. Unfortunately, these protocols offer a number of different ways to run, so called, man in the middle attacks (MITM). Some previous publications have mentioned and have taken advantage of one or another of these design flaws. In this work, we provide a deep dive into the various issues at hand and show the extent of the problem. We consider that the growing reliance of networks on these protocols represent an underestimated and ill covered threat. We have run a number of experiments (300) to test various implementations. We also have looked at ways to detect these attacks thanks to Zeek (aka Bro), a free open source network based intrusion detection system. We will make the attack code as well as the Zeek scripts available to the research community in a format that makes replication of our results possible by researchers while not easy to use by script kiddies. The results of this work are currently under submission at a conference devoted to offensive technologies.

1.6 Papers published in 2020 in the area of System and Software Security

[PS20] Sebastian Poeplau, Aurélien Francillon "Symbolic execution with SymCC: Don't interpret, compile!", Proceedings of the 29th USENIX Security Symposium - Distinguished Paper Award Winner

[PA20] Andrea Possemato, Yanick Fratantonio "Towards HTTPS Everywhere on Android: We Are Not There Yet", Proceedings of the 29th USENIX Security Symposium

[CA20] Giovanni Camurati, Aurélien Francillon, François-Xavier Standaert, "Understanding Screaming Channels: From a Detailed Analysis to Improved Attacks", IACR Transactions on Cryptographic Hardware and Embedded Systems (CHES 2020) - Google Bughunter Hall of Fame Honorable Mention

[CO20-1] Nassim Corteggiani, Aurélien Francillon "HardSnap: Leveraging hardware snapshotting for embedded systems security testing", Proceedings of the 50th IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2020)

[DA20] Savino Dambra, Leyla Bilge, Davide Balzarotti "SoK: Cyber Insurance - Technical Challenges and a System Security Roadmap", Proceedings of the IEEE Symposium on Security & Privacy

[CO20-3] Nassim Corteggiani, Giovanni Camurati, Marius Muench, Sebastian Poeplau, Aurelien Francillon, "SoC Security Evaluation: Reflections on Methodology and Tooling" IEEE Design & Test

[AG20] Hojjat Aghakhani, Fabio Gritti, Francesco Mecca, Martina Lindorfer, Stefano Ortolani, Davide Balzarotti, Giovanni Vigna, Christopher Kruegel "When Malware is Packin' Heat; Limits of Machine Learning Classifiers Based on Static Analysis Features", Proceedings of the Network and Distributed System Security (NDSS) Symposium

[MA20] Alessandro Mantovani, Simone Aonzo, Xabier-Ugarte Pedrero, Alessio Merlo, Davide Balzarotti "Prevalence and Impact of Low-Entropy Packing Schemes in the Malware Ecosystem", Proceedings of the Network and Distributed System Security (NDSS) Symposium

[CO20-2] Emanuele Cozzi, Pierre-Antoine Vervier, Matteo Dell'Amico, Yun Shen, Leyla Bilge, Davide Balzarotti "The Tangled Genealogy of IoT Malware", Annual Computer Security Applications Conference (ACSAC)

[SR20-1] Iskander Sanchez-Rola, Davide Balzarotti, Igor Santos "Cookies from the Past: Timing Server-Side Request Processing Code for History Sniffing", ACM Digital Threats: Research and Practice Journal (DTRAP)

[SR20-2] Iskander Sanchez-Rola, Davide Balzarotti, Christopher Kruegel, Giovanni Vigna, Igor Santos "Dirty Clicks: A Study of the Usability and Security Implications of Click-related Behaviors on the Web", Proceedings of The Web Conference 2020

[CH20-1] Elisa Chiapponi, Onu Catakoglu, Olivier Thonnard, Marc Dacier, “HoPLA: a Honeypot Platform to Lure Attackers”, Proc. of Computer & Electronics Security Applications Rendez-vous (C&ESAR) Deceptive Security, Rennes (France), 2020, available on line at www.cesar-conference.org

[CH20-2] Elisa Chiapponi, Marc Dacier, Massimiliano Todisco, Onu Catakoglu, Olivier Thonnard, “Botnets sizes: when maths meet myths”, Proc. of 1st workshop on Cyber Forensics and Threat Investigations Challenges in emerging infrastructures (CFTIC), December 14-17, 2020, Dubai (United Emirates).

2. CRYPTOGRAPHY AND PRIVACY ENHANCING TECHNOLOGIES

Faculty Members: Antonio Faonio, Melek Önen, Massimiliano Todisco

In 2020, the cryptography and privacy enhancing technologies group recruited two new assistant professors: Antonio Faonio and Massimiliano Todisco.

Research in our group has been carried out around the following themes:

- Privacy preserving Neural Networks;
- Crypto-biometrics;
- Leakage resilient cryptography;
- Machine learning for leak detection.

2.1 Privacy-preserving Neural Network Classification

Research Team
Beyza Bozdemir, Orhan Ermiş, Melek Önen
Research conducted in collaboration with
TU-Delft

As part of the PAPAYA H2020 project, we are investigating privacy preserving Neural Networks (NN) inference solutions. Performing NN predictions at untrusted servers poses a serious risk to the privacy of the data involved (see [CPMRÖPCM20] for some of the challenges raised by GDPR compliance requirements). Existing solutions propose to either encrypt data with Fully Homomorphic Encryption (FHE) enabling the execution of operations over encrypted data or use two-party computation that allow the computation of a function among two parties without revealing their inputs. Having studied existing solutions, in [TBÖ20], we propose a hybrid protocol named SwaNN which switches the computations between HE and 2PC (Two-party computation. We make use of partially homomorphic encryption (more specifically the Paillier encryption scheme) to perform linear operations over encrypted data. The non-linear operations are supported thanks to the 2PC. The combination of these two cryptographic tools helps maintain the accuracy of the prediction. SwaNN is designed to support two different settings: a client-server setting and a non-colluding two-server setting. While in the client-server setting the majority of the operations are delegated to the server and the client helps the server in intermediate steps, in the two-server setting, all operations are performed by the two servers with a balanced workload.

Furthermore, in [BEÖ20] we propose ProteiNN, a privacy-preserving neural network classification solution in a one-to-many scenario whereby one model provider outsources a machine learning model to the cloud server for its many different customers, and wishes to keep the model confidential while controlling its users. On the other hand, these customers take advantage of this machine learning model without revealing their sensitive inputs and the corresponding results. The solution employs homomorphic proxy re-encryption. A homomorphic proxy re-encryption (H-PRE) scheme is a public-key encryption scheme that, on the one hand, thanks to its homomorphic property, enables operations over encrypted data, and, on the other hand, allows a third-party proxy to transform ciphertexts encrypted with one public key into ciphertexts encrypted with another public key without leaking any information about the underlying plaintext. With such a technique, the third-party cloud server can easily serve multiple different queriers without having access to the in-puts, the results, and the model, and only the destined querier can decrypt the result. ProteiNN also makes use of a simple additive encryption scheme in order tolet the model provider keep control over the queriersand to prevent potential collusions among ProteiNN parties.

2.2 Crypto-biometrics

Research Team
Oubaïda Chouchane, Orhan Ermiş, Nicholas Evans, Alberto Ibarrondo, Madhu Kamble, Melek Önen, Jose Patino, Massimiliano Todisco
Research conducted in collaboration with
Idemia

Today, interactive services can easily collect, process, disclose and transmit our personal biometric data. Biometrics refers to the measurement and statistical analysis of individuals based on their intrinsic biological or behavioural characteristics. Biometrics - non-exhaustive list - include face, fingerprints, iris, retina, voice. By inferring identity from biological and behavioural characteristics, biometrics are an appealing, efficient, convenient and natural alternative to more traditional and now-outdated token-based authentication paradigms.

Despite the clear advantages and proliferation of biometrics technology, persisting concerns regarding intrusions into privacy have dented public confidence. The communication of biometric data across untrusted, public networks, and their processing on insecure, third- party cloud infrastructure, therefore, poses grave concerns, given the potential for our data to be exposed and intercepted by fraudsters or hackers.

Machine learning technologies are not traditionally designed with security and privacy into consideration. However, as innovation and legislation move forward, there is a requirement for this to alter. Privacy preservation techniques are urgently needed in order to fundamentally change the way biometric signals are processed in order to preserve privacy from the very moment of signal capture. To accomplish this objective, we must bring encryption to every stage of the processing tool-chain.

The aim is to research computationally feasible and reliable approaches through deep learning and cryptography that make use of biometric data for person recognition.

We continued in 2020 the collaborative work on cryptography and voice biometrics with the Audio Security and Privacy group through the ANR-DFG RESPECT and H2020 MCSA TReSPAsS-ETN projects. Both projects will deliver a new type of security protection through generalised presentation attack detection (PAD) technologies and privacy preservation through computationally feasible encryption solutions (see section 3.2.1).

We are investigating the potential of 3rd generation artificial neural networks (ANNs), known as spiking neural networks (SNNs), for privacy-preserving voice biometric recognition. These networks are known as spiking neural networks (SNNs), which more closely mimic the behaviour of biological neurons. Spiking neurons (SNs) are an intermediate model between biological neurons and artificial neurons, which are abstracted from biology by keeping only the most important principles, and thus, keeping a low computational complexity. Simpler computing capabilities - SNNs have the advantages to have only additive weight operations and sparse activations - will help inclusion of cryptographic primitives.

Secure multi-party computation (MPC) can be considered as suitable cryptographic tools for the design of novel primitives as they allow some functions to be evaluated by one or multiple parties without discovering any additional than what is needed. Unfortunately, some complex operations cannot be supported by existing practical cryptographic tools and therefore may need to be approximated with simpler operations. Some preliminary results showed the very early potential of using MPC in combination with SNN architectures for the detection of spoofing attacks for voice biometrics.

We are also considering privacy preserving inference with binarized neural networks (BNN). BNNs provide efficient implementations of Convolutional Neural Networks. This makes them particularly suitable to perform fast and memory-light inference of neural networks running on resource-constrained devices. Motivated by the growing interest in CNN-based biometric recognition on potentially insecure devices, or as part of strong multi-factor authentication for sensitive applications, we currently investigate a new method that performs secure inference of BNN relying on MPC. While existing solutions offer security in a semi-honest security setting for BNN, we propose to develop a scheme that aborts against one malicious adversary by leveraging replicated secret sharing.

2.3 Leakage Resilient Cryptography

Research Team
Antonio Faonio
Research conducted in collaboration with
Aarhus University, Imperial College London, National University of Singapore, University of Luxembourg and Sapienza University of Rome.

The security analysis of cryptographic primitives typically relies on the assumption that the underlying secrets (including, e.g., secret keys and internal randomness) are uniformly random to the eyes of the attacker. In reality, however, this assumption may simply be false due to the presence of so-called side-channel attacks (see for example section 1.2.3 of this very document), where an adversary can obtain partial information (also known as leakage) on the secret state of an implementation of a cryptographic scheme, by exploiting physical phenomena. Leakage-resilient cryptography aims at bridging this gap by allowing the adversary to launch leakage attacks in theoretical models too. The last decade has seen an impressive amount of work in this area, thanks to which we now dispose of a large number of leakage-resilient cryptographic primitives in different leakage model.

The most well-known cryptographic model, the so-called Bounded Leakage Model, assumes that the leakage is a "shirking" function of the secret material, namely a function that outputs L bits where L is smaller than the size of the cryptographic key (for example, L < 1024 bits for RSA). A considerable limitation of the Bounded Leakage Model is the fact that, in real-world side-channel attacks, the leakage obtained by the attacker is rarely bounded in length. For instance, the power trace on a physical implementation of AES typically consists of several Megabytes of information, which is much larger than the length of the secret key (256 bits). This motivates a more general notion of noisy leakage, where there is no upper bound on the length of the leakage but instead we assume it is somewhat noisy, in the sense that it does not reveal too much information about the secret material. It turns out that the level of noisiness of the leakage can be measured in several ways, each yielding a different leakage model. Indeed, we identified in the literature at least 6 different ways to model noisy leakage.

We investigated the relations between the bounded leakage model and the (many kind of) noisy-leakage models. Our result shows that many cryptographic primitives from previous works which have only been proved to be resilient against bounded leakage are also secure against noisy leakage, with only a small loss in the parameters. The core of our method is a master theorem that shows that noisy-leakage functions can be simulated by bounded-leakage functions. The corollaries of our master theorem are many: we can improve the state of art, upgrading from security against bounded-leakage attacks to security against noisy-leakage attacks, of forward-secure storage, leakage-resilient one-way functions, public-key encryption, randomness extractors, symmetric non-interactive key exchange, leakage-resilient secret sharing scheme and secure two-party computation. The results of this project have been recently published at EUROCRYPT'21 [BFORSSV21].

2.4 Machine Learning for Leak Detection

Research Team
Sofiane Lounici, Melek Önen
Research conducted in collaboration with
SAP

Public code platforms like GitHub are exposed to several different attacks, and in particular to the detection and exploitation of sensitive information (such as passwords or API keys). While both developers and companies are aware of this issue, there is no efficient open-source tool performing leak detection with a significant precision rate. Indeed, a common problem in leak detection is the amount of false positive data (i.e., noncritical data wrongly detected as a leak), leading to an important workload for developers manually reviewing them. In [LRMNTÖ20] together with colleagues at SAP, we propose an approach to detect data leaks in open-source projects with a low false positive rate thanks to the use of different machine learning techniques. First a regular expression scanner searches through the source code for potential leaks, looking for any correspondence with a set of programming patterns. Then, machine learning models filter the potential leaks by detecting false positive data, before a human reviewer can check the classified data manually to correct possible wrongly classified data. These machine learning models are using various techniques such as data augmentation, code stylometry and

reinforcement learning. We demonstrate that our tool, while producing a negligible false negative rate, decreases the false positive rate to, at most, 6% of the output data.

2.5 Other Topics

Research Team
Orhan Ermis, Melek Önen, Dimitrios Vasilopoulos
Research conducted in collaboration with
TNO, Bogazici University

2.5.1 Proofs of Reliability

In [VÖME20], we describe and evaluate two Proofs of Data Reliability (PDR) schemes that are cryptographic protocols that provide assurance to a user that a cloud storage system correctly stores her data and has provisioned sufficient redundancy to be able to guarantee reliable storage service. These schemes named as POROS and PORTOS on the one hand guarantee the retrieval of the outsourced data in their entirety through the use of proofs of data possession and on the other hand ensure the actual storage of redundancy. Moreover, POROS and PORTOS delegate the burden of generating the redundancy to the cloud as well. The security of both solutions is proved in the face of a rational adversary whereby the cheating cloud provider tries to gain storage savings without increasing its total operational cost.

2.5.2 Quantum-Safe Onion Routing

This year we performed a study on the use of post-quantum cryptographic primitives for the Tor network in order to make it safe in a quantum world [TRVHÖ20]. The Onion Router Tor aims to obfuscate the anonymity of its users when accessing or communicating over the Internet. In principle, when using Tor, the messages or website connection requests are sent through a network of relays and after multiple 'hops' reach their destination. So, if Alice wants to send Bob a message, but does not want an eavesdropper to know that she initiated the contact, Alice can use Tor. The cryptographic schemes used today and in Tor are based on hard mathematical assumptions e.g., Discrete Logarithm Problem and integer factorization \cite{KatzLindell}. These schemes are assumed to be secure against classical adversaries, as solving them with the currently known algorithms cost exponential time. However, with a quantum computer solving these problems becomes feasible. The transition from current cryptography to post-quantum cryptography needs to be started as soon as possible as quantum computers pose a threat to current and in particular public-key cryptographic algorithms. With this aim, the underlying keying material has first been analysed. We observe that breaking the security of the algorithms/protocols that use long- and medium-term keys (usually RSA keys) have the highest impact in security. Therefore, in our publication we investigate the cost of quantum-safe variants. These include key generation, key encapsulation and decapsulation. Six different post-quantum cryptographic algorithms that ensure level 1 NIST security are evaluated. We further target the Tor circuit creation operation and evaluate the overhead of the post-quantum variant. This comparative study is performed through a reference implementation based on SweetOnions that simulates Tor with slight simplifications. We show that a quantum-safe Tor circuit creation is possible and suggest two versions - one that can be used in a purely quantum-safe setting, and one that can be used in a hybrid setting.

2.5.3 Dynamic group key agreement with blockchain

Group key agreement protocols where participants can agree on a common secret key in an insecure channel have gained significant importance as a secure communication facilitator. Starting with Diffie-Hellman where two parties can agree on a secret key, several protocols have been developed which enable multiple parties to agree on a common key. However, such protocols were mostly designed for static groups, where the members of the group do not change until the end of communication session. Therefore, if the members in the group change, the entire protocol should be executed from the beginning for all participants in the group. On the other hand, some protocols provide additional functionalities to handle that re-execution overhead. Such protocols are called Dynamic Group Key Agreement (DGKA) protocols. Although DGKA protocols are more efficient compared to static ones, there exist some other several factors which affect their performance -- the first one is the way of broadcasting key agreement parameters and the second being the validation of participant identities via verification of received parameters. The blockchain technology is decoupled from cryptocurrencies and transformed into a wider domain with the concept of Distributed Ledger Technology (DLT) such as in Hyperledger Fabric (HF). Essentially, HF is a generic decentralized application development platform where transaction history is shared among peers as computation nodes in the network. DLT allows decentralization, greater transparency and easier auditability in a distributed setting. In [TEGA20], we propose a dynamic group key agreement protocol called B-GKAP which is an improved

version of a previously proposed protocol KAP-PBC and integrates the HF platform to improve key computation performance while keeping important security properties of known DGKA protocols. The main rationale for using the blockchain technology is to offload computational burden in a trustworthy and distributed manner for resource-constrained environments. The HF provides capabilities, e.g., Fabric Channels, of a permission-based blockchain platform to realize this extension in an efficient and secure way. The relevant benefits of our implemented approach are shown with the complexity analysis carried out for different aspects such as communication and computation in our experiments.

2.5.4 DDoS detection using time-series analysis for SDN

Available distributed denial of service (DDoS) attack detection mechanisms for software defined networks (SDNs) are generally based on statistical and machine learning approaches. Statistical methods employ a threshold to discriminate attacks from normal traffic. On the other hand, machine learning techniques use model training to generalize normal patterns. Compared to machine learning, statistical algorithms are simpler; however, defining optimal thresholds is challenging. To address this problem, in [FEA20], we propose a DDoS attack detection and defense scheme based on statistics and time-series analysis which is integrated into the SDN controller. First, we extract key features from the flow table of OpenFlow switches to analyze their time-series representation. Therefore, the controller monitors each switch separately and applies the proposed algorithm to detect anomalies due to the DDoS attack event in particular switches. Such anomaly can be detected using the number of unique source IP addresses (USIP) feature, since during the attack phase this value increases significantly compared to the normal traffic. In addition, when an attack occurs, it is possible to observe changes in the number of unique destination IP addresses (UDIP) feature; however, this change is not as significant as the change in the USIP feature. By considering the dramatic growth in the total number of packets (TPACK) in the flow table during the attack, the number of UDIP value substantially reduces when it is normalized by the TPACK values. Therefore, we employ normalized UDIP as the second feature to detect DDoS attacks. The upcoming value of the USIP time-series is estimated by using an auto regressive integrated moving average (ARIMA) model and the error of estimation is examined for chaotic behavior. Finally, according to the aforementioned method, a binary anomaly score is assigned to the traffic instance. Another binary anomaly score is obtained by using the dynamic threshold method based on the exponential filter and the NUDIP time-series. The product of two mentioned binary scores is used to identify DDoS attack in each switch. After detecting the attack samples, the scheme activates the countermeasure mechanism. The chaotic behavior and the dynamic threshold method resolve the problem of constant threshold in the previous works. Moreover, by monitoring each switch, the source of the attack can be identified and prompt countermeasure would be applied by modifying the packet forwarding policy at that switch.

2.6 Papers published in 2020 in the area of Cryptography and Privacy Enhancing Technologies

[BEÖ20] Beyza Bozdemir, Orhan Ermis, Melek Önen. “ProteiNN: Privacy-preserving one-to-many Neural Network classifications”, SECRYPT 2020, 17th International Joint Conference on Security and Cryptography, 8-10 July 2020, Lieusaint-Paris, France (Online conference)

[BFORSSV21] Gianluca Brian, Antonio Faonio, Maciej Obremski, Joao Ribeiro, Mark Simkin, Maciej Skorski, Daniele Venturi. "The Mother of All Leakages: How to Simulate NoisyLeakages via Bounded Leakage (Almost) for Free", accepted paper at EUROCRYPT'21, October 2021.

[CPMRÖPCM20] Renata M. de Carvalho M., Camillo Del Prete, Yod Samuel Martin, Rosa M. Araujo Rivero, Melek Önen, Francesco Paolo Schiavo, Ángel Cuevas Rumín, Haris Mouratidis, “Protecting citizens' personal data and privacy: a joint effort from GDPR EU cluster research projects”, SN Computer Science, Vol.1, Article number: 217 (2020).

[FEA20] Ramin Fadaei Fouladi, Orhan Ermiş, Emin Anarim, “A DDoS attack detection and defense scheme using time-series analysis for SDN", Journal of Information Security and Applications, 2020, vol. 54

[LMNRTÖ21] Sofiane Lounici, Carlo Maria Negra, Marco Rosa, Slim Trabelsi, Melek Önen. “Optimizing leak detection in open-source platforms with machine learning techniques”, ICISSP 2021, 7th International Conference on Information Systems Security and Privacy, 11-13 February 2021

[TBÖ20] Gamze Tillem, Beyza Bozdemir, Melek Önen. “SwANN: Switching among cryptographic tools for privacy-preserving Neural Network predictions”, SECRYPT 2020, 17th International Joint Conference on Security and Cryptography, 8-10 July 2020, Lieusaint-Paris, France

[TEGA20] Yaşar Berkay Taçyıldız, Orhan Ermiş, Gürkan Gür, Fatih Alagöz, “Dynamic group key agreement for resource-constrained devices using blockchains", 2nd International Workshop on Application Intelligence and Blockchain Security (AIBlock) 2020, in conjunction with ACNS 2020.

[TRVHÖ20] Zsolt Tujner, Thomas Rooijakkers, Maran van Heesch, Melek Önen. “QSOR: Quantum-safe onion routing”, SECRYPT 2020, 17th International Joint Conference on Security and Cryptography, 8-10 July 2020, Lieusaint-Paris, France (Online conference).

[VÖME20] Dimitrios Vasilopoulos, Melek Önen, Refik Molva, Kaoutar Elkhiyaoui. “Proofs of data reliability: Verification of reliable data storage with automatic maintenance”, Security and Privacy, December 2020, Wiley.

3. BIOMETRICS AND DIGITAL MEDIA

Faculty members: Jean-Luc Dugelay, Nicholas Evans, Melek Önen and Massimiliano Todisco

In this research axis, EURECOM works on signal processing (both image/video & speech/audio) for security applications, with a special emphasis on biometrics and, more particularly, on people recognition, privacy preservation, as well as DeepFake and presentation attack detection.

In 2020, the research in this area has focused on two topics related to Imaging Security, Safety and Privacy:

1. Facial Image processing & Biometrics
2. Image Forensics & Data Hiding

and on four topics related to audio and speech:

3. Voice biometrics
4. DeepFake and presentation attack detection
5. End-to-end and evolutive learning, explainability and interpretability
6. Privacy, anonymisation and pseudonymisation

3.1 Facial Image Processing & Biometrics

Research Team
Carmen Bisogni, Jean-Luc Dugelay, Chiara Galdi, Khawla Mallat, Luca Ulrich
Research conducted in collaboration with
Polito, University of Salerno, Cenatav

We are working on several topics related to Facial Image Processing, including:

- Adversarial attacks in face recognition, in collaboration with U. Salerno; visit of Carmen Bisogni during Spring 2020;
- Facial Soft Biometrics from videos, in collaboration with CENATAV Cuba (collaboration initialized within the context of the project IDENTITY); Visit of A. Morales and Fabiola Becerra during Oct.-Dec. 2019;
- Facial beauty, in collaboration with Polito.; visit of Luca Ulrich during Spring 2019;
- Thermal face image analysis within the context of the FUI project COOPOL;
- 3D Face recognition and multimodal biometrics within the context of the EU project PROTECT;
- Presentation Attack Detection in collaboration with the company SURYS;
- This year, we started some R&D activities on facial image analysis in Healthcare.

We are investigating several new cameras in facial image processing and video surveillance, in particular FLIR dual cameras for biometrics (project FUI COOPOL) and video surveillance (project ANR OKLOS), and more recently event-based cameras. Thermal imaging has substantially evolved, during the recent years, to be established as a complement, or even occasionally as an alternative to conventional visible light imaging, particularly for face analysis applications. Facial landmark detection is a crucial prerequisite for facial image processing. Given the upswing of deep learning based approaches, the performance of facial landmark detection has been significantly improved. However, this uprise is merely limited to visible spectrum based face analysis tasks, as there are only few research works on facial landmark detection in thermal spectrum. This limitation is mainly due to the lack of available thermal face databases provided with full facial landmark annotations. In this paper, we propose to tackle this data shortage by converting existing face databases, designed for facial landmark detection tasks, from visible to thermal spectrum that will share the same provided facial landmark annotations. Using the synthesized thermal databases along with the facial landmark annotations, two different models are trained using active appearance models and deep alignment networks. Evaluating the models trained on synthesized thermal data on real thermal data, we obtained facial landmark detection accuracy of 94.59% when tested on low quality thermal data and 95.63% when tested on high quality thermal data with a detection threshold of 0.15×IOD.

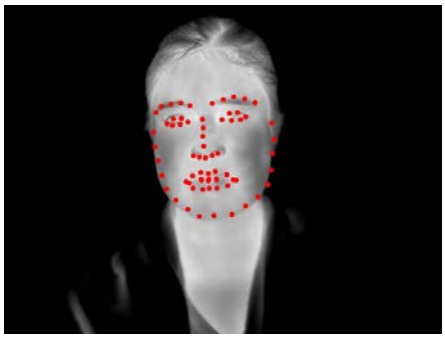


Image extracted from [MAL 20]

Cross-spectrum face recognition, e.g. visible to thermal matching, remains a challenging task due to the large variation originating from different domains. In article [DON 20], we proposed a subspace projection hashing (SPH) to enable the cross-spectrum face recognition task. The intrinsic idea behind SPH is to project the features from different domains onto a common subspace, where matching the faces from different domains can be accomplished. Notably, we proposed a new loss function that can (i) preserve both inter-domain and intra-domain similarity; (ii) regularize a scaled-up pairwise distance between hashed codes, to optimize projection matrix. Three datasets, Wiki, EURECOM VIS-TH paired face and TDFace are adopted to evaluate the proposed SPH. The experimental results indicate that the proposed SPH outperforms the original linear subspace ranking hashing (LSRH) in the benchmark dataset (Wiki) and demonstrates a reasonably good performance for visible-thermal, visible-near infrared face recognition, therefore suggests the feasibility and effectiveness of the proposed SPH.

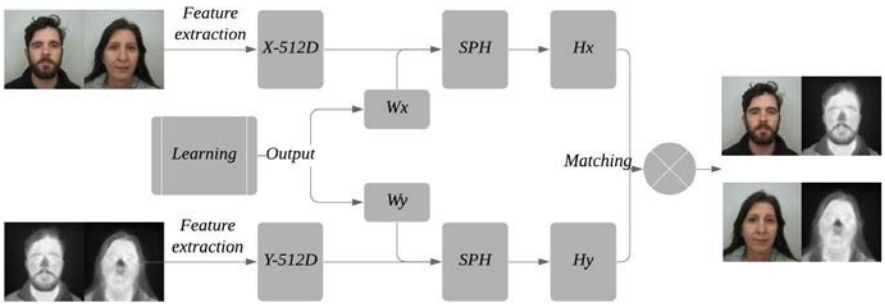


Image extracted from [DON 20]

Most existing works regarding facial demographic estimation are focused on still image datasets, although nowadays the need to analyze video content in real applications is increasing. We propose in [BEC 20] to tackle gender, age and ethnicity estimation in the context of video scenarios. Our main contribution is to use an attribute-specific quality assessment procedure to select most relevant frames from a video sequence for each of the three demographic modalities. Selected frames are classified with fine-tuned MobileNet models and a final video prediction is obtained with a majority voting strategy. Our validation on three different datasets and our comparison with state-of-the-art models, show the effectiveness of the proposed demographic classifiers and the quality pipeline, which allows to reduce both: the number of frames to be classified and the processing time in practical applications; and improves the soft biometrics prediction accuracy.

Common sense usually considers the assessment of female human attractiveness to be subjective. Nevertheless, in the past decades, several studies and experiments showed that an objective component in beauty assessment exists and can be strictly related, even if it does not match, with proportions of features. Proportions can be studied through analysis of the face, which relies on landmarks, ie, specific points on the facial surface, which are shared by everyone, and measurements between them. In this work, several measures have been gathered from studies in the literature considering datasets of beautiful women to build a set of measures that can be defined as suggestive of female attractiveness. The resulting set consists of 29 measures applied to a public dataset, the Bosphorus database, whose faces have been both analyzed by the developed methodology based on the expanded set of measures and judged by human observers. Results show that the set of chosen measures is significant in terms of attractiveness evaluation, confirming the key role of proportions in beauty

assessment; furthermore, the sorting of identified measures has been performed to identify the most significant canons involved in the evaluation.

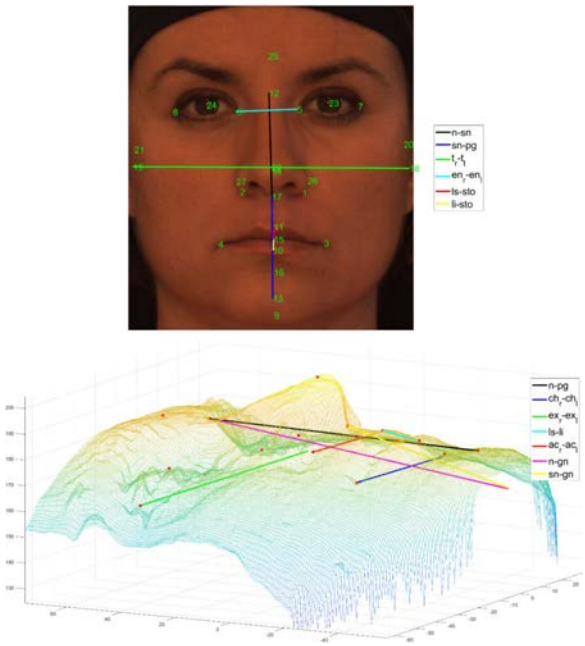


Image extracted from [ULR 20]

Pervasive and useR fOcused biomeTrics bordEr projeCT (PROTECT) is an EU project funded by the Horizon 2020 research and Innovation Programme. The main aim of PROTECT was to build an advanced biometric-based person identification system that works robustly across a range of border crossing types and that has strong user-centric features. This work took place at the end of the project and includes the case study of the multibiometric verification system developed within PROTECT. The system has been developed to be suitable for different borders such as air, sea, and land borders. The system covers two use cases: the walkthrough scenario, in which the traveller is on foot; the drive-through scenario, in which the traveller is in a vehicle. Each deployment includes a different set of biometric traits and this study illustrates how to evaluate such multibiometric systems in accordance with international standards and, in particular, how to overcome practical problems that may be encountered when dealing with multibiometric evaluation, such as different score distributions and missing scores.

Image extracted from [GAL 20]

3.2 Image Forensics & Data Hiding

Research Team
Alexandre Berthet, Jean-Luc Dugelay, Chiara Galdi, Anis Trabelsi
Research conducted in collaboration with
SURYS, University of Monash

We are working on several topics related to Image Forensics and Data Hiding:

- Detection of malicious modifications on images within the context of the ANR project DEFACTO; As in many other domains, the last trend is to develop AI-based methods;
- Detection of Deepfakes, in collaboration with the company SURYS;
- Data hiding in images, in collaboration with U. Monash;
- Visual encryption, in collaboration with U. of Tech. Guangzhou.

Access to technologies like mobile phones contributes to the significant increase in the volume of digital visual data (images and videos). In addition, photo editing software is becoming increasingly powerful and easy to use. In some cases, these tools can be utilized to produce forgeries with the objective to change the semantic meaning of a photo or a video (e.g. fake news). Digital image forensics (DIF) includes two main objectives: the detection (and localization) of forgery and the identification of the origin of the acquisition (i.e. sensor identification). Since 2005, many classical methods for DIF have been designed, implemented and tested on several databases. Meantime, innovative approaches based on deep learning have emerged in other fields and have surpassed traditional techniques. In the context of DIF, deep learning methods mainly use convolutional neural networks (CNN) associated with significant preprocessing modules. This is an active domain and two possible ways to operate preprocessing have been studied: prior to the network or incorporated into it. None of the various studies on the digital image forensics provide a comprehensive overview of the preprocessing techniques used with deep learning methods. Therefore, the core objective of the article [BER 20] is to review the preprocessing modules associated with CNN models.

A technique is put forward to hide data into an animated GIF by exploiting the transparent pixels [WON 20]. Specifically, a new frame is crafted based on the data to be embedded. The newly crafted frame is inserted between 2 existing frames, and the delay time of the affected frames are adjusted accordingly to achieve complete imperceptibility. To the best of our knowledge, this is the first attempt to hide data into an animated GIF by exploiting the transparent pixel. Regardless of the characteristics of the animated GIF image, the proposed method can completely preserve the quality of the image before and after hiding data. The hiding capacity achieved by the proposed method is scalable, where more information can be embedded by introducing more frames into the animated GIF. While file size expansion is inevitable, reverse zero run length is adopted to suppress the expansion. The proposed method is reversible, i.e., the original image can be recovered.

We proposed a novel image transformation scheme to protect the visual information. By mimicking an arbitrarily chosen reference image, a secret image is visually changed and can be exactly recovered from the transformed image when needed. Unlike the blockwise visual encryption methods, the proposed transformation scheme modifies the secret image by bit plane replacement and reordering so that no block effect is introduced. In particular, one or more bit planes are hidden into the other bit planes so that the most significant one(s) can be vacated. Besides replacing the vacated bit plane(s) and reordering the others according to the reference image, histogram modification may be performed to conceal the secret content if needed. To exactly recover the secret image, the required information is recorded and reversibly hidden into the transformed image by adopting a reversible data hiding algorithm. The experimental results on three image sets show that their content can be semantically changed to prevent leakage of visual information. Moreover, the applicability and efficiency of the proposed scheme have been validated by comparing the existing visual encryption schemes.

Image extracted from [WU 20]

3.3 Voice biometrics

Research team
Oubaïda Chouchane, Nicholas Evans, Wanying Ge, Madhu Kamble, Andreas Nautsch, Melek Önen, Jose Patino, Hemlata Tak, Massimiliano Todisco

The department has continued collaborative work in voice biometrics during 2020 through the ANR-DFG RESPECT and H2020 TReSPAsS-ETN projects. The joint-national RESPECT project (http://www.respect-project.eu) encompasses core research in multi-biometric audio-visual recognition technologies, i.e. research to improve recognition reliability, as well as security and privacy. These topics are covered in additional detail in Sections 2.2 and 3.4. RESPECT is coordinated by Massimiliano Todisco with collaboration from Hochschule Darmstadt and Hochschule Ansbach in Germany and INRIA in France. The 48-month, 3.7m€ TReSPAsS-ETN project (http://www.trespass-etn.eu) also coordinated by EURECOM and led by Massimiliano Todisco, is innovating next generation secure and privacy-preserving authentication technologies in consideration of the legal and ethical issues surrounding the use of biometrics for person recognition. Like all ITN projects, TReSPAsS-ETN is training a new cohort of early-stage researchers (ESRs) through a blend of traditional research training, workshops, summer schools, industry networking and start-up creation events. Involving 14 leading European research institutes and partner organisations, and having started in January 2020, TReSPAsS-ETN has now recruited 14 ESRs, two of which are hosted at EURECOM.

In 2020 the department embarked upon a new I4U (Institutes and Universities across 4 continents) joint submission to the 2020/21 speaker recognition evaluation (SRE), the premier evaluation campaign in speaker recognition administered by the National institute of Standards and Technology in the US. I4U facilitates the collaboration by a large number of research groups, each of whom produce scores that are later fused before submission to NIST. Our system uses state-of-the-art deep speaker embeddings and the use of novel paradigms for channel compensation and domain adaptation. A joint paper covering the I4U submission will be submitted for publication in 2021.

3.4 DeepFake detection / anti-spoofing

Research team
Nicholas Evans, Wanying Ge, Madhu Kamble, Andreas Nautsch, Jose Patino, Hemlata Tak, Massimiliano Todisco
Research conducted in collaboration with
University of Eastern Finland, the National Institute of Informatics (Japan), I2R in (Singapore), Inria (France), Hochschule Darmstadt and Hochschule Ansbach (Germany)

All biometric systems are vulnerable to presentation or spoofing attacks, namely the circumvention, manipulation or deception of biometrics technology by fraudsters who impersonate the biometric characteristics of another person in order to gain illegitimate access to sensitive or protected facilities or resources. The development of presentation attack detection (PAD) systems, which aim to protect biometrics technology from attack, has been a core focus for over a decade. We are co-founders of the ASVspoof initiative, the Automatic Speaker Verification Spoofing and Countermeasures challenge series which aims to promote the consideration of spoofing and to foster progress in PAD for voice biometrics. During 2020, our work in developing constant Q cepstral coefficients (CQCCs) for anti-spoofing received the ISCA Award for the Best Paper published in Computer Speech and Language between 2015 and 2019.

In 2020 we focused on identifying the artefacts extracted by CQCCs that account for its performance advantage over competing representations [TAK20b]. Our study focused on examining the attention given to different sub-band components by each front-end and why the CQCC front-end performs well in detecting some spoofing attacks, but poorly in detecting others. We found that different spoofing attacks exhibit artefacts at different frequencies where the CQCC front-end has higher resolution than other front-ends. The same findings also suggest that conventional cepstral analysis smooths information across the full band and dilutes localised information. Our follow-up work published in [TAK20c] demonstrated the benefit of fusing relatively traditional subband classifiers, each tuned to the detection of different spoofing attacks that exhibit artefacts in different, specific sub-bands and non-linear approaches to score-level fusion. Even without the use of complex deep learning, our approach outperforms all but two of the 48 competing systems submitted to the ASVspoof 2019 logical access challenge.

In collaboration with Hochschule Darmstadt and Hochschule Ansbach, Germany, we also explored the application of image-based anti-spoofing solutions to speech data [GON20]. This is achieved by treating speech data as a visual spectro-temporal decomposition. Motivated by the need to improve countermeasure generalisability, namely countermeasure robustness to new forms of spoofing attack with previously unseen characteristics, the work explored the application of texture descriptors to the analysis of speech spectrogram images. We proposed in [GON20] a common fisher vector feature space, often used in image classification tasks, as a means of extracting spoofing artefacts from speech signals. Experiments performed on the ASVspoof 2019 database confirm its use for the detection of both known as well as unknown spoofing attacks.

In collaboration with others from the University of Eastern Finland, the National Institute of Informatics in Japan, I2R in Singapore and Inria in France, EURECOM has continued its role as co-organiser of the bi-annual ASVspoof challenge series. Following the most successful edition yet that was organised in 2019, attracting the participation of 63 teams from research institutions and industry worldwide [WAN20], preparation began in earnest during 2020 for the 4th edition scheduled to run during 2021. Technical goals for the 2021 edition include a shift towards non-ideal conditions, e.g. logical access attacks mounted over variable telephony channels and physical access attacks launched in real, rather than simulated acoustic environments. Outcomes during 2020 include the publication of a special issue on Advances in Automatic Speaker Verification Anti-spoofing in Elsevier Computer Speech and Language in 2020 co-edited by the ASVspoof organising committee and joint work on the metrics used for evaluation.

In contrast to common practice throughout the broader biometrics community which focuses on the assessment of PAD solutions in isolation from recognition systems, since 2018 we have advocated the tandem/joint assessment of independently developed PAD and recognition systems. This seems critical since PAD will always impact upon the performance of a recognition system and it is an estimate of their performance when combined that is of the greatest interest. Since both PAD and recognition systems are deployed together in a holistic system, such as in banking applications to detect fraud, a risk-based evaluation framework is also essential. We compared two approaches to risk assessment which differ in terms of whether or not the threshold of the voice biometrics system is fixed or can be changed such that the two systems can be jointly optimised. In a risk-based assessment methodology, different decision policies for which risk should be minimised predetermine the default risk of a particular policy. These results show that some PAD solutions only serve to increase risk, whereas others are effective in improving security. This work was published in [KIN20].

3.5 End-to-end and evolutive learning, explainability and interpretability

Research team
Nicholas Evans, Wanying Ge, Andreas Nautsch, Jose Patino, Hemlata Tak, Massimiliano Todisco
Research conducted in collaboration with
Laboratoire d'Informatique de l'Université du Mans (France)

End-to-end learning marks the shift away from handcrafted representations and manually or empirically optimised systems towards learning strategies where (almost) every parameter and component of a classification system is learned automatically and jointly. Neural architecture search and evolutive learning take this idea one step further where even the architecture of a neural network is itself learned automatically. While the latter are currently more blue sky research, end-to-end learning techniques are now pushing the state of the art in an array of different speech processing tasks, including voice biometrics and anti-spoofing. In collaboration with colleagues at the Laboratoire d'Informatique de l'Université du Mans, EURECOM is exploring these strategies within the scope of the ANR ExTENSoR project which started in March 2020. Through ExTENSoR we have developed a deep neural network based solution to anti-spoofing that ingests raw speech signals, thereby avoiding the use of hand-crafted feature extraction [TAK20a]. The solution is based upon the recently proposed RawNet2 architecture which exploits residual blocks and gated recurrent units for the extraction of embeddings for automatic speaker verification. It has potential to learn cues that are not detectable using more traditional countermeasure solutions. The work demonstrates the potential of end-to-end approaches that utilise automatic feature learning to improve performance, both for known spoofing attack types as well as for worst-case, more practical scenarios involving previous unseen spoofing attacks for which representative training data is unavailable. When fused with more traditional classifiers, we obtained the second-best results reported for the ASVspoof 2019 logical access condition.

In parallel and by building upon previous work conducted in collaboration with industry, we are also investigating the potential of evolutive topologies. This work, which started in the fourth quarter of 2020, is aligned to both the ANR ExTENSorR and

H2020 Marie Curie TReSPAsS-ETN project and is exploring NAS based solutions to anti-spoofing. The searching, learning and evolution of network topologies learned automatically has the potential to capture information contained within speech signals not captured by more traditional, hand-crafted neural network solutions. We have already demonstrated the competitiveness of these less complex, hierarchical as well as non-hierarchical, inherently recursive network architecture solutions and plans to publish this work are already afoot.

3.6 Privacy, anonymisation and pseudonymisation

Research team
Nicholas Evans, Andreas Nautsch, Jose Patino, Massimiliano Todisco
Research conducted in collaboration with
National Institute of Informatics (Japan), Laboratoire Informatique d'Avignon, Inria (France)

Stemming from the implementation of the European general data protection regulation (GDPR), efforts to research and develop privacy preservation solutions for speech data have increased in recent years. Speech data is inherently personal data since it encodes potentially sensitive information such as a speaker's age, gender, ethnic origin, geographical background, health or emotional state, political orientation, or religious beliefs, etc. Automatic speaker recognition systems can also infer a speaker's identity. The group's interest in privacy research started in 2018 and has explored both crypto-biometrics solutions as well as anonymisation and pseudonymisation. Whereas encryption-based systems aim to protect privacy-sensitive data from being intercepted, anonymisation and pseudonymisation aim to sanitize or suppress privacy-sensitive data from a speech signal so no privacy is leaked even if the speech signal is intercepted in clear form. Details of the department's work in crypto-biometrics is covered in Section 2.2.

Our work in anonymisation is supported by the ANR Harpocrates and VoicePersonae projects. The ANR VoicePersonae project started in 2018 and is led by the National Institute of Informatics (NII) in Japan, with the participation of EURECOM and the Laboratoire Informatique d'Avignon (LIA) in France. VoicePersonae has a focus on advanced speaker identity modeling and on improving security and privacy in speech processing through research in voice biometrics. In collaboration with colleagues at Laboratoire Informatique d'Avignon and INRIA, we formulated the ANR Harpocrates project to launch a competitive challenge to help foster progress in the research and development of anonymisation solutions for speech data. With our international colleagues, we formed the VoicePrivacy initiative and associated challenge series [TOM20b] to promote the development of anonymisation solutions for speech technology. VoicePrivacy gathered a new community to define the task, the evaluation methodology, and benchmarking solutions through a series of challenges. The first challenge was held during 2020 [TOM20a] and culminated in a workshop held in conjunction with Odyssey 2020, the Speaker and Language Recognition Workshop. It attracted the participation of over 80 delegates from both academia and industry worldwide. Results show that anonymisation solutions can suppress personally identifiable information contained within speech signals, but also that anonymisation is far from being a solved problem; automatic speaker recognition systems trained with anonymised speech can still infer speaker identity, meaning that some speaker discriminative information remains in anonymised speech utterances. More detailed findings from the challenge are in preparation for submission to Elsevier Computer Speech and Language through a special issue on privacy preservation co-edited with Harpocrates, VoicePersonae and VoicePrivacy colleagues.

We also developed our own solution to anonymisation [PAT20] which was adopted as a contrastive baseline for the VoicePrivacy challenge. Our approach was motivated by the need: (i) for a computationally efficient approach with potential to be implemented on a battery powered mobile device; (ii) to provide a baseline to help demonstrate the benefit of more complex, sophisticated solutions; (iii) to lower the cost of entry to anonymisation research for potential challenge participants and provide additional inspiration. It is based upon traditional signal processing techniques and the McAdams factor which acts to shift the pole positions and hence the spectral envelope of a speech utterance in order to deceive an automatic speaker recognition system while protecting quality and naturalness. As an added benefit, it is also training data independent. Our solution was released as open source and has already been adopted by VoicePrivacy participants among other researchers.

We have also invested in the design of metrics for the assessment of privacy preservation and anonymisation solutions. Work in collaboration with Laboratoire Informatique d'Avignon (LIA) [NOE20], developed new metrics and visualisations to help assess de-identification performance and the distinguishability of anonymised voices. The zero-evidence biometric

recognition assessment (ZEBRA) framework published in [NAU20a] introduces new metrics based upon forensic science methods. We introduced an approach to gauge the level of expected privacy disclosure in protected speech data and showed that even privacy safeguards that are strong in an average sense may leave some individuals with relatively weak protection. We also proposed categorical tags to help provide more intuitive insights into privacy preservation performance for the less technically minded. Together with an overview of the department's broader activities in privacy preservation, this work was also presented at the ITG Workshop on Speech Assistants [NAU20b].

The department has also invested heavily in community actions linked to privacy preservation for speech technologies. Senior Research Fellow Andreas Nautch, recipient of the 2020 European Biometric Max Snijder Award, was re-elected to serve a second term as co-chair (with Tom Bäckström from Aalto) of the newly formed ISCA Special Interest Group (SIG) in Security and Privacy in Speech Communication (SPSC). The department hosted a SIG-supported Speech meets legal exerts event at EURECOM which was attended by Thomas Zerdick from the European Data Protection Supervisor and is organising (with Stephan Sigg, also from Aalto) a Lorentz Center workshop on the topic of Speech as Personally Identifiable Information: Ethics & Usability in vis-à-vis with Security & Privacy.

3.7 Papers published in 2020 in the area of Biometrics and Digital Media

[MAL 20] Mallat, Khawla; Dugelay, Jean-Luc, "Facial landmark detection on thermal data via fully annotated visible-to-thermal data synthesis", IJCB 2020, International Joint Conference on Biometrics, September 28 - October 1, 2020 (Virtual Conference)

[DON 20] Wang, H.; Dong, X.; Zhe, J.; Dugelay, J.-L.; Tistarelli, M., "Cross-spectrum face recognition using subspace projection hashing", ICPR 2020, 25th International Conference on Pattern Recognition, 10-15 January 2021, Milan, Italy (Virtual Conference)

[BEC 20] Becerra-Riera, F.; Morales-González, A.; Mendez-Vazquez, H.; Dugelay, J.-L., "Attribute-based quality assessment for demographic estimation in face videos", ICPR 2020, 25th International Conference on Pattern Recognition, 10-15 January 2021, Milan, Italy (Virtual Conference)

[ULR 20] Ulrich, L.; Dugelay, J.-L.; Vezzetti, E.; Moos, S.; Marcolin, F., "Perspective morphometric criteria for facial beauty and proportion assessment", Applied Sciences 2020, 10, 8; doi:10.3390/app10010008

[GAL 20] Galdi, C.; et al. "PROTECT: Pervasive and user focused biometrics border project – a case study", IET Biometrics, Vol.9, N°6, November 2020

[BEC 20] Becerra-Riera, F.; Morales-González, A.; Mendez-Vazquez, H.; Dugelay, J.-L., "Attribute-based quality assessment for demographic estimation in face videos", ICPR 2020, 25th International Conference on Pattern Recognition, 10-15 January 2021, Milan, Italy (Virtual Conference)

[BER 20] Berthet, A.; Dugelay, J.-L., "A review of data preprocessing modules in digital image forensics methods using deep learning" VCIP 2020, IEEE International Conference on Visual Communications and Image Processing, 1-4 December 2020, Macau, China (Virtual Conference)

[WON 20] Wong, K.; Nazeeb; Dugelay, J.-L., "Complete quality preserving data hiding in animated GIF with reversibility and scalable capacity functionalities", IWDW, 19th International Workshop on Digital-forensics and Watermarking, 25-27 Novembre 2020, Melbourne, Australia (Virtual Conference)

[WU 20] Wu, H.-T.; Jia, R.; Dugelay, J.-L.; He, J., "Reversible image visual transformation for privacy and content protection", Multimedia Tools and Applications, 14 October 2020
Facial landmark detection on thermal data via fully annotated visible-to-thermal data synthesis
IJCB 2020, International Joint Conference on Biometrics, September 28–October 1, 2020 (Virtual Conference)

[GON20] Gonzalez-Soler, L., Patino, J., Gomez-Barrero, M., Todisco, M., Busch, C. and Evans, N., "Texture-based presentation attack detection for automatic speaker verification", in Proc. WIFS, IEEE International Workshop on Information Forensics and Security, 6-11 December 2020

[TAK20a] Tak, H., Patino, J., Todisco, M., Nautsch, A., Evans, N., Larcher, Anthony, “End-to-end anti-spoofing with RawNet2”, arXiv:2011.01108, 2020, also submitted to ICASSP 2021

[PAT20] Patino, J., Tomashenko, N., Todisco, M., Nautsch, A., Evans, N., “Speaker anonymisation using the McAdams coefficient”, arXiv:2011.01130, 2020, also submitted to ICASSP 2021

[TAK20b] Tak, H., Patino, J., Nautsch, A., Evans, N., Todisco, M., “An explainability study of the constant Q cepstral coefficient spoofing countermeasure for automatic speaker verification”, ODYSSEY 2020, Speaker and Language Recognition Workshop, 2020

[TOM20a] Tomashenko, N., Lal Srivastava, B. M.; Wang, X.; Vincent, E., Nautsch, A., Yamagishi, J., Evans, N., Patino, J., Bonastre, J.-F; Noé, P.-G.; Todisco, M., “Introducing the VoicePrivacy initiative”, in Proc. INTERSPEECH, Annual Conference of the International Speech Communication Association, 2020

[NAU20a] Nautsch, A., Patino, J., Tomashenko, N., Yamagishi, J., Noé, Paul-Gauthier; Bonastre, Jean-François; Todisco, M., Evans, N., “The privacy ZEBRA: Zero Evidence Biometric Recognition Assessment”, in Proc. INTERSPEECH, Annual Conference of the International Speech Communication Association, 2020

[TAK20c] Tak, H., Patino, J., Nautsch, A., Evans, N., Todisco, M., “Spoofing attack detection using the non-linear fusion of sub-band classifiers”, in Proc. INTERSPEECH, Annual Conference of the International Speech Communication Association, 2020

[NOE20] Noé, P.-G.; Bonastre, J.-F.; Matrouf, D.; Tomashenko, N., Nautsch, A., Evans, N., “Speech pseudonymisation assessment using voice similarity matrices”, in Proc. INTERSPEECH, Annual Conference of the International Speech Communication Association, 2020

[KIN20] Kinnunen, T.; Delgado, H.; Evans, N., Lee, K.-A.; Vestman, V.; Nautsch, A., Todisco, M., et al., “Tandem assessment of spoofing countermeasures and automatic speaker verification: fundamentals”, IEEE/ACM Transactions on Audio, Speech and Language Processing, 2020

[WAN20] Wang, Xin; Yamagishi, J., Todisco, M., Delgado, H.; Nautsch, A., Evans, N., Sahidullah, M.; Vestman, V.; et al., “ASVspoof 2020: A large-scale public database of synthesized, converted and replayed speech”, Elsevier Computer Speech and Language, Vol. 64, November 2020

[BAC20] Bachhav, P., Todisco, M., Evans, N., “Artificial bandwidth extension using conditional variational auto-encoders and adversarial learning”, in Proc. ICASSP, International Conference on Acoustics, Speech, and Signal Processing, 2020

[NAU20b] Nautsch, A., Todisco, M., Patino, J., Evans, N., “Audio security and privacy”, ITG Workshop, 2020

[TOM20b] Tomashenko, N., Srivastava, B. M. L.; Wang, X.; Vincent, E., Nautsch, A., Yamagishi, J., Evans, N., Patino, J., Bonastre, J.-F., Noé, P.-G.; Todisco, M., “The VoicePrivacy 2020 challenge evaluation plan”, VOICEPRIVACY 2020

4. STAFF 2020

Professors (6)	Visiting Scientists (1)	Phd Students (30)
Davide Balzarotti Marc Dacier Jean-Luc Dugelay Nicholas Evans Refik Molva (emeritus) Aurélien Francillon	Carmen Bisogni (Univ. of Salerno)	Pierre Ayoub Samuel Aubertin Alexandre Berthet Beyza Bozdemir Giovanni Camurati Elisa Chiapponi Oubaïda Chouchane Emmanuele Cozzi Savino Dambra Fioraldi Andrea Moritz Eckert Wanying Ge Alberto Ibarrondo Mohamed Jamel Mikhail Klementev Mirabet Nelida Sofiane Lounici Mohamad Mansouri Alessandro Mantovani Dario Nisi Anis Trabelsi Feras Al Kassar Andrea Oliveri Paul Olivier Sebastian Pöplau Andrea Possemato Anazim Rahman Hemlata Tak Florent Moriconi Kamaruzaman Maskat
Assistant Professors (4)	Postdocs (3)	
Antonio Faonio Yanick Fratantonio Melek Önen Massimiliano Todisco	Simone Aonzo Ayse Unsal Jose Patino	
Research Associates (7)	Interns (12)	
Nassim Corteggiani Matteo dell'Amico Orhan Ermis Chiara Galdi Kamble Madhu Khawla Mallat Andreas Nautsch	Gilang Hamidy Arnaud Barral Maxime Bouthors Emna Abid Mehdi Kachouri Zied Becha Dhia Farrah Mohamed Njeh Ala Tarhouni Jiahao Hu Lorenzo Maffia Antonino Vitale	

DATA SCIENCE DEPARTMENT

Introduction

The objectives of the Data Science Department are defined through an interdisciplinary approach to research, merging contributions from computer science, web science, machine learning and statistics, artificial intelligence, and addressing numerous applied problems. In the past year, the research scope of the department consolidated along the several lines of work developed in the last 5 years. On the fundamental side, the theme of computational statistics for computer simulation has been established as a natural complement to the body of work on computational methods for Bayesian statistics. In addition, the genomics research line has been flourishing, merit of the ambitious OligoArchive FET project: this constitutes the basis of an emerging research activity in computational biostatistics and genomics. The key application domain related to medicine and health has been growing steadily as well, with both fundamental and practical research in robust methods for machine learning.

In the past year, our the department has promoted the application of fundamental research (machine learning and AI, knowledge and data management) in the following verticals:

- Digital preservation
- Medical imaging and e-health
- Life and environmental sciences
- Genomics and biostatistics
- Creative industries, cultural heritage and tourism
- Merchandising and retail
- Autonomous transportation systems

The Data Science department has reached a steady state regime, which ranges from research to teaching activities. The past year has been particularly favourable in terms of international visibility in key application domains, such as digital preservation and sustainable AI.

Summary of Achievements

Overall, the past year has seen new EU projects, new industrial research projects with EURECOM members and with new industrial actors, and existing projects being carried out, which is a striking evidence of the visibility and outreach achieved by all department members, thanks to their research activities and scientific achievements.

- Publicly funded projects
  - ANR JCJC “ECO-ML”
  - ANR JCJC “InfClean”
  - ANR ANTRACT
  - ANR ASRAEL
  - CHISTERA CIMPLE
  - H2020 Odeuropa
  - H2020 FET “OligoArchive”
  - H2020 MeMAD
  - H2020 SILKNOW
- Industrial projects
  - CIFRE Huawei
  - 2x CIFRE SAP
  - KPMG contract “KaBAN”
  - IMT Futur & Ruptures “AutoClean”
  - 2x CIFRE Orange
  - Direct contract, Orange
  - CIFRE Orkis
  - 2x CIFRE Renault Software Labs
  - Direct contract, Amadeus
  - CIFRE Amadeus

- Direct contract, LINKS Foundation
- Direct contract, NEC Corporation
- Direct contract, Google.org
- Direct contract, International Fact Checking Network

- PhD Defenses:
  - Enrico Palumbo. Knowledge graph embeddings for recommender systems, 04-2020 (supervised by Raphael Troncy)
  - Quan-Dong Vu. Models and solutions of strategic resource allocation problems: Approximate equilibrium and online learning in Blotto games, 06-2020 (supervised by Patrick Loiseau)
  - Gia-Lac Tran. Advances of deep Gaussian processes: Calibration and sparsification, 12-2020 (supervised by Maurizio Filippone)

- Awards:
  - Most cited paper 2018-2021 "Pattern Recognition", for the work "A comparative evaluation of outlier detection algorithms: Experiments and analyses"
  - 2 x Distinguished Reviewer ICML 2020
  - Distinguished Reviewer at VLDB 2020
  - Best demo award at DBA 2020
  - 1st ranked system in the TRECVID Video Summarization 2020 Challenge
  - 2nd ranked system (out of 8) in the MediaEval Memorability 2020 Challenge
  - 3rd ranked system (out of 28) in the ISWC SemTab 2020 Challenge
  - Google Faculty Research award
  - Intel Innovator award

- Event organization:
  - Hosted and co-organized the second workshop on Functional Inference and Machine Intelligence (FIMI) with the Institute of Statistical Mathematics, Tokyo, Japan.
  - Co-organized the second workshop on AI for Smart TV Content Production, Access and Delivery (AI4TV) colocated with ACM Multimedia 2020, https://memad.eu/ai4tv2020/
  - Co-organized the first workshop on Semantic and Knowledge Graph advances for Journalism (SKG4J) colocated with CIKM 2020, https://almoslmi.github.io/SemanticJournalism/
  - Co-organized the first workshop on Semantics for Online Misinformation Detection, Monitoring, and Prediction (Semiform) colocated with ISWC 2020, https://d2klab.github.io/semiform2020/

- Chair role: Conferences, Workshops
  - TheWebConference - TWC 2022 - General Chair
  - Artificial Intelligence and Statistics Conference - AISTATS 2020 - Area Chair
  - Medical Image Computing and Computer-Assisted Interventions (MICCAI) 2020 - Program Co-chair (https://www.miccai2020.org/en/ORGANIZING-COMMITTEE.html)
  - IEEE International Symposium on Biomedical Imaging (ISBI) 2021 - Organizing Committee
  - Conference for Truth and Trust Online - TTO 2021 - Program co-Chair
  - European Semantic Web Conference (ESWC 2020) - Poster and Demo Chair
  - International Conference on Computer Software and Applications (COMSAC 2020), Autonomous Systems Track Co-chair

Press coverage

DAGOBAH: semantic annotation of tabular data
- Hello Future: https://hellofuture.orange.com/fr/Dagobah-un-tableau-ne-parle-que-par-celui-qui-sait-lannoter/
- IMT (EN): https://imtech.wp.imt.fr/en/2020/10/19/dagobah-tables-ai-will-understand/
- IMT (FR): https://imtech.wp.imt.fr/2020/06/22/dagobah-les-tableaux-lia-comprendra/

Odeuropa: re-creating an encyclopedia of the old scents
- France: Le Monde, Le Figaro, France Inter (radio), France 5 (TV)
- International: New York Times, Forbes, The Guardian, RTBF, La Republica, La Stampa, El Pais, etc.
- Complete coverage: https://odeuropa.eu/horizon-2020/press/

CoronaCheck fact checking
- Corriere della Sera (IT), Italian first national newspaper on 20/03/2020 and 14/12/2020
- TechXplore (EN) https://techxplore.com/news/2020-04-coronacheck-website-combats-misinformation.html
- IMT (EN): https://imtech.wp.imt.fr/2020/04/02/coronacheck-demeler-le-vrai-du-faux-sur-lepidemie-de-covid-19/
- IMT(FR): https://imtech.wp.imt.fr/en/2020/04/14/coronacheck-separating-fact-from-fiction-in-the-covid-19-epidemic/
- Complete coverage: https://coronacheck.eurecom.fr/en/press

1. MACHINE LEARNING RESEARCH

1.1 Computational aspects in Statistical Inference

Summary: In this research track, the main general question that we aim to address is how to use data and expert domain knowledge to make decisions. Today, we have access to so much data generated by a variety of sensors, but we are facing difficulties in using these data in a sensible way. Machine Learning and Statistics offer the main tools to help make sense of data, and novel techniques in this domain will be used and developed throughout this project. Quantification of risk and decision-making require accurate quantification of uncertainty, which is a major challenge in many areas of sciences involving complex phenomena like in finance, environmental and life sciences. In order to accurately quantify uncertainty, we employ flexible and accurate tools offered by modern statistical models. However, today's diversity and abundance of data make it difficult to use these models in practice. The goal of this project is to propose new ways to better manage the interface between computational and statistical models - which in turn will help get accurate quantification of the confidence in predictions based on observed data.

The novel contributions we made this year are as follows:

- Accelerating optimization of approximate inference for Deep Learning: Last year we started a new research direction in which we were investigating ways to accelerate inference for Bayesian deep neural/convolutional networks. The idea is to use a reparameterization of the weights of these networks inspired by the literature on kernel methods, which uses pseudo-random matrices for which it is cheap to perform computations. Such a reparameterization has the advantage of using far less parameters than the traditional parameterization of these networks; as a result, the proposed reparameterization avoids all the pathologies (and in particular over-regularization) stemming from the over-parameterization of these models when employing variational inference, while tremendously accelerating computations. This year we have finalized this work and we had it published at NeurIPS 2020 (A* conference), whereas a reduced version of this was presented at Bayesian Deep Learning workshop at NeurIPS at the end of 2019.

- Novel Sparsification for Scalable Gaussian Processes: Approximations to Gaussian processes based on inducing variables, combined with variational inference techniques, enable state-of-the-art sparse approaches to infer Gaussian processes at scale through mini batch-based learning. In this work, we address one limitation of sparse Gaussian processes, which is due to the challenge in dealing with a large number of inducing variables without imposing a special structure on the inducing inputs. In particular, we introduced a novel hierarchical prior, which imposes sparsity on the set of inducing variables. We treated our model variationally, and we experimentally showed considerable computational gains compared to standard sparse Gaussian processes when sparsity on the inducing variables is realized considering the nearest inducing inputs of a random mini-batch of the data. Through an extensive experimental validation we demonstrated that our approach is competitive with the state-of-the-art, but crucially it enables the possibility to implement sparse Gaussian processes with a large number of inducing points without incurring a prohibitive computational cost. For example, we showcased sparse Gaussian processes for a large-scale classification problem (millions of observations) using hundreds of thousands of inducing points; we are not aware of other approaches that can handle such a large set of inducing inputs without imposing some special structure on them (e.g., grid) or without considering one-dimensional inputs. This work is currently under review in an A conference.

- Fully Bayesian Deep Gaussian Processes: Variational inference techniques based on inducing variables provide an elegant framework for scalable posterior estimation in Gaussian process (GP) models. Besides enabling scalability, one of their main advantages over sparse approximations using direct marginal likelihood maximization is that they provide a robust alternative for point estimation of the inducing inputs, i.e. the location of the inducing variables. In this work we challenged the common wisdom that optimizing the inducing inputs in the variational framework yields optimal performance. We showed that, by revisiting old model approximations such as the fully-independent training conditionals endowed with powerful sampling-based inference methods, treating both inducing locations and GP hyper-parameters in a Bayesian way can improve performance significantly. Based on stochastic gradient Hamiltonian Monte Carlo, we developed a fully Bayesian approach to scalable GP and deep GP models, and demonstrated its state-of-the-art performance through an extensive experimental campaign across several regression and classification problems. This work is currently under review in an A conference.

- Novel Priors for Bayesian Deep/Conv Nets: The Bayesian treatment of neural networks dictates that a prior distribution is specified over their weight and bias parameters. This poses a challenge because modern neural networks are characterized by a large number of parameters, and the choice of these priors has an uncontrolled effect on the induced functional prior, which is the distribution of the functions obtained by sampling the parameters from their prior distribution. Our investigation into this issue started by realizing that this is a hugely limiting aspect of Bayesian deep learning, and with this work we wanted to tackle this limitation in a practical and effective way. Our proposal is to reason in terms of functional priors, which are easier to elicit, and to "tune" the priors of neural network parameters in a way that they reflect such functional priors. Gaussian processes offer a rigorous framework to define prior distributions over functions, and we propose a novel and robust framework to match their prior with the functional prior of neural networks based on the minimization of their Wasserstein distance. We provided vast experimental evidence that coupling these priors with scalable Markov chain Monte Carlo sampling offers systematically large performance improvements over alternative choices of priors and state-of-the-art approximate Bayesian deep learning approaches. We consider this work a considerable step in the direction of making the long-standing challenge of carrying out a fully Bayesian treatment of neural networks, including convolutional neural networks, a concrete possibility. This work is currently under review in the Journal of Machine Learning Research, and we have a smaller version of this work accepted at the Advances in Approximate Bayesian Inference (AABI 2020) symposium.

- Novel Deep Gaussian Processes for spatio-temporal modeling: Spatial processes with nonstationary and anisotropic covariance structure are often used when modelling, analyzing and predicting complex environmental phenomena. Such processes may often be expressed as ones that have stationary and isotropic covariance structure on a warped spatial domain. However, the warping function is generally difficult to fit and not constrained to be injective, often resulting in ‘space-folding.’In this work, we proposed modelling an injective warping function through a composition of multiple elemental injective functions in a deep-learning framework. Inspired by recent methodological and technological advances in deep learning and deep Gaussian processes, we employed approximate Bayesian methods to make inference with these models at scale. This work has been done in collaboration with the University of Wollongong, Australia, and it has just been accepted for publication in the very selective and prestigious Journal of the American Statistical Association (JASA) - impact factor 3.898.

Team: Maurizio Filippone, Pietro Michiardi, Dimitrios Milios, Simone Rossi, Gia-Lac Tran, Ba-Hien Tran.
Funding: AXA Chair
References:

- S. Rossi, S. Marmin, M. Filippone. Walsh-Hadamard Variational Inference for Bayesian Deep Learning. In Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems, NeurIPS, 2020.
- S. Rossi, S. Marmin, M. Filippone. Efficient Approximate Inference with Walsh-Hadamard Variational Inference. In NeurIPS 2019, 32nd Conference on Neural Information Processing Systems, Workshop on Bayesian Deep Learning, 2019.
- B.-H. Tran, S. Rossi, D. Milios, M. Filippone. All You Need is a Good Functional Prior for Bayesian Deep Learning. arXiv:2011.12829.
- G.-L. Tran, D. Milios, P. Michiardi, M. Filippone. Sparse within Sparse Gaussian Processes using Neighbor Information. arXiv:2011.05041.
- S. Rossi, M. Heinonen, E. V. Bonilla, Z. Shen, M. Filippone. Sparse Gaussian Processes Revisited: Bayesian Approaches to Inducing-Variable Approximations. arXiv:2003.03080.
- B.-H. Tran, S. Rossi, D. Milios, M. Filippone. All You Need is a Good Functional Prior for Bayesian Deep Learning. arXiv:2011.12829.
- A. Zammit-Mangion, T.L. Ng, Q.H. Vu, M. Filippone. Deep Compositional Spatial Models. Journal of the American Statistical Association, to appear.

1.2 Machine Learning on Optical Processing Units

Summary: Information and Communication Technologies (ICT) are constantly producing advancements that translate into a variety of societal changes including improvements to economy, better living conditions, access to education, well-being, and entertainment. The widespread use and growth of ICT, however, is posing a huge threat to the sustainability of this development, given that the energy consumption of current computing devices is growing at an uncontrolled pace.

Within ICT, machine learning is currently one of the fastest growing fields, given its pervasive use and adoption in smart cities, recommendation systems, finance, social media analysis, communication systems, and transportation. Apart from isolated application-specific attempts, the only general solution to tackle the sustainability of computations in machine learning is Google's Tensor Processing Unit (TPU), which has been opened to general use through a cloud system in 2018. This is an interesting and effective direction to push a transistor-based technology to address some of the issues above pertaining to the sustainability of computing for machine learning, and it is inspiring other companies and start-ups to follow this trend.

ECO-ML's ambition is to radically change this and to propose a novel angle of attack to the sustainability of computations in machine learning. The models that will be studied in ECO-ML are Gaussian Processes (GPs), Deep Gaussian Processes (DGPs) and Bayesian Deep Neural Networks (DNNs), as they offer attractive flexibility and quantification of uncertainty.

In the last year, we have come across the work that the French company LightOn has done on the development of novel Optical Processing Units (OPUs). OPUs perform a specific matrix operation in hardware exploiting the properties of scattering of light, so that in practice this happens at the speed of light. Not only this is the case, but the consumption of OPUs is much lower than current computing devices, while allowing for the possibility to operate with large Gaussian random matrices, orders of magnitude larger than current computing devices. GP and DGP models are perfect candidates to benefit from the principles behind OPUs, but there is a need to make advancements on the design and inference of these models for this to become a reality. We expect to produce and release the first implementation of GPs and DGPs using OPUs, and to demonstrate that this leads to considerable acceleration in model training and prediction times while reducing power consumption with respect to the state-of-the-art. We expect to advance the state-of-the-art in GP, DGP and DNN modeling and inference by developing novel model approximations and inference tailored to exploit OPU computing, but that will also trigger advances in the theory of approximation for these models. Finally, we expect to showcase a variety of modeling applications in environmental and life sciences, demonstrating that our approach leads to competitive performance with the state-of-the-art, while achieving sound quantification of uncertainty and fast model training and prediction times in a sustainable way.

In 2020, we have mostly worked on some theoretical aspects of polynomial kernels, with the aim to obtain guarantees on their approximation using random features. We are about to complete a very long and in-depth study on this, which we aim to submit for publication to the Journal of Machine Learning Research.

Team: Maurizio Filippone, Sebastien Marmin, Jonas Wacker, Bogdan Kozyrskiy.
Funding: ANR JCJC “ECO-ML” and 3IA Chair
References:

- R. Ohana, J. Wacker, J. Dong, S. Marmin, F. Krzakala, M. Filippone, L. Daudet. Kernel computations from large-scale random features obtained by Optical Processing Units. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020.

1.3 Statistical learning for computer simulation

Summary: The aim of this project is to enhance the power of computer simulation via the development of relevant, novel machine learning algorithms. Computer simulation has been widely used in many fields of science and engineering for modeling complex time-evolving phenomena, such as climate, weather, weather, epidemics, traffic, pedestrian movement, and so on. However, there are several challenges to deploy such simulators in practice. These include 1) whether the

simulator is reliable and 2) how to manage the high computational costs of simulations. This project aims at developing and using novel machine learning algorithms to deal with such challenges.

In the past year, we worked on several research projects:

1) Machine learning approach to calibrating a simulation model.

To make a simulation model reliable, we need to appropriately select parameters that define the simulation model. This is in general hard to do manually, since it requires several try-and-errors of setting candidate parameters and comparing the simulation results to real data. We developed a machine learning algorithm that automates this parameter tuning procedure, which results in a more reliable simulator. The result was published at AISTATS 2020, a top-tier machine learning conference, and we are working on its extensions (work in progress).

Funding: NEC corporation Team: Motonobu Kanagawa
External Collaborators: Keiichi Kisamori (NEC/AIST, Japan), Keisuke Yamazaki (AIST, Japan)
References:

- Simulator calibration under covariate shift with kernels K Kisamori, M Kanagawa, K Yamazaki International Conference on Artificial Intelligence and Statistics (AISTATS), 1244-1253, 2020

2) Machine learning algorithm for emulating a simulation model, with application to tsunami early warning

Computer simulators are usually computationally very expensive, requiring rich computational resources and computational time. This becomes problematic in particular when simulators are supposed to be used in real-time forecasting. An illustrative example is tsunami early warning, the task of forecasting the inundation of a tsunami just after it happens. While high-quality simulators for tsunami modeling are available, they need too long time (~hours) to simulate for this purpose (a tsunami arrives in minutes). Thus, there is a great need for accelerating computer simulators.

We developed a machine learning method for accelerating computer simulations. This method learns to imitate the behavior of a simulation model in a black-box manner: this is called ``emulation" in the literature. Since the learned machine learning model is computationally much cheaper to run, it can be used as a surrogate for the simulation model in real-time forecasting. We applied our method for a tsunami simulator for the task of inundation forecasting, and our approach performed much better than an existing method for this purpose. We are currently mathematically analyzing the developed method to provide theoretical guarantees, and preparing a submission to the Journal of Machine Learning Research (JMLR).

Funding: NEC corporation
Team: Motonobu Kanagawa
External Collaborators: Takafumi Kajihara (NEC, Japan), Wataru Suzuki (NIED, Japan)

Other ongoing projects include the following:

3) Theoretical foundations of Gaussian processes and kernel methods

Gaussian processes and kernel methods are the two widely used approaches in machine learning based on positive definite kernels. It is well known that these two formalisms are closely related; for instance, the estimator of kernel ridge regression is identical to the posterior mean of Gaussian process regression. However, they have been studied and developed almost independently by two essentially separate communities, and this makes it difficult to seamlessly transfer results between them. To overcome this potential difficulty, this work discusses the connections and equivalences between the two approaches in a unifying manner. We also provide discussions on subtle philosophical and theoretical differences between the two approaches.

Team: Motonobu Kanagawa
External collaborators: Philipp Hennig (University of Tuebingen and Max Planck Institute, Germany), Dino Sejdinovic (University of Oxford, UK), Bharath Sriperumbudur (Penn State University, USA)

References:

- Kanagawa, M., Hennig, P., Sejdinovic, D., & Sriperumbudur, B. K. Gaussian processes and kernel methods: A review on connections and equivalences. arXiv preprint arXiv:1807.02582. (In revision for resubmission to the Journal of Machine Learning Research.)

4) A novel methodology for counterfacutal causal inference

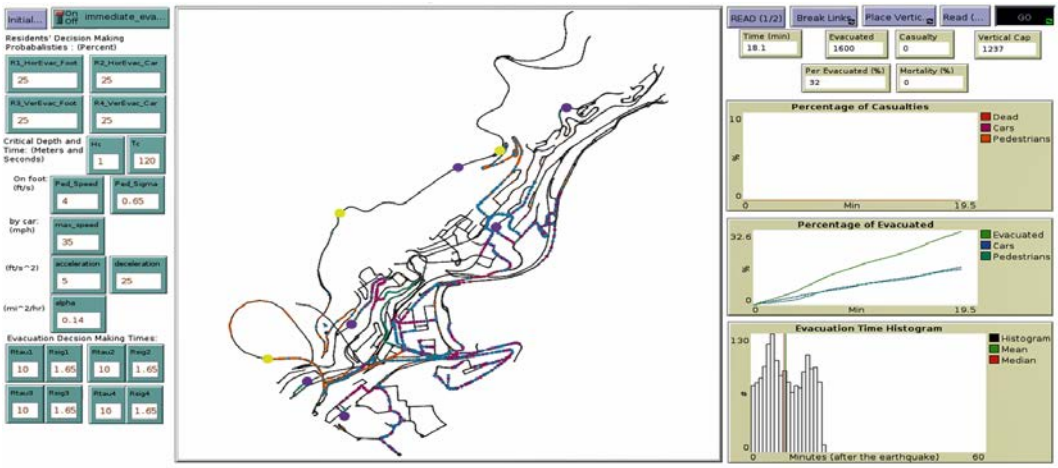
Counterfactual inference has become a ubiquitous tool in online advertisement, recommendation systems, medical diagnosis, and econometrics. Accurate modeling of outcome distributions associated with different interventions---known as counterfactual distributions---is crucial for the success of these applications. In this work, we propose to model counterfactual distributions using a novel Hilbert space representation called counterfactual mean embedding (CME). The CME embeds the associated counterfactual distribution into a reproducing kernel Hilbert space (RKHS) endowed with a positive definite kernel, which allows us to perform causal inference over the entire landscape of the counterfactual distribution. Based on this representation, we propose a distributional treatment effect (DTE) which can quantify the causal effect over entire outcome distributions. Our approach is nonparametric as the CME can be estimated consistently from observational data without requiring any parametric assumption about the underlying distributions. We also establish a rate of convergence of the proposed estimator which depends on the smoothness of the conditional mean and the Radon-Nikodym derivative of the underlying marginal distributions. Furthermore, our framework also allows for more complex outcomes such as images, sequences, and graphs. Lastly, our experimental results on synthetic data and off-policy evaluation tasks demonstrate the advantages of the proposed estimator.

Team: Motonobu Kanagawa
External Collaborators: Krikamol Muandet (Max Planck Institute, Germany), Sorawit Saengkyongam (University of Copenhagen, Denmark)
References:

- Muandet, K., Kanagawa, M., Saengkyongam, S., & Marukatat, S. Counterfactual Mean Embeddings. arXiv preprint arXiv:1805.08845. (Accepted with minor revision by the Journal of Machine Learning Research.)

5) Simulations for evaluation planning in the city of Monaco

We developed a prototype computer simulator for evacuation planning (Figure below). This simulator has been developed for the purpose of modeling possible evacuation scenarios, in which Monaco citizens evacuate from their residential or working areas to safer evacuation destinations (e.g., shelters). Such simulations are necessary for emergency managers to understand the quality or the problem of a candidate evacuation plan since possible incidents that cause evacuations (e.g., flood, fire, and attack) are rare events, and thus the only way of studying evacuations is via simulations. Moreover, evacuation simulations are also useful in identifying possible problems in the structures of the city that prevent prompt evacuation, such as narrow loads that become bottlenecks of evacuation pathways and cause traffic jams.



Funding: Monaco Government
Team: Motonobu Kanagawa, Pietro Michiardi, Kensuke Mitsuzawa

6) Simulations for a defined-contribution pension fund utilizing intergenerational risk sharing

We propose a novel design for a collective defined-contribution (CDC) pension fund, and perform simulations to study its properties. Our pension design is an overlapping generation model in which each generation maintains its own benefit account, which keeps track of the evolution of the asset belonging to that generation. Our pension fund is also equipped with a collective buffer account based on a funding-ratio-linked declaration rate, which effectively smooth the returns (i.e., benefits and losses) of consecutive generations, enabling intergenerational risk sharing. In the pension design, there are two main parameters that determine a constant-mix strategy for investment and the funding-ratio declaration rate. To determine these parameters, we formulate an optimization problem based on the expected utility of a social planner, and apply Bayesian optimization, a modern machine learning approach to global optimization. Simulation studies show that the evolution of an individual benefit account in our pension fund is much smoother than the corresponding individual investment plan without any return-smoothing mechanism, demonstrating the intergenerational risk sharing in our pension design. Moreover, simulations demonstrate that our pension design is more robust in a volatile financial market than the corresponding individual investment design.

Team: Motonobu Kanagawa
External collaborators: An Chen, Fangyuan Zhang (University of Ulm, Germany)
- Preparing a paper in submission to an insurance/economics journal

1.4 Robust Machine Learning

Summary: Despite the numerous success stories of machine learning (ML)-based systems, their application remains to be challenging. This can be explained by three factors. First, data from production environments is complex and expensive to collect and process. Second, many application domains rely on critical, high-risk, high-impact decisions, where there is little tolerance for errors. Thirdly, the current lack of reproducibility of results among the research community hinders the translation of new methods into practice. Despite advances in the field to address these factors, there is still a need for robust and reliable methods that can be safely deployed. This research line focuses on the development of reliable learning frameworks by addressing the problems of data complexity and low tolerance to errors. On the methodological side, it has been structured around three axes: 1) Efficient use of data, 2) Error assessment to assist high-risk decision making, and 3) Definition and setup of evaluation and validation frameworks. From an application perspective, the research line on robust machine learning is targeting but not limited to applications in healthcare, as this is a field particularly challenged by the problems previously discussed.

Efficient use of data: Multi-task Learning

Multi-task learning has gained popularity due to the advantages it provides with respect to resource usage and performance. Nonetheless, the joint optimization of parameters with respect to multiple tasks remains an active research topic. Sub-partitioning the parameters between different tasks has proven to be an efficient way to relax the optimization constraints over the shared weights, may the partitions be disjoint or overlapping. However, one drawback of this approach is that it can weaken the inductive bias generally set up by the joint task op-timization. In this work, we present a novel way to partition the parameter space without weakening the inductive bias. To address these problems, we have proposed a novel method, denoted Maximum Roaming, which is inspired by dropout that randomly varies the parameter partitioning, while forcing them to visit as many tasks as possible at a regulated frequency, so that the network fully adapts to each update. Through a thorough experimental setup we show that the regularization brought by roaming has more impact on performance than usual partitioning optimization strategies.

Team: Lucas Pascal, Pietro Michiardi, Maria A. Zuluaga
Funding: PhD CIFRE with Orkis

References:

- L. Pascal, P. Michiardi, X. Bost, B. Huet, M.A. Zuluaga. Maximum Roaming Multi-Task Learning. Accepted to: The 35th AAAI Conference on Artificial Intelligence

Efficient use of data: Simplifying the annotation process of complex data

The use of deep learning techniques for 3D brain vessel image segmentation has not been as widespread as for the segmentation of other organs and tissues. The use of deep learning techniques for 3D brain vessel image segmentation has not been as widespread as for the segmentation of other organs and tissues. This can be explained by two factors. First, deep learning techniques tend to show poor performances at the segmentation of relatively small objects compared to the size of the full image. Second, due to the complexity of vascular trees and the small size of vessels, it is challenging to obtain the amount of annotated training data typically needed by deep learning methods. We have proposed a novel annotation-efficient deep learning vessel segmentation framework. The framework avoids pixel-wise annotations, only requiring patch-level labels to discriminate between vessel and non-vessel 2D patches in the training set, in a setup similar to the CAPTCHAs used to differentiate humans from bots in web applications. The user-provided annotations are used for two tasks: 1) to automatically generate pixel-wise labels for vessels and background in each patch, which are used to train a segmentation network, and 2) to train a classifier network. The classifier network allows to generate additional weak patch labels, further reducing the annotation burden, and it acts as a noise filter for poor quality images. Our experimental results demonstrated that the framework achieves state-of-the-art accuracy, while reducing the annotation time by up to 80% with respect to learning-based segmentation methods using pixel-wise labels for training.

Team: Maria A. Zuluaga
Funding: N/A
External Collaborators: INRIA Sophia Antipolis (France), University College London (UK), UCL Institute of Neurology (UK) and University of Siena (Italy).
References:
Journal paper under review

Error assessment: Anomaly detection

Detecting unexpected behavior on multivariate data is an active research, where a large set of learning-based methods have been proposed over the past years. As data has reached a complexity that no longer allows the use of more traditional methods, most recently, the ability of unsupervised anomaly detection methods based on deep learning. Among these, methods based on Recursive Neural Networks (RNNs) and Generative Adversarial Networks (GANs) are very popular. However, the former are well-known for being computationally greedy and for requiring a significant amount of time to be trained, whereas the latter are problematic to train, due their lack of stability, impeding deployment in production. A production environment requires high scalability and stability that obliges to rethink the key requirements that deep learning methods should meet. We addressed these constraints by proposing a novel method based on an Autoencoder (AE) architecture, which learns under a scheme inspired by GANs. The setup of an encoder-decoder architecture within an adversarial training framework allows to combine the advantages of autoencoders and adversarial training, while compensating for the limitations of each technique, making the resulting method fast to train, scalable and robust, while achieving a high performance.

Team: Julien Audibert, Pietro Michiardi, Maria A. Zuluaga
Funding: PhD contract with Orange
References:

- J. Audibert, P. Michiardi, F. Guyard, S. Marti, M.A. Zuluaga. USAD: Unsupervised Anomaly Detection in multivariate time series. In: SIGKDD 2020

Error assessment: Monitoring model's performance

We have been studying the development of models that can be used to monitor and maintain (model selection) other systems. Model monitoring refers to the task of constantly tracking a model's performance to determine when it degrades, becoming obsolete. Once a degradation in performance is detected, model maintenance and improvement take place to update the deployed model by rebuilding it, recalibrating it or, more generally, by doing model selection. The contributions in this line are on two fronts. First, in the development of monitoring models in the context of time-series forecasting. Second, in the monitoring of image processing pipelines for the analysis of cardiac magnetic resonance images.

Team: Maria A. Zuluaga
Collaborators: Rosa Candela, Pietro Michiardi, Maurizio Filippone
Funding: Monaco Government
References:

- R. Candela, P. Michiardi, M. Filippone, MA. Zuluaga. Model Monitoring and Dynamic Model Selection in Travel Time-series Forecasting. In: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases ECML-PKDD 2020
- Conference paper under review on clinical pipelines monitoring

AI for Healthcare: Wearable devices

The joint use of smart cities sensors (SCS) and wearables and wearable biometric monitoring devices (WBMDs) data has the potential to provide a holistic view of a subject's condition and context in real time, within natural settings and in a continuous fashion, thus providing complementary information to standard, more precise, but more costly and less available clinical information. This setup represents a big step forward in the collection of health-related data from subjects and patients. However, SCSs and WDBMs continuous data flow makes it is nearly impossible to analyze manually. Without the possibility of curating data, the problems linked to it are exacerbated. We work in the development of efficient methods that can jointly exploit SCSs and WBMDs data in two different applications: 1) remote diagnosis of COVID-19 through the analysis of coughing and 2) detection of challenging behaviours in children with autism.

Team: Maria A. Zuluaga
Internal Collaborators: Massimiliano Todisco (Security Department)
External Collaborators: IMT Saint-Étienne, INRIA Sophia Antipolis, CHU Grenoble, CHU Lenval (Nice)

1.5 Learning interpretable representations

Interpretable Machine Learning

Summary: Model interpretability has become an important factor to consider when applying machine learning in critical application domains. In medicine, law, and predictive maintenance, to name a few, understanding the output of the model is at least as important as the output itself. However, a large fraction of models currently in use (e.g. Deep Nets, SVMs) favor predictive performance at the expense of interpretability.

The lack of transparency and accountability of predictive models can have (and has already had) severe consequences. For example, there have been cases of people incorrectly denied parole, poor bail decisions leading to the release of dangerous criminals, pollution models stating that highly polluted air was safe to breathe, and many more.

In our work, we develop our research along two lines.

We focus on predictive rule learning for challenging applications where data is unbalanced. For rules, interpretability translates into simplicity, and it is measured as a function of the number of rules and their size (average number of atoms): such proxies are easy to compute, understandable, and allow comparing several rule-based models.

Our goal is to learn a set of rules from the training set that (i) effectively predict a given target, (ii) generalize to unseen data, (iii) and are interpretable, i.e., a small number of short rules. The first objective is particularly difficult to meet in the presence of imbalanced data. In this case, most rule-based methods fail at characterizing the minority class. Additional data issues

that hinder the application of rule-based methods are data fragmentation, overlaps between imbalanced classes, and presence of rare examples.

As a result, we have produced a rule-based method that is (i) versatile and effective in dealing with both balanced and imbalanced data, (ii) interpretable, as it produces small and compact rule sets, and (iii) scalable to big datasets.

A second line of work relates to studying latent representation of data. Representation learning aims at learning data representations such that it is easier to extract useful information when building classifiers or other predictive tasks. Representation learning seeks to obtain the following properties: i) expressiveness: a reasonably-sized representa- tion should allow to distinguish among a high number of dif- ferent input configurations; ii) abstractness: learned repre- sentations should capture high-level features; iii) invariance: representation should be invariant to local changes of input configurations; iv) interpretability: learned representations should allow each dimension to be informative about the given task. These properties are at the core of disentangled representations. In disentangled representation learning, the main assumption is that high-dimensional observations x are the result of a (possibly nonlinear) transformation applied to a low dimensional latent variable of independent generative factors, called ground-truth factors, capturing semantically meaningful concepts. Input observations can be thought of as the result of a probabilistic generative process, where latent variables z are first sampled from a prior distribution p(z), and then the observations x are sampled from p(x|z). The goal is to learn a representation of the data that captures the generative factors.

In this line of work, we focus on deep generative models, and in particular those based on variational autoencoders (VAE), to learn disentangled representations

In this work, we propose a novel generative model that uses a conditional prior and has theoretical identifiability guarantees. We show that our method naturally imposes an optimality constraint, in information theoretic terms, on the conditional prior: this improves the regularization on the function that maps input observations to latent variables, which translates in tangible improvements of disentanglement in practice.

Since assuming to have access to auxiliary variables for each input observations, both at training and testing time, is not practical in many applications, we also propose a semi-supervised variant of our method.

Team: Graziano Mita, Maurizio Filippone, Pietro Michiardi
Funding: CIFRE contract with SAP Labs France (Sophia Antipolis)
References:

- Graziano Mita, Paolo Papotti, Maurizio Filippone, Pietro Michiardi, “LIBRE: Learning Interpretable Boolean Rule Ensembles”, https://arxiv.org/abs/1911.06537, (published in AISTATS 2020)
- Graziano Mita, Maurizio Filippone, Pietro Michiardi, “Learning Optimal Conditional Priors For Disentangled Representations”, https://arxiv.org/abs/2010.09360, (in submission)

Learning multi-modal representations

Summary: This line of work stems from the context of Advanced Driving Assistance Systems, Autonomous Driving (AD), robotics, and general autonomous cognitive systems. At the high level, we tackle the problem of sensor data fusion, which is the task of combining sensory data or data derived from disparate sources such that the resulting information has less uncertainty than would be possible when these sources were used individually. The term uncertainty reduction in this case can mean more accurate, more complete, or more dependable, or refer to the result of an emerging view, such as stereoscopic vision (calculation of depth information by combining two-dimensional images from two cameras at slightly different viewpoints).

The sensor fusion task is fundamental to gather the information from any available sensor and to construct a reliable model of the surrounding environment (i.e. World Model) to predict and apprehend all the discrete elements (i.e. objects) and the infrastructure (i.e. road, walls) of which it is composed. The fusion task can be performed on data from a set of heterogeneous or homogeneous sensors, soft sensors, and their history values: these sensors can include exteroceptive (Camera, LIDAR, RADAR, ultrasounds, etc) and proprioceptive sources (Odometers, Inertial Measurement Units, GNSS, etc), HD-Maps or distributed sensors (i.e. not embedded in the cognitive system).

The key idea of our work is to tackle the challenges described above by learning generative models that span multiple data modalities, which are motivated by the desire to learn more useful, generalisable representations that faithfully capture common underlying factors between the modalities. We characterise successful learning of such models as the fulfilment of four criteria: i) implicit latent decomposition into shared and private subspaces, ii) coherent joint generation over all modalities, iii) coherent cross-generation across individual modalities, and iv) improved model learning for individual modalities through multi-modal integration.

An additional goal that spans the overall research we describe in this section is to impose interpretability constraints, by learning disentangled representations. Since the objective of any generative model is essentially to capture underlying data generative factors, the disentangled representation would mean a single latent unit being sensitive to variations in single generative factors. This gives rise to an important research question: "how to achieve disentanglement without losing the reconstruction abilities underlying the typical training of generative models?" Furthermore, the ultimate goal of our research is to improve latent representations through multi-modal fusion: such latent representations are then fed to downstream components such as, e.g., object detection and scene segmentation models.

Team: Matthieu Da Silva-Filarder, Maurizio Filippone, Pietro Michiardi
Funding: CIFRE contract with Renault Software Labs (Sophia Antipolis)
References:

- 3 patents submitted

1.6 Stochastic optimization for large-scale machine learning

Markov Chain MonteCarlo, through the lenses of Stochastic differential equations and Hamiltonian Dynamics

Summary: Stochastic gradient descent (SGD) has become crucial to modern machine learning. SGD optimizes a function by following noisy gradients with a decreasing step size. Classical results indicate that this procedure provably reaches the optimum of the function (or local optimum, when it is nonconvex). Stochastic gradient descent has enabled efficient optimization with massive data, since one can often obtain noisy-but-unbiased gradients very cheaply by randomly subsampling a large dataset. Recently, stochastic gradients (SG) have also been used in the service of scalable Bayesian Markov Chain Monte Carlo (MCMC) methods, where the goal is to generate samples from a conditional distribution of latent variables given a data set.

Our work stems from the quest for studying MCMC methods through the lenses of statistical physics theoretical tools. In particular, we view the process underlying MCMC method as the discretization of a continuous-time stochastic process, which we describe using the language of differential equations. In particular, we focus on Hamiltonian dynamics, which is a sophisticated formulation of classical mechanics describing the time evolution of a physical system subject to kinetic and potential forces. In our work, we introduce a novel view on the stochasticity deriving from using mini-batch stochastic gradients, which we interpret as a partitioning scheme of an infinitesimal operator representing the gradient. Our theory allows: to explain how and why existing schemes work, often demystifying the original explanations attributing the advantages of existing proposals to mechanisms to combat noisy estimates of the loss function gradient; to study the role of high-order integration schemes, and the impact on the numerical stability of the algorithms to simulate a dynamical system; to propose novel approaches to MCMC, which perform better than alternative and that are simpler to compute.

Team: Giulio Franzese, Dimitrios Milios, Maurizio Filippone, Pietro Michiardi
Funding: n/a
References:

- Giulio Franzese, Rosa Candela, Dimitrios Milios, Maurizio Filippone, Pietro Michiardi, "Isotropic SGD: a Practical Approach to Bayesian Posterior Sampling", https://arxiv.org/abs/2006.05087

1.7 Distributional Reinforcement Learning

Summary: The context of this project is that of designing and analyzing methods to enable safe autonomous vehicle operations. Over the course of it's advancement, RL has been shown to be able to tackle the task of control in progressively more and more complex environments. However one of the limitations inherent to any RL problem, is the need to define

the elements of the Markov Decision Process (MDP) model, which RL algorithms are well-adapted to solve. This poses the problem that any control algorithm is only as good as the environment in which it has been trained. Because MDPs usually seek to model real-world interactions with inherent simplifications, they are susceptible to change, when the MDP model's limitations become apparent. In the case where the defining elements of an MDP changes, so does the optimal policy. Because of the infamous sample inefficiency of RL, generalizing to different environments can prove extremely challenging.

In this line of work, we introduce a training paradigm for a reinforce-ment learning agent in the model-free setting, which we name concurrent learning, based on expliciting potentially useful behaviours, in order to increase the ability of an agent to tackle changes to the stochastic parameters of an MDP. Our aim is to allow the agent to make use of the multitude of behaviours that it learns during its training and exploration phase, inorder to access them more easily later on when performing in a modified version of the MDP in which it has been trained.

The main insight of our approach is to improve upon expected-value algorithms which optimize agents' behaviour with regards to the expectation of return, and may miss out nuances in the dynamics of the environment when the uncertainty parameters of the dynamics are changed. We argue for the practical utility of learning suboptimal behaviours now, in the case that the knowledge gained by the agent may be transferred to a similar environment.

Team: Ugo Lecerf, Maurizio Filippone, Pietro Michiardi
Funding: CIFRE contract with Renault Software Labs (Sophia Antipolis)
References: n/a

2. KNOWLEDGE GRAPHS AND DATA MANAGEMENT

2.1 Novel Hardware for Storage and Databases

Summary: The demand for data-driven decision making coupled with the need to retain data to meet regulatory compliance requirements has resulted in a rapid increase in the amount of archival data stored by enterprises. As data generation rate far outpaces the rate of improvement in storage density of media like HDD and tape, research on alternate media technologies, like film, synthetic DNA, and glass, for long-term data archival has received a lot of attention recently.

In 2020, we collaborated with EUPALIA, an innovative French startup specializing in digital preservation to demonstrate how analog media like paper, microform, or film, can be used for long-term data archival. More specifically, while researchers have developed novel layout and encoding techniques for archiving databases on these new media types, one key question remains unaddressed: How do we ensure that the decoders developed today will be available and executable by a user who is restoring an archived database several decades later in the future, on a computing platform that potentially does not even exist today? In our work, we introduced Universal Layout Emulation (ULE), a new approach for future-proof, long-term database archival that advocates archiving decoders together with the data to ensure successful recovery. In order to do so, ULE brings together concepts from Data Management and Digital Preservation communities by using emulation for archiving decoders. In order to show that ULE can be implemented in practice, we designed Micr'Olonys, an end-to-end long-term database archival system that can be used to archive databases using visual analog media like film, microform, and archival paper.

In October 2019, we started a three year project OligoArchive, where we will research all relevant aspects of using DNA as a digital storage medium. We have established a consortium of six partners across three countries (UK, France, Ireland) bringing together the necessary expertise to explore encoding different types of data in DNA, scalable DNA synthesis to store data, experimental techniques to manipulate the data, efficient sequencing and decoding approaches to read back the data as well as automation. Last year, we made substantial progress in developing scalable decoding techniques for recovering data from DNA. The process of retrieving data from DNA is computationally bottlenecked by a key read consensus stage that effectively performs an edit similarity join to identify millions of unique consensus strings from hundreds of millions of noisy copies. Current join algorithms are unable to scale to such datasets due to the inherent complexity of edit distance computation, and due to their single-threaded design rendering them unable to exploit accelerators like GPUs. Last year, we developed OneJoin, a cross-architecture edit similarity join that can exploit multicore CPUs, integrated GPUs, and multi-vendor discrete GPUs using a single code base. Central to the effectiveness of OneJoin is the use of oneAPI–an open, standards-based unified programming model for achieving portable data parallelism. We showed that OneJoin can provide up to 21×improvement in performance over other state-of-the-art joins and reduce the overall DNA data decoding time from several hours to just a few minutes. This work has been recognized by Intel with an Intel Innovation Award. This work was highlighted with an invited talk during Intel DevSummit and a highlight article on Intel website.

In 2020, we have also demonstrated that some of our solutions can also be applied to problems in computational genomics. An example of this is our recent work on Accel-Align, a new sequencer aligner that we developed for the task of mapping billions of reads in DNA storage. On evaluating Accel-Align with genomic data, we demonstrated that the aligner can provide a 10x speedup over state-of-the-art aligners at comparable accuracy. This makes Accel-Align a cost-effective tool for NGS-based precision medicine platforms.

Team: Raja Appuswamy, Eddy Ghabach (from October 2020), Yiqing Yan (PhD), Eugenio Marinelli (Masters)
Funding: EU FET OligoArchive
References:

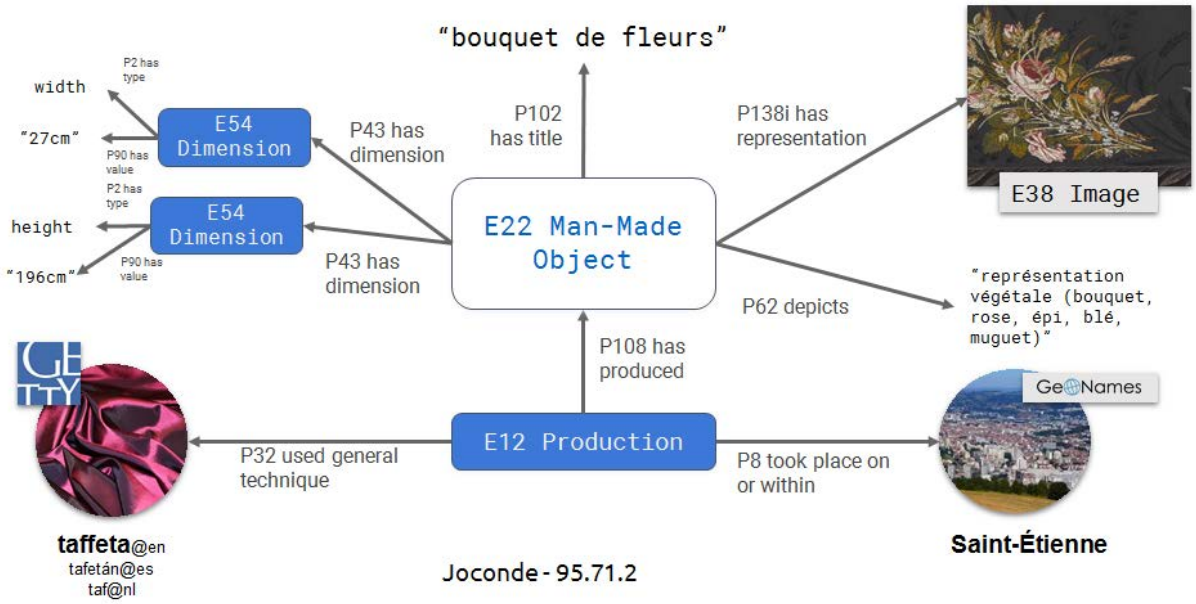
- Marinelli, Eugenio, Appuswamy, Raja: OneJoin: Cross-Architecture, Scalable Edit Similarity Join for DNA Data Storage Using oneAPI, Under submission
- Yan, Yiqing, Chatuvedi, Nimisha, Appuswamy, Raja: Accel-Align: A Fast Sequence Mapper and Aligner based on the Seed–Embed–Extend Method, Under submission
- Appuswamy, Raja; Joguin, Vincent.: Universal Layout Emulation for Long-Term Database Archival. CIDR 2021
- Melpomeni, Dimopoulou; Antonini, Marc; Barbry, Pascal; Appuswamy, Raja. Storing digital data into DNA: A comparative study of Quaternary code construction. EUSIPCO 2020

- Anadiotis, Angelos., Appuswamy, Raja, Ailamaki, Anastasia, Bronshtein, Ilya, Avni, Hillel, Dominguez-Sal, David, Goikhman, Shay, Levy, Eliezer: A system design for elastically scaling transaction processing engines in virtualized servers, VLDB 2020

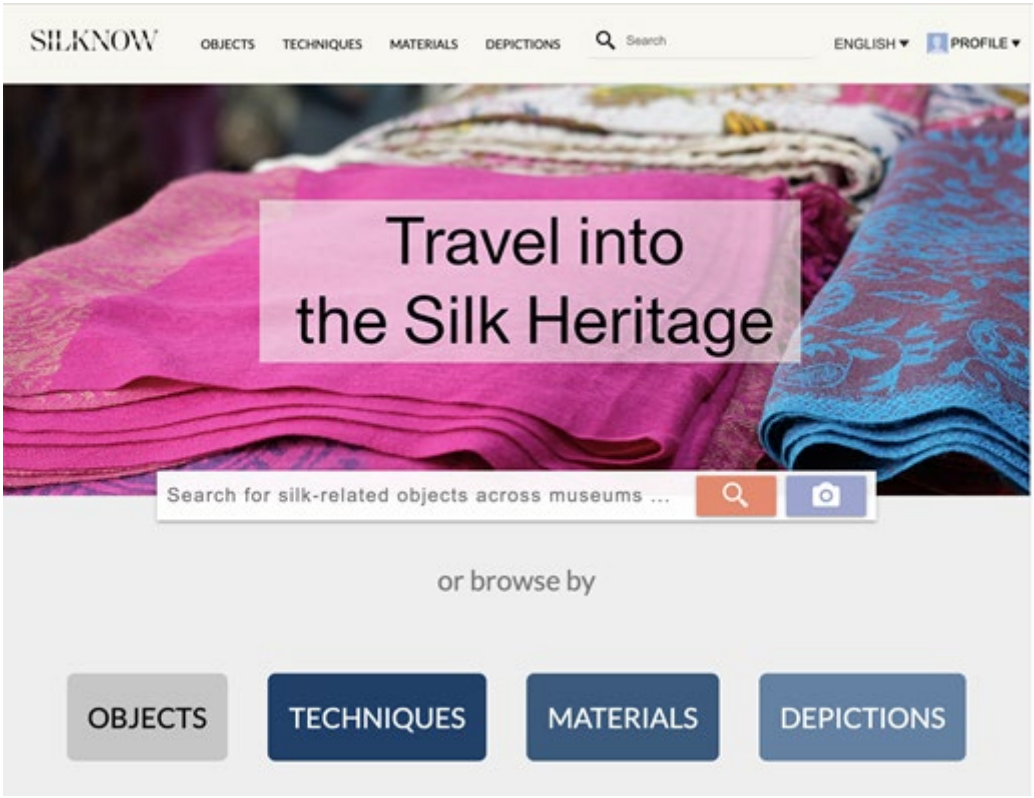
2.2 Building and Accessing Knowledge Graphs

Building a Knowledge Graph for the Silk Cultural Heritage

Summary: SILKNOW is a research project that aims at improving the understanding, conservation and dissemination of European silk heritage from the 15th to the 19th century. In this work, we designed and developed the SILKNOW knowledge graph that uses Semantic Web technologies. The underlying data model is based on CIDOC-CRM and data mappings which are realised and implemented with conversion tools developed for SILKNOW. The full integration pipeline consists also of our own crawling software to retrieve the original data from both public sources and project partners. Finally, we demonstrate the usefulness of the knowledge graph for generating labeled data that are used to train supervised algorithms on images and texts that predict missing metadata in the knowledge graph.



We also designed and developed the ADASilk exploratory search engine, named after Ada Lovelace (1815-1852), the mathematician who anticipated some of the main features of modern computing some 100 years before its advent. In her notes, she wrote that such a computation machine weaves algebraic patterns, just as the Jacquard loom weaves flowers and leaves. The application is available online at https://ada.silknow.org.

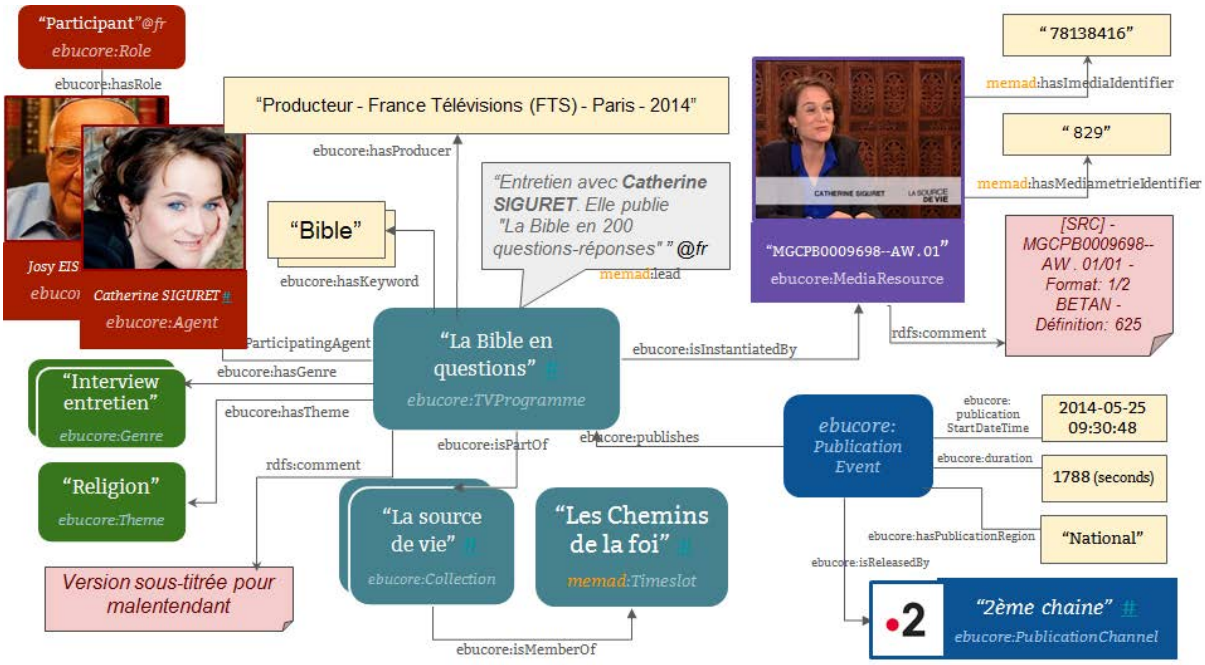


Team: Thomas Schleider, Thibault Erhart, Pasquale Lisena, Raphael Troncy
Funding: H2020 "SILKNOW"

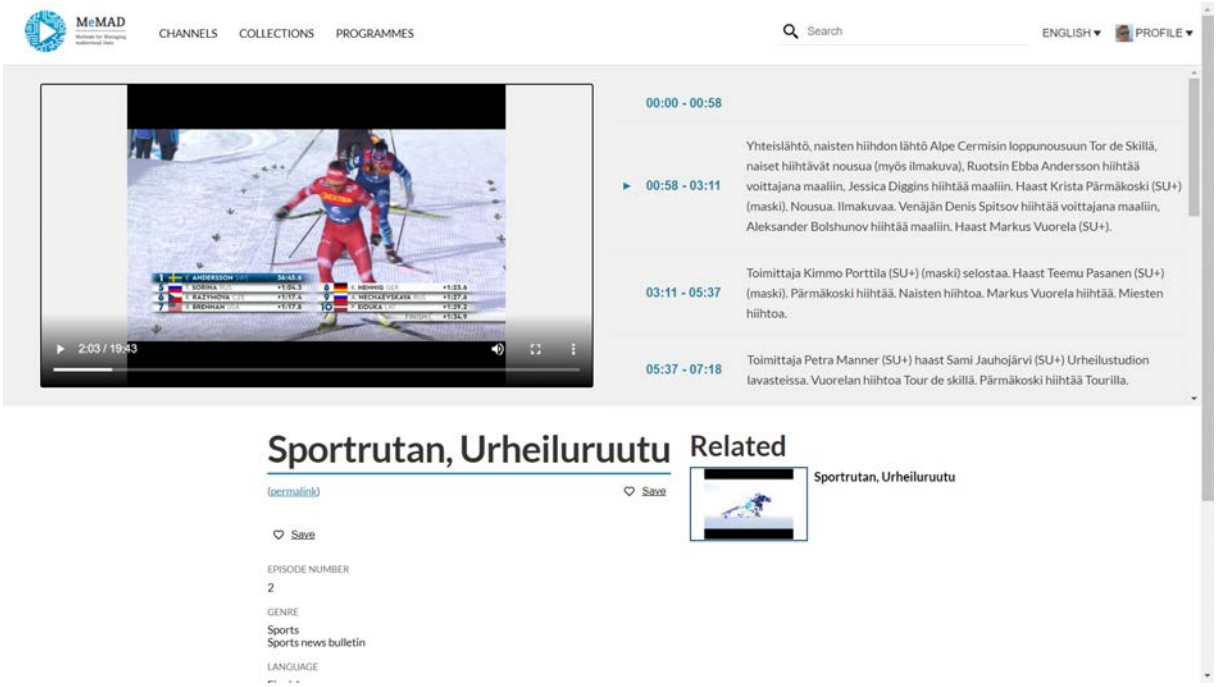
Building a Knowledge Graph for Describing TV and Radio Programmes

Summary: Multimedia systems typically contain digital documents of mixed media types, which are indexed on the basis of strongly divergent metadata standards. This severely hampers the inter-operation of such systems. Therefore, machine understanding of metadata coming from different applications is a basic requirement for the inter-operation of distributed multimedia systems. Furthermore, the content will be processed by automatic multimedia analysis tools which have their own formats for exchanging their results. One of the main goals of MeMAD is to enrich seed video content with additional content that come from diverse sources including broadcast archives, web media, news and photo stock agencies or social networks.

The general methodology that we follow consists in: i) semantifying the legacy metadata coming with audiovisual content (program metadata coming from the producer, the broadcaster and/or the archive) and ii) automatically extracting concepts and entities from the true subtitles or the text generated by automatic speech recognition on the audiovisual content. The resulting knowledge graph can then be used to infer additional information in order to enrich and hyperlink key video content moments.



We designed and developed an exploratory search engine enabling users to search and browse large collections of audiovisual material using both legacy metadata and automatically generated metadata based on multimodal content analysis. The exploratory search engine will soon integrate an explainable content recommendation algorithm.

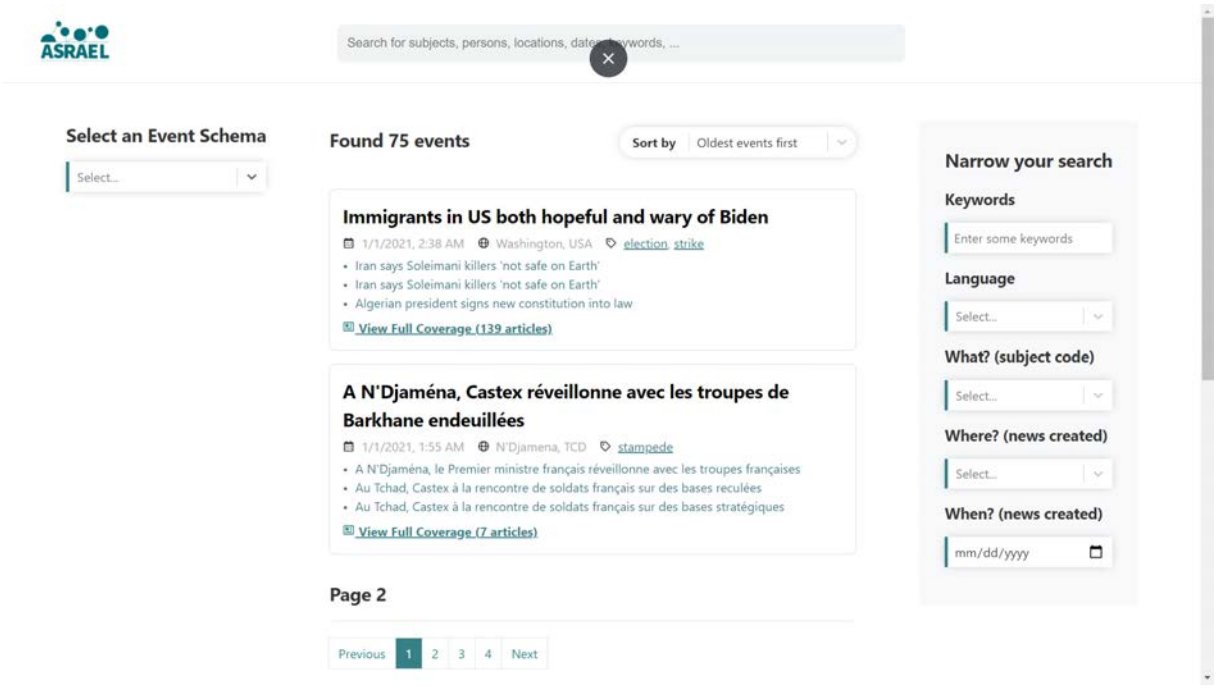


Team: Ismail Harrando, Alison Reboud, Raphael Troncy
Funding: H2020 "MeMAD"

Building a Knowledge Graph for News Organizations

Summary: News agencies produce thousands of multimedia stories describing events happening in the world that are either scheduled such as sports competitions, political summits and elections, or breaking events such as military conflicts, terrorist attacks, natural disasters, etc. When writing up those stories, journalists refer to contextual background and to compare with past similar events. However, searching for precise facts described in stories is hard. In this work, we propose a general

method that leverages the Wikidata knowledge base to produce semantic annotations of news articles. Next, we describe a semantic search engine that supports both keyword based search in news articles and structured data search providing filters for properties belonging to specific event schemas that are automatically inferred. The search engine is available at http://asrael.eurecom.fr/search-engine-beta/



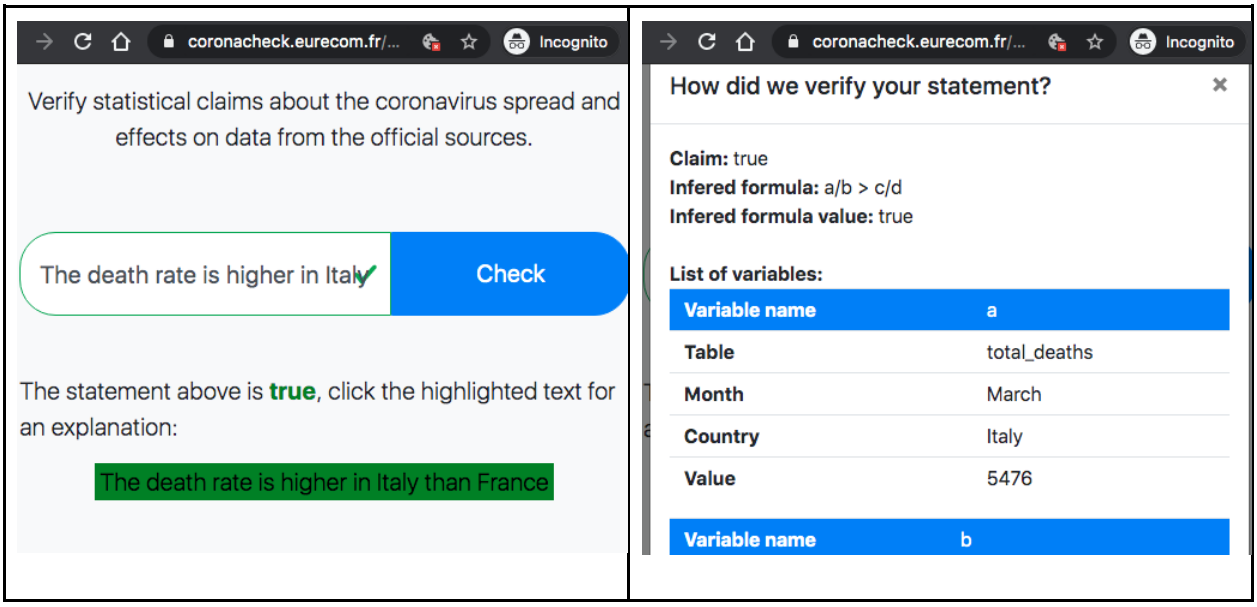
Team: Thibault Ehrhart, Raphael Troncy
Funding: ANR "ASRAEL"

2.3 Misinformation and Information Quality

Large-Scale, Data-Driven Claim Verification

Summary: Organizations spend significant amounts of time and money to manually fact check text documents summarizing data. The goal of our system (Scrutinizer) is to reduce verification overheads by supporting human fact checkers in translating text claims into SQL queries on an database. Scrutinizer coordinates teams of human fact checkers. It reduces verification time by proposing queries or query fragments to the users. Those proposals are based on claim text classifiers, that gradually improve during the verification of a large document. In addition, Scrutinizer uses tentative execution of query candidates to narrow down the set of alternatives.

The verification process is controlled by a cost-based optimizer. It optimizes the interaction with users and prioritizes claim verifications. For the latter, it considers expected verification overheads as well as the expected claim utility as training samples for the classifiers. We evaluate the Scrutinizer system using simulations and a user study with professional fact checkers, based on actual claims and data. Our experiments consistently demonstrate significant savings in verification time, without reducing result accuracy.



One instance of the system is available online at https://coronacheck.eurecom.fr for checking Coronavirus disease (COVID-19) textual claims since March 2020. The website has attracted more than 15k users and has been covered in national newspapers (https://coronacheck.eurecom.fr/en/press). We report in the figure above screenshots of the system for the checking of a single claim (left) and its explanation (right) generated as output based on the evidence in the data. A video presentation is available at https://www.youtube.com/watch?v=jzHU8CDydbc

Team: Mohammed Saeed, Paolo Papotti
Funding: ANR "InfClean"
References:

- Georgios Karagiannis, Mohammed Saeed, Paolo Papotti, Immanuel Trummer: Scrutinizer: Fact Checking Statistical Claims. Proc. VLDB Endow. 13(12): 2965-2968 (2020)
- Georgios Karagiannis, Mohammed Saeed, Paolo Papotti, Immanuel Trummer: Scrutinizer: A Mixed-Initiative Approach to Large-Scale, Data-Driven Claim Verification. Proc. VLDB Endow. 13(11): 2508-2521 (2020)
- Georgios Karagiannis, Mohammed Saeed, Paolo Papotti and Immanuel Trummer. Scrutinizer: A System for Checking Statistical Claims. BDA (2020) [Best demo award]

RuleHub: A Public Corpus of Rules for Knowledge Graphs

Summary: Entity-centric knowledge graphs (KGs) are now popular to collect facts about entities. KGs have rich schemas with a large number of different types and predicates to describe the entities and their relationships. On these rich schemas, logical rules are used to represent dependencies between the data elements. While rules are useful in query answering, data curation, and other tasks, they usually do not come with the KGs. Such rules have to be manually defined or discovered with the help of rule mining methods. We believe this rule-collection task should be done collectively to better capitalize our understanding of the data and to avoid redundant work conducted on the same KGs. For this reason, we have created RuleHub, our extensible corpus of rules for public KGs. RuleHub provides functionalities for the archival and the retrieval of rules to all users, with an extensible architecture that does not constrain the KG or the type of rules supported. We are populating the corpus with thousands of rules from the most popular KGs and report on our experiments on automatically characterizing the quality of a rule with statistical measures.

The corpus is available at https://rudik.eurecom.fr/

Team: Naser Ahmadi, Thi-Thuy-Duyen Truong, Le-Hong-Mai Dao, Paolo Papotti
Funding: ANR "InfClean"
References:

- Naser Ahmadi, Viet-Phi Huynh, Venkata Vamsikrishna Meduri, Stefano Ortona, Paolo Papotti: Mining Expressive Rules in Knowledge Graphs. ACM J. Data Inf. Qual. 12(2): 8:1-8:27 (2020)

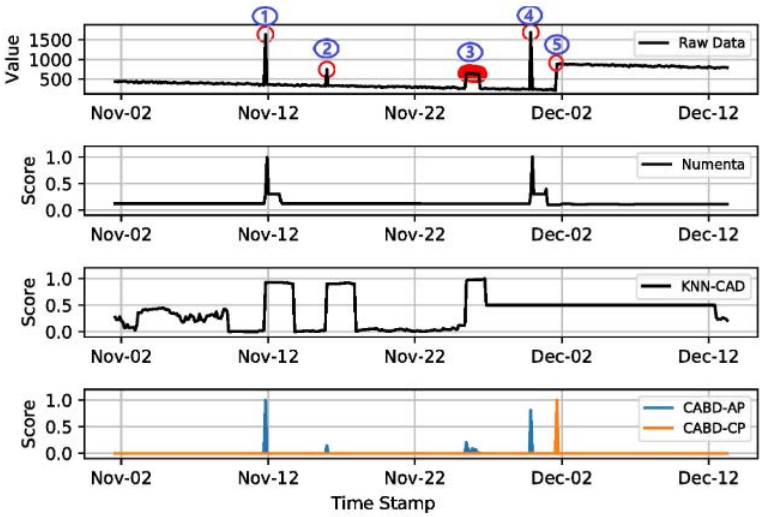
- Naser Ahmadi, Thi-Thuy-Duyen Truong, Le-Hong-Mai Dao, Stefano Ortona, Paolo Papotti: RuleHub: A Public Corpus of Rules for Knowledge Graphs. ACM J. Data Inf. Qual. 12(4): 21:1-21:22 (2020)

User-driven Error Detection for Time Series with Events

Summary: Anomalies are pervasive in time series data, such as sensor readings. Anomaly detection is an important task in several domains, such as intrusion detection systems, financial fraud detection, and Internet of Thing (IoT). It has been estimated that such data has from 2.3% to 26.9% error rate. Applications built upon imprecise time series can potentially result in losses in the millions of dollars to businesses.

Existing methods for anomaly detection cannot distinguish between anomalies that represent data errors, such as incorrect sensor readings, and notable events, such as the watering action in soil monitoring. The first plot from the top in the figure visually presents this problem in real ultrasonic sensor data obtained from an IoT company. The sensor is plugged on the top of a tank to monitor its liquid level (y axis) over time (x axis). As shown in the figure, sudden changes appear in isolation (e.g., 08-November) and as small groups (24 to 26-November). These abnormal values are sensor errors, and should be fixed or removed from the data set. On the other hand, the data change reflecting the filling of the tank (30-November) should be preserved. Anomaly detection results for three algorithms are reported also in the figure. Points labeled as 1-4 are anomalies, while point 5 represents a water filling event. We can observe that only our method, the one at the bottom, is able to recognize and distinguish anomalies and events.

In addition, the quality performance of most existing detection methods highly depends on the configuration parameters, which are dataset specific. In this work, we exploit active learning to detect both errors and events in a single solution that aims at minimizing user interaction. For this joint detection, we introduce an algorithm that accurately detects and labels anomalies with a non-parametric concept of neighborhood and probabilistic classification. Given a desired quality, the confidence of the classification is then used as termination condition for the active learning algorithm. Experiments on real and synthetic datasets demonstrate that our approach achieves F-score above 80% in detecting errors by labeling 2 to 5 points in one data series. We also show the superiority of our solution compared to the state-of-the-art approaches for anomaly detection. Finally, we demonstrate the positive impact of our error detection methods in downstream data repairing algorithms.



Team: Kim-Hung Le, Paolo Papotti
Funding: CIFRE contract with Greencityzen (Marseille)
References:

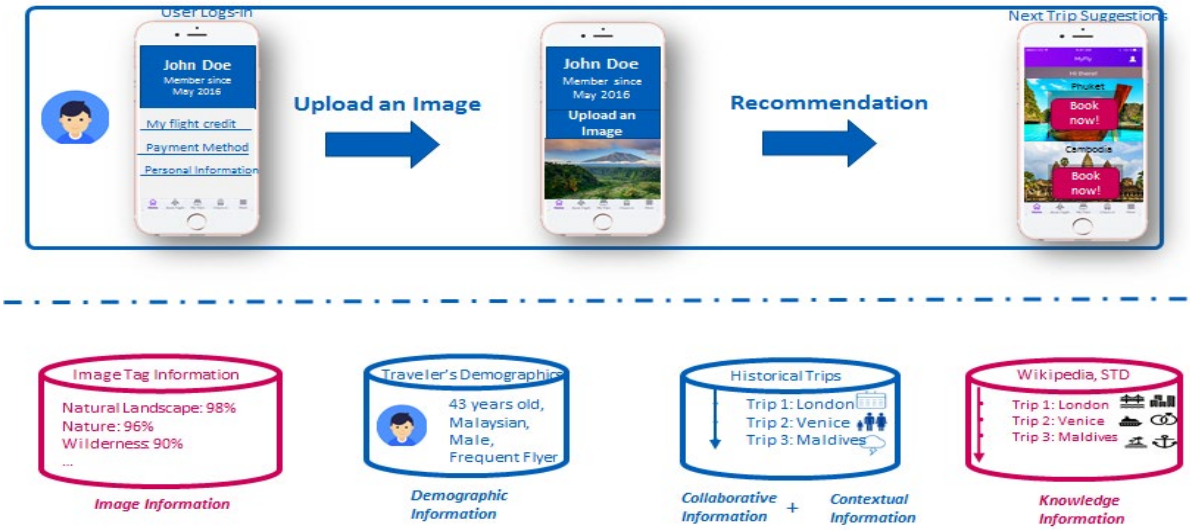
- Kim-Hung Le, Paolo Papotti: User-driven Error Detection for Time Series with Events. ICDE 2020: 745-757

2.4 Recommender Systems

Location Embeddigns for Next Trip Recommendation

Summary: The amount of information available in social media and specialized blogs has become useful for a user to plan a trip. However, the user is quickly overwhelmed by the list of possibilities offered to him, making his search complex and time-consuming. Recommender systems aim to provide personalized suggestions to users by leveraging different type of information, thus assisting them in their decision-making process. Recently, the use of neural networks and knowledge

graphs have proven to be efficient for items recommendation. In our work, we propose an approach that leverages contextual, collaborative and content information in order to recommend personalized destinations to travelers. We compare our approach with a set of state of the art collaborative filtering methods and deep learning based recommender systems.



Two Stages Approach for Tweet Engagement Prediction

Summary: We also participate in the 2020 RecSys Challenge on the task of predicting user engagement facing tweets. This approach relies on two distinct stages. First, relevant features are learned from the challenge dataset. These features are heterogeneous and are the results of different learning modules such as handcrafted features, knowledge graph embeddings, sentiment analysis features and BERT word embeddings. Second, these features are provided in input to an ensemble system based on XGBoost. This approach, only trained on a subset of the entire challenge dataset, ranked 22 in the final leaderboard.

Gitlab: https://gitlab.eurecom.fr/dadoun/RecSys_Challenge_2020

Team: Amine Dadoun, Raphael Troncy
Funding: CIFRE Amadeus

2.5 Multimodal Information Extraction

TOMODAPI: A Topic Modeling API to Train, Use and Compare Topic Models

Summary: From LDA to neural models, different topic modeling approaches have been proposed in the literature. However, their suitability and performance is not easy to compare, particularly when the algorithms are being used in the wild on heterogeneous datasets. In this work, we introduce ToModAPI (TOpic MOdeling API), a wrapper library to easily train, evaluate and infer using different topic modeling algorithms through a unified interface. The library is extensible and can be used in Python environments or through a Web API.

The library is open sourced at https://github.com/D2KLab/ToModAPI

Team: Pasquale Lisena, Raphael Troncy
Funding: ANR "ASRAEL"

DAGOBAH: a Tool for Semantically Interpreting and Annotating Tabular Data

Summary: We present new approaches used in the DAGOBAH system to perform automatic semantic table interpretation. DAGOBAH semantically annotates tables with Wikidata entities and relations to perform three tasks: Columns-Property

Annotation (CPA), Cell-Entity Annotation (CEA) and Column-Type Annotation (CTA). In our system, the initial scores from entity disambiguation influence the CPA output, which, in turn, influences the output of the CEA. Finally, the CTA is computed using the type hierarchy available in the knowledge graph in order to annotate columns with the most suitable fine-grained types. This approach that leverages mutual influences between annotations allows DAGOBAH to obtain very competitive results on all tasks of the SemTab2020 challenge.

Team: Jixiong Liu, Raphael Troncy
Funding: CIFRE Orange

Predicting Media Interestingness or Memorability

Summary: We propose a multimodal approach for the MediaEval 2020 Predicting Media memorability task. Our best approach is a weighted average method combining predictions made separately from visual, audio, textual and visiolinguistic representations of videos. Our best model achieves Spearman scores of 0.101 and 0.078, respectively, for the short and long term predictions tasks.

Github: https://github.com/MeMAD-project/media-memorability

Team: Alison Reboud, Ismail Harrando, Raphael Troncy
Funding: H2020 "MeMAD"

Detecting and Tracking Faces of Celebrities in TV Content

Summary: We further researched and developed our generic face recognition method to detect and recognise not only famous people within a TV corpus, such as the one provided in the MeMAD project or the "Actualitée Francaise" corpus from the ANTRACT project, but also frequently recurring persons even if we do not name them. We evaluated our approach on a number of celebrities such as Charle De Gaulle, The Queen Elizabeth II, Brigitte Bardot and Jean-Paul Belmondo. The average precision is 67% while the recall is up to 93% on the MeMAD corpus with 15 selected celebrities.

The algorithm can also detect and cluster the regular occurrence of a recurrent person. Once tagged, using our interactive tool available at http://facerec.eurecom.fr/visualizer/, the algorithm will similarly detect and recognize all occurrences of this person.

Github: https://github.com/D2KLab/Face-Celebrity-Recognition

Team: Pasquale Lisena, Raphael Troncy
Funding: ANR "ANTRACT", H2020 MeMAD

Automatically Generating Summaries of TV Series

Summary: We propose a fan-driven and character-centered approach for the 2020 TRECVID Video Summarization Task. In terms of data, besides the provided videos, scripts and master shot boundaries, our approach relies on fan-made content, more precisely on the BBC EastEnders episode synopsis from its Fandom Wiki. We also use BBC EastEnders characters' images crawled from a search engine to train a face recognition system. All our runs use the same method, but with varying constraints of the number of shots and the maximum duration. The shots included in the summaries are the ones whose transcripts and visual content have the highest similarity with sentences from the synopsis. Overall, we consider that the results obtained speak in favour of using fan-made content as a starting point for such a task. As we did not try to optimize for tempo and contextuality, we believe there is some margin for improvement here, however the task of answering unknown questions remains challenging.

Github: https://github.com/MeMAD-project/trecvid-vsum

Team: Alison Reboud, Ismail Harrando, Raphael Troncy
Funding: H2020 "MeMAD"

3. PUBLICATIONS

3.1 Book and Edited Proceedings

[1] Al-Moslmi, Tareq; Troncy, Raphaël; Freitas, André; Ceolin, Davide; Abolohom, Abdullatif: SKG4J 2020: 1st international workshop on semantic and knowledge Graph advances for journalism. In CIKM 2020, 29th ACM International Conference on Information and Knowledge Management, 19-23 October 2020

[2] Troncy, Raphaël; Laaksonen, Jorma T; Tavakoli, Hamed R; Nixon, Lyndon; Mezaris, Vasileios; Hosseini, Mohammad: AI4TV 2020: 2nd International Workshop on AI for smart TV content production, access and delivery. In MM 2020, 28th ACM International Conference on Multimedia, 12-16 October 2020, Seattle, USA

[3] Harth, Andreas; Presutti, Valentina; Troncy, Raphaël; Acosta, Maribel; Polleres, Axel; Fernández, Javier D.; Parreira, Josiane Xavier; Hartig, Olaf; Hose, Katja; Cochez, Michael: The Semantic Web: ESWC 2020 Satellite Events, May 31-June 4, 2020, Heraklion, Crete, Greece. Lecture Notes in Computer Science book series (LNCS, volume 12124)

[4] Martel, Anne L; Abolmaesumi, Purang; Stoyanov, Danail; Mateus, Diana; Zuluaga, Maria A; Zhou, S. Kevin; Racoceanu, Daniel; Joskowicz, Leo: Medical image computing and computer assisted intervention. In MICCAI 2020, 23rd International Conference on Medical Image Computing and Computer Assisted Intervention, 4-8 October 2020, LNCS 12261- 12267 Lima, Peru

3.2 Conference and Workshops

[5] Lisena, Pasquale; Troncy, Raphaël: Representing complex knowledge for exploration and recommendation: The case of classical music information. In "Applications and Practices in Ontology Design, Extraction, and Reasoning". Studies on the Semantic Web Series (SSWS), Volume 49, IOSPress, December 2020

[6] Nguyen, Phuong Anh; Wu, Jiaxin; Ngo, Chong-Wah; Francis, Danny; Huet, Benoit: VIREO @ Video browser showdown 2020. In 26th International Conference on MultiMedia Modeling, 5-8 January 2020, Daejon, Korea.

[7] Le, Kim-Hung; Papotti, Paolo: User-driven error detection for time series with events. In 36th IEEE International Conference on Data Engineering (ICDE), 20-24 April, Dallas, Texas, USA

[8] Dimopoulou, Melpomeni; Antonini, Marc; Barbry, Pascal; Appuswamy, Raja: Storing digital data into DNA: A comparative study of quaternary code construction. In 45th International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 4-8 May 2020, Barcelona, Spain

[9] Ohana, Ruben; Wacker, Jonas; Dong, Jonathan; Marmin, Sébastien; Krzakala, Florent; Filippone, Maurizio; Daude, Laurent: Kernel computations from large-scale random features obtained by optical processing units. In IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4-8 May 2020, Barcelona, Spain

[10] Kisamori, Keiichi; Kanagawa, Motonobu; Yamazaki, Keisuke: Simulator calibration under covariate shift with Kernels. In AISTATS 2020, 23rd International Conference on Artificial Intelligence and Statistics, June 3-5, 2020, Palermo, Sicily, Italy

[11] Mita, Graziano; Papotti, Paolo; Filippone, Maurizio; Michiardi, Pietro: LIBRE: Learning interpretable boolean rule ensembles. In AISTATS 2020, 23rd International Conference on Artificial Intelligence and Statistics, 3-5 June 2020, Palermo, Sicily, Italy

[12] Cappuzzo, Riccardo; Papotti, Paolo; Thirumuruganathan, Saravanan: Creating embeddings of heterogeneous relational datasets for data integration tasks. In SIGMOD/PODS 2020, ACM International Conference on Management of Data, 14-19 June 2020, Portland, Oregon, USA

[13] Chabot, Yoan; Labbé, Thomas; Liu, Jixiong; Troncy, Raphaël: DAGOBAH : Un système d'annotation sémantique de données tabulaires indépendant du contexte. In IC 2020, 31èmes Journées francophones d'Ingénierie des Connaissances, 30 June 2020-3 July, Angers, France

[14] Audibert, Julien; Michiardi, Pietro; Guyard, Frederic; Marti, Sebastien; Zuluaga, Maria A.: USAD: Unsupervised anomaly detection on multivariate time series. In KDD 2020, 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, August 23-27, 2020, San Diego, USA

[15] Anadiotis, Angelos-Christos; Appuswamy, Raja; Ailamakiy, Anastasia; Bronshtein, Ilan; Avni, Hillel; Dominguez-Sal, David; Goikhman, Shay; Levy, Eliezer: A system design for elastically scaling transaction processing engines in virtualized servers. In VLDB 2020, 46th International Conference on Very Large Data Bases, 31 August-4 September 2020, Tokyo, Japan

[16] Karagiannis, Georgios; Saeed, Mohammed; Papotti, Paolo; Trummer, Immanuel: Scrutinizer: A mixed-initiative approach to large-scale, data-driven claim verification. In VLDB 2020, 46th International Conference on Very Large Data Bases, 31 August-4 September 2020, Tokyo, Japan

[17] Candela, Rosa; Michiardi, Pietro; Filippone, Maurizio; Zuluaga, Maria A: Model monitoring and dynamic model selection in travel time-series forecasting. In ECML-PKDD 2020, The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, 14-18 September 2020, Ghent, Belgium

[18] Rossi, Simone; Marmin, Sébastien; Filippone, Maurizio: Supplement material for "Walsh-Hadamard Variational Inference for Bayesian Deep Learning". In NeurIPS 2020, 34th Conference on Neural Information Processing Systems, 6-12 December 2020

[19] Huynh, Viet-Phi; Liu, Jixiong; Chabot, Yoan; Labbé, Thomas; Monnin, Pierre; Troncy, Raphaël: DAGOBAH: Enhanced scoring algorithms for scalable annotations of tabular data. In SEMTAB 2020, Semantic Web Challenge on Tabular Data to Knowledge Graph Matching (SemTab 2020), co-located with the 19th International Semantic Web Conference (ISWC 2020), 5 November 2020, Athens, Greece

[20] Lisena, Pasquale; Harrando, Ismail; Kandakji, Oussama; Troncy, Raphaël: ToModAPI: A topic modeling API to train, use and compare topic models. In NLP-OSS 2020, 2nd Workshop for Natural Language Processing Open Source Software, 19 November 2020

[21] Rossi, Simone; Marmin, Sébastien; Filippone, Maurizio: Walsh-hadamard variational inference for Bayesian deep learning. In NeurIPS 2020, 34th Conference on Neural Information Processing Systems, 6-12 December 2020, Vancouver, Canada

[22] Reboud, Alison; Harrando, Ismail; Laaksonen, Jorma; Troncy, Raphaël: Predicting media memorability with audio, video, and text representations. In MEDIAEVAL 2020, Multimedia Evaluation Benchmark, 14-15 December 2020

[23] Chatterjee, Sujoy; Pasquier, Nicolas; Nanty, Simon; Zuluaga, Maria A: Multi-objective consensus clustering framework for flight search recommendation. In Lecture Notes in Networks and Systems, Vol 141, Springer

[24] Harrando, Ismail; Reboud, Alison; Lisena, Pasquale; Troncy, Raphaël; Laaksonen, Jorma; Virkkunen, Anja; Kurimo, Mikko: Using fan-made content, subtitles and face recognition for character-centric video summarization. In TRECVID 2020, November 2020

[25] Karagiannis, Georgios; Saeed, Mohammed; Papotti, Paolo; Trummer, Immanuel: Scrutinizer: Fact checking statistical claims. In VLDB 2020, 46th International Conference on Very Large Data Bases, 31 August-4 September 2020, Tokyo, Japan

[26] Pascal, Lucas; Huet, Benoit; Bost, Xavier; Zuluaga, Maria A.: Maximum roaming multi-task learning for glaucoma detection from color fundus photographs. In REFUGE2 2020, 2nd Retinal Fundus Glaucoma Challenge, in Conjunction with MICCAI 2020, 8 October 2020, Lima, Peru

3.3 Journal

[27] Palumbo, Enrico; Monti, Diego; Rizzo, Giuseppe; Troncy, Raphaël; Baralis, Elena: entity2rec: Property-specific knowledge graph embeddings for item recommendation. In Expert Systems with Applications, January 2020.

[28] Ahmadi, Naser; Huynh, Viet-Phi; Meduri, Vamsi; Ortona, Stefano; Papotti, Paolo: Mining expressive rules in knowledge graphs. In ACM Journal of Data and Information Quality (JDIQ), Vol. 1, N°1, May 2020

[29] Ahmadi, Naser; Truong, Thi-Thuy-Duyen; Dao, Le-Hong-Mai; Ortona, Stefano; Papotti, Paolo: RuleHub: a public corpus of rules for knowledge graphs. In Journal of Data and Information Quality, Special issue " Metadata Discovery for Assessing Data Quality", July 2020

[30] Ma, Da; Cardoso, Manuel J.; Zuluaga, Maria A.; Modat, Marc; Powell, Nick M.; Wiseman, Frances K.; Cleary, Jon O.; Sinclair, Benjamin; Harrison, Ian F.; Siow, Bernard; Popuri, Karteek; Lee, Sieun; Matsubara, Joanne A.; Sarunic, Marinko V.; Beg, Mirza Faisal; Tybulewicz, Victor L.J.; Fisher, Elizabeth M.C.; Lythgoe, Mark F.; Ourselin, Sébastien: Substantially thinner internal granular layer and reduced molecular layer surface in the cerebellum of the Tc1 mouse model of Down Syndrome – a comprehensive morphometric analysis with active staining contrast-enhanced MRI. In NeuroImage, N°223, August 2020

[31] Pérez, M. Casanova-Salas, P; Twardo, P.; León, A.; Mladenic, D; Massri, B.M.; Troncy, Raphaël; Ehrhart, Thibault; Lo Cicero, G.; Vitella, M.; Gaitán, M.; Gimeno, J.; Ribes, E.; Fernández, M; Portalés, C: From historical silk fabrics to their interactive virtual representation and 3D printing. In Sustainability, Vol.12, Issue 18, September 2020

[32] Roy, Senjuti Basu; Chen, Lei; Morishima, Atsuyuki; Monedero, James; Bourhis, Pierre; Charoy, François; Danilevsky, Marina; Das, Gautam; Demartini, Gianluca; Dubey, Abishek; Elbassuoni, Shady; Gross-Amblard, David; Hoareau, Emilie; Inoguchi, Munenari; Kenworthy, Jared; Kitahara, Itaru; Lee, Dongwon; Li, Yunyao; Borromeo, Ria Mae; Papotti, Paolo; Rao, Raghav; Roy, Sudeepa; Senellart, Pierre; Tajima, Keishi; Thirumuruganathan, Saravanan; Tommasi, Marion; Umemoto, Kazutoshi; Wiggins, Andrea; Yoshida, Koichiro; Amer-Yahia, Sihem: Making AI machines work for humans in FoW. In ACM SIGMOD Record, Vol. 49, N°2, December 2020

3.4 HDR, PhD Thesis and Technical Reports

[33] Palumbo, Enrico: Knowledge graph embeddings for recommender systems. PhD Thesis, Sorbonne Université, 2020

[34] Vu, Quan-Dong: Models and solutions of strategic resource allocation problems: Approximate equilibrium and online learning in Blotto games. PhD Thesis, Sorbonne Université, 2020

[35] Tran, Gia-Lac: Advances of deep Gaussian processes: Calibration and sparsification. PhD Thesis, Sorbonne Université, 2020

[36] Muandet, Krikamol; Kanagawa, Motonobu; Saengkyongam, Sorawit; Marukatat, Sanparith: Counterfactual mean embeddings. Submitted to ArXiV, 1st March 2020

[37] Rossi, Simone; Heinonen, Markus; Bonilla, Edwin; Shen, Zheyang; Filippone, Maurizio: Rethinking sparse Gaussian processes: Bayesian approaches to inducing-variable approximations. Submitted to ArXiv, 9 March 2020

[38] Milios, Dimitrios: Michiardi, Pietro; Filippone, Maurizio: A variational view on bootstrap ensembles as Bayesian inference. Submitted to ArXiV, 8 June 2020

[39] Franzese, Giulio; Candela, Rosa; Milios, Dimitrios; Filippone, Maurizio; Michiardi, Pietro: Isotropic SGD: a practical approach to Bayesian posterior sampling. Submitted on ArXiV, 9 June 2020

[40] Dadoun, Amin; Troncy, Raphaël: Many-to-one recurrent neural network for session-based recommendation. In Submitted on ArXiV, 25 August 2020

[41] Tran, Gia-Lac; Dimitrios, Milios; Michiardi, Pietro; Filippone, Maurizio: Sparse within sparse gaussian processes using neighbor information. Submitted to ArXiV, 12 November 2020

[42] Tran, Ba-Hien; Rossi, Simone; Milios, Dimitrios; Filippone, Maurizio: All you need is a good functional prior for Bayesian deep learning. Submitted to ArXiV, 25 November 2020

[43] Dadoun, Amin; Harrando, Ismail; Lisena, Pasquale; Reboud, Alison; Troncy, Raphaël: Two stages approach for tweet engagement prediction. In Technical Report, CoRR abs/2008.10419, 24 August 2020

3.5 Invited Talks

[44] Karagiannis, Georgios; Saeed, Mohammed; Papotti, Paolo; Trummer, Immanuel: Scrutinizer: Fact checking statistical claims. In BDA 2020, 36eme Conférence sur la Gestion de Données, Principes, Technologies et Applications, 27-29 Octobre 2020, France

[45] Kanagawa, Motonobu: Simulator calibration under covariate shift with Kernels. In FIMI 2020, Workshop on Functional Inference and Machine Intelligence, 17-19 February 2020, Sophia Antipolis, France

[46] Filippone, Maurizio: Random feature expansions for deep Gaussian processes. In FIMI 2020, Workshop on Functional Inference and Machine Intelligence, 17-19 February 2020, Sophia Antipolis, France

[47] Papotti, Paolo: Explainable checking for statistical and property claims. In Webinar "Data analytics & AI", 28 May 2020, IMT, France (https://www.imt-bs.eu/webinar-data-analytics-ai/)

[48] Papotti, Paolo: Explainable fact checking for statistical and property claims. In Data analytics Seminar, 15 July 2020, Qatar Computing Research Institute, Qatar

[49] Papotti, Paolo: Explainable fact checking for statistical and property claims. In DEXA 2020, Keynote speech at 31st International Conference on Database and Expert Systems Applications, 14-17 September 2020, Bratislavia, Slovakia

[50] Papotti, Paolo: Explainable fact checking for statistical and property claims. In Invited Talk, 15 October 2020, Inria-Ecole Polytechnique CEDAR Seminar, France

[51] Papotti, Paolo; Trummer, Immanuel: Fact checking statistical claims. In TTO 2020, Conference for Truth and Trust Online, 16-17 October 2020

[52] Appuswamy, Raja; Marinelli, Eugenio: OneOligo: Use oneAPI to accelerate DNA data storage. Invited talk

4. THE DATA SCIENCE DEPARTMENT TEAM

Professors (8):

- Pietro Michiardi
- Maurizio Filippone
- Bernard Merialdo (Professor Emeritus)
- Raja Appuswamy
- Motonobu Kanagawa
- Paolo Papotti
- Raphael Troncy
- Maria A. Zuluaga

Postdoc / Engineers (6):

- Thibault Ehrhart
- Pasquale Lisena
- Dimitrios Milios
- Giulio Franzese
- Daniele Venzano (external contractor)
- Eddy Ghabach

PhD Students (24):

- Naser Ahmadi (KPMG)
- Julien Audibert (CIFRE Orange)
- Rosa Candela (Amadeus)
- Riccardo Cappuzzo (IMT Futur&Ruptures)
- Matthieu Da Silva-Filarder (CIFRE Renault Software Labs)
- Amine Dadoun (CIFRE Amadeus)
- Ismail Harrando (H2020 MeMAD)
- Ugo Lecerf (CIFRE Renault Software Labs)
- Jixiong Liu (CIFRE Orange)
- Graziano Mita (CIFRE SAP)
- Florent Moriconi (CIFRE Amadeus)
- Lucas Pascal (CIFRE Orkis)
- Enrico Palumbo (LINKS Foundation)
- Alison Reboud (H2020 MeMAD)
- Simone Rossi (AXA Chair)
- Mohammed Saeed (ANR JCJC InfClean)
- Thomas Schleider (H2020 SILKNOW)
- Lionel Tailhardat (Orange)
- Gia-Lac Tran (AXA Chair)
- Ba-Hien Tran (AXA Chair)
- Bogdan Kozyrskiy (3IA Chair)
- Davit Gogolashvili (EU ITN Windmill)
- Jonas Wacker (ANR JCJC ECO-ML)
- Yiqing Yan (OligoArchive)

COMMUNICATION SYSTEMS DEPARTMENT

Introduction

Communication networks are facing unprecedented changes driven by both human usage (e.g. the emergence of (i) immersive VR experience) and machine usage (smart cities, ITS, industry 4.0 etc.). From an architectural perspective the mobile networking world is experiencing a convergence of internet, broadcast, internet-of-things (IoT), and classical telecommunication frameworks. The network traffic is now dominated by machines which also sometimes must operate with each other, forcing the development of robust connected machine networks (cars, robots) with applications to systems as diverse as intelligent transport systems and the factory of the future.

In order to provide the expected services to a growing population of machine users (alongside human driven traffic), the redesign of communications strategies, radio spectrum sharing policies, QoS control modes and eventually the overall architecture of mobile networks must be carried out, in the context of beyond 5G evolutions (6G).

The primary research goal of the Communication Systems department at EURECOM is two fold: First to make advances at the fundamental and applied levels in domains ranging from information and communication theory, signal processing, graph theory and learning for networks, and secondly turn concepts into practical system-oriented innovations by leveraging the open-source experimental platforms developed by our department.

Consequently, research in the department is organized along the activities of the two research group below:

- Foundation & Algorithms: We develop or exploit solid scientific tools for the analysis of communications systems and the enhancement of their performance at multiple layers. We push the fundamental limits of communications over networks with novel transmissions and information decoding strategies. We develop concepts for 6G networks including ideas such as ultra-directional (massive MIMO, smart surfaces) communication, cache-aided (and coded-chaching) networking, connected robotics, robust multi-agent decision networks.

- Networked Systems: We bring life to our ideas by real-life experimentation and prototyping. We develop and implement communication schemes and techniques with the goal of advancing the performance of post-4G (5G and beyond) standardized systems. We push open source 5G protocol design onto the global stage via the OpenAirInterface software alliance. We also address new challenges posed by emerging novel wireless applications (e.g. broadband multimedia and IoT) and scenarios (e.g. vehicular, radio over ITS,..) by developping efficient networking methods. The architecture of backhaul and core networks under the new prism of Mobile Edge Computing is considered and open source implementations are promoted.

1. ORGANIZATION OF THE REPORT

The research report of the Communication Systems Dept. is presented in the following parts:

- NETWORKED SYSTEMS
- RESOURCE OPTIMIZATION AND CROSS LAYER DESIGN
- LOCALIZATION
- COMMUNICATION THEORY AND TECHNIQUES
- LEARNING FOR WIRELESS NETWORKS
- EXPERIMENTAL ACTIVITIES
- COLLABORATIVE PROJECTS
- PUBLICATIONS

Note that the EURECOM Communication Systems Department is active in a large number of EU (both industrial oriented and ERC) and French (ANR) funded research projects and well as bilateral directly funded projects from several industry partners. The Dept is also present in standardization bodies: ETSI groups (Software Radio, Cognitive Systems, and Intelligent Transport Systems) and Car to Car Forum. Such activities are referred to in the report, where appropriate.

2. NETWORKED SYSTEMS

A substantial portion of the department's activities focusses on systems oriented research and aims at bringing ideas to life through experimentation.

It addresses Protocols and network architecture as well as PHY and MAC design on the wireless access.

One of the two research groups composing the department, called the Networked Systems Group directly contributes to the OPENAIRINTERFACE developments in order to setup real time and emulation experimental platforms.

It also studies and experiments Service oriented networking via MOSAIC5G platforms based on Virtualization and Software Defined Networking

Various application domains and systems are addressed by our research, amongst them:

- Public Safety Networks, Alerting Networks, Moving Cells
- LTE Advanced Networks for Broad Band Services and Machine 2 Machine Communications (NB IoT, LTE M)
- Vehicular Networks (V2X), LTE D2D
- Cooperative Intelligent Transportation Systems
- Internet of Things (IoT)

The Research Group put efforts in contributing actively to the standardization bodies: ETSI ITS, OneM2M, W3C, CAR2CAR, and IETF.

We present our research at the light of different themes that are developed in the framework of projects. Since some projects encompass several research themes they are mentioned several times.

- COOPERATIVE INTELLIGENT TRANSPORTATION SYSTEMS (C-ITS)
- 5G NETWORKING (SDN, NFV, Cloud RAN, Orchestration, …)
- INTERNET of THINGS
- EXPERIMENTAL MIMO SYSTEMS

Finally the Research Group put efforts in contributing actively to the standardization bodies ETSI ITS, C2C, W3C, oneM2M and IETF.

The group was involved in the following projects: SYMPA, CONECT, PFT, FLEX, MCN, CNES SMILE, Netcom, , 4GinVITRO, HIGHTS, direct contracts with Orange Labs, direct contracts with INTEL, direct contracts with Renault, direct contracts with Toyota, direct contracts with the C2CCC. Projects started in 2018 5G-EVE, 5GENESIS, EMPOWER, 5G-CROCO and SAFE4Rail2 are under progress.

The group keeps on developing the protocol stacks software OpenAirInterface: LTE/5G protocol stack on radio interface, in particular LTE-D2D/V2X, DMMIPv6, Non Access Stratum protocols, Cloud RAN protocols, 802.11p MAC, 802.21 mechanisms integration. The LTE Core Network has been virtualized as a main achievement of the year. This work is described in the section dedicated to the OpenAirInterface open source platform

2.1 Cooperative Intelligent Transportation Systems (C-ITS)

We see today urban traffic growing faster than the capacity of road infrastructures and posing a challenge to traffic safety and to the sustainability of our mobility. In this context, cooperative mobility strategies are expected to become a key enabler for a better usage of the available road infrastructures. Considering the estimation by the American Automotive Association of the yearly cost of traffic safety to 160 billion USD, or the yearly cost of traffic congestion to 50 billion EUR by the European Commission, local road authorities are facing a strategic challenge to develop wireless vehicular communication solutions and propose cooperative ITS applications to the public.

Research activities in the field of C-ITS have been conducted in various domains, among them this year:

- Data-Driven Dependable Vehicular Communication & Networking
- Data Driven Mobility & Traffic Modeling and Planning
- Active Standardization Contributions in C-ITS domains

The research conducted for C-ITS have been financed by National, European projects, as well as industrial contracts and contracts with the standardization bodies.

2.1.1 Data Driven Vehicular Communication and Networking

We conducted research in enhancing both WIFI and Cellular V2X technologies, and initiated activities on NR V2X. On the Cellular V2X side, in [Hae20a], we proposed a LTE V2X architecture for Wireless Train Control and Monitoring System (TCMS). In [Hae20b], we extended the OAI V2X architecture to support a stub architecture enabling MAC-level emulations, and evaluated the performance of the LTE-V2X scheduler under challenging scenarios. In [Hae20c], we surveyed communication and networking challenges for 5G V2X-based connected automated vehicles.

On the ITS-G5 side, we focused on defining advanced distributed resource managements for dependable V2X communications. In [KhaHae20a], we developed an efficient and fair resource allocation considering multiple V2X services. In [KhaHae20b] we proposed a multi-objective orchestrator for cooperative multi-service resource allocations. These activities led to a Patent jointly proposed between Renault and EURECOM [KhaHae20c].

We finally further extended the concept and architecture of Knowledge-centric Vehicular network in [DunHae20b], and applied it to data-driven hazard detection in [DunHae20c]. Finally, we demonstrated the benefit of Knowledge-centric vehicular networking on Federated Machine Learning by proposing an orchestration architecture for FML in [DunHae20d].

2.1.2 Data Driven Mobility & Traffic Modeling and Planning

We conduct research in both developing advanced realistic vehicular traffic scenarios as well as models for mixed traffic scenarios between vehicles and motorcycles. In particular, we proposed in [CodHae20a], we proposed SAGA, an automated citywide scenario generation for SUMO. In [CodHae20b], we applied Multimodal Fundamental Diagrams to analyze spatio-temporal traffic challenges in multimodal transport planning.

In [DunHae20a], we evaluated the impact of comfort-based re-routing in urban scenarios, where automated vehicles would rather avoid certain areas of the city due to the complex geometry or potential hazard. In [DigHae20], we proposed a hybrid federated Machine Learning (H-FML) mechanism to optimize federated learning of hazardous conditions under highly dynamic mobility.

2.1.3 Standardization efforts in C-ITS

We contributed to standards in ETSI ITS and the CAR 2 CAR, first to integrate into ETSI ITS standards the proposed multi-service resource orchestrations published in [KhaHae20a] and [KhaHae20b], as well as evaluating the impact of the new IEEE Next Generation Vehicular (IEEE 802.11bd) [SasHae20]. Our IETF draft on challenging of IPv6 for vehicular network has also been published in a journal in [Hae20d], and expert the draft to be published as RFC within 2021.

2.1.4 Enhancements to OpenAirInterface.org (OAI)

Within the H2020 SAFE4Rail2 project, we developed a multi-layer emulation architecture for LTE V2X communications on OAI. In [Hae20b], we published a MAC-level Stub architecture, and also developed a PHY-level Proxy architecture, which will be published in 2021.

2.1.5 Enhancements to the iTETRIS C-ITS Simulation Platform

Through industrial contracts with Renault, we further extended the features of the iTETRIS simulation platform. We notably integrated an ETSI ITS compliant DCC architecture and integrated the support for LTE V2X as well.

2.1.6 Enhancements to the IoT Platform

Through two H2020 projects (CONCORDA and 5G-CROCO), we developed an industry-grade IoT/V2X Geoservice architecture for Hazard detection services tailored to Cooperative Connected Automated Vehicles in a large-scale cross-border environment. Stemming from the existing oneM2M/W3C compliant IoT architecture, we developed a flexible

architecture supporting both ETSI ITS and IoT format for cooperative information exchange between vehicles [DatHae20a]. We then developed a hierarchical MEC/Cloud architecture for low latency CCAM services [DatHae20b]. We finally completed the integration of the ETSI ITS Security PKI into the Geoserver.

2.2 5G Networking

2.2.1 Network Slicing and Virtualization

5G will provide a paradigm shift to establish a flexible communication system. This includes the accommodation of multiple concurrent services with different requirements (e.g., data rate, latency, reliability, number of connections, ...) within the 5G system. To this end, multiple logical/virtual networks have to be created in the RAN to isolate the services, while running all of them on the same physical network.

To demonstrate the flexibility of the future RAN, we implemented a prototype [SchNik20a] based on the OpenAirInterface and FlexRAN platforms to show how a RAN can be dynamically customized without service interruptions for different slices. In particular, our solution considers the slice requirements and can adapt the slicing algorithm without interrupting other slices in the network, allowing an efficient resource usage while respecting isolation and performance requirements, in particular latency. Also, this includes dynamic end-to-end slicing by automatically adding core networks as required by the slice owner. Consider Figure 1. A slice owner wishes to embed their URLLC slice. The controller connects to a dedicate core network of the slice owner (see also the deployment on the right side), and inserts the new slice while adapting the slice algorithm. Reductions of the average round-trip time of up to 30% can be observed, as shown in Figure 2.

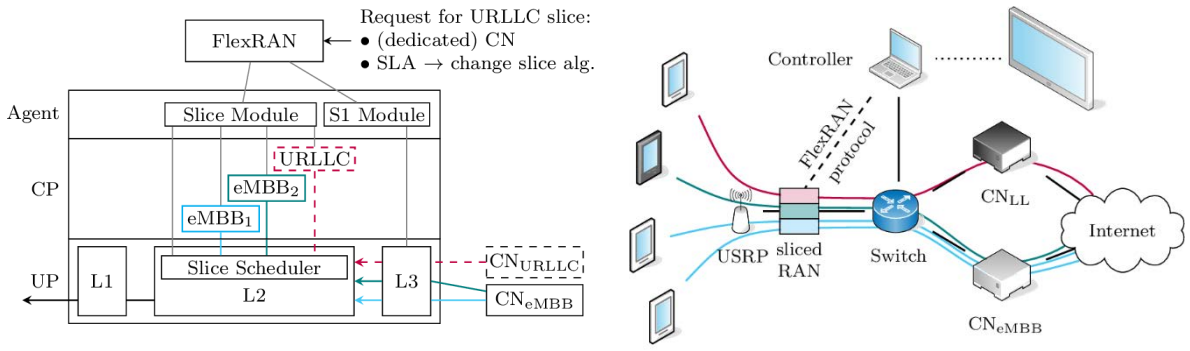


Figure 1: Dynamic slice deployment (left) and testbed setup (right)

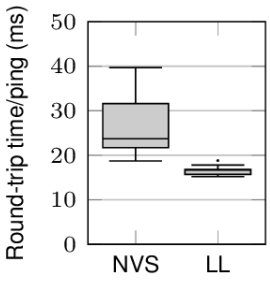


Figure 2: Latency reduction for one slice by loading the appropriate slice algorithm

We further enhanced this prototype [SchNik20b] by extending the RAN with network applications to make the RAN more "intelligent". This is achieved by deploying RAN applications in an agent that can be automatically deployed from a network store, similar to a mobile app store, as shown in Figure 3 (left). Likewise, network functions (e.g., user scheduling functionality) can be automatically downloaded and deployed from such store upon request from the RAN, as shown in Figure 3 (right). The demo included a monitoring application that observed the traffic of selected users and created dedicated low-latency slices when their traffic matched a particular pattern (see Figure 4), resulting in latency reductions as already shown in Figure 2.

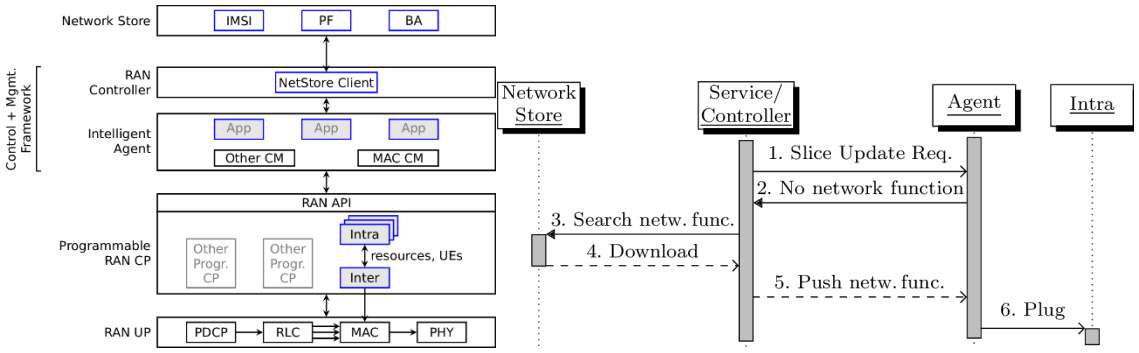


Figure 3: Network store for deploying custom RAN applications (left) and deployment sequence diagram (right)

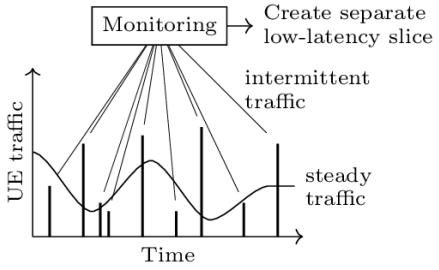


Figure 4: Traffic monitoring/analysis to auto-deploy slices

Finally, we generalized the concept of a flexible RAN: we designed the RAN service engine [SchNik21] (shown in Figure 5) that allows services to customize and extend RAN functionality using containerized micro-services. This is achieved through micro-SDKs that abstract key RAN control endpoints, and which can then be used by the services to flexibly customize and extend the RAN in order to steer control plane behavior. The RAN service engine represents a generalization of the above prototype: multiple service owners can use service-specific RAN controllers to capture service characteristics and steer ``their'' part of the RAN, while the service engine enforces isolation and preserves security.

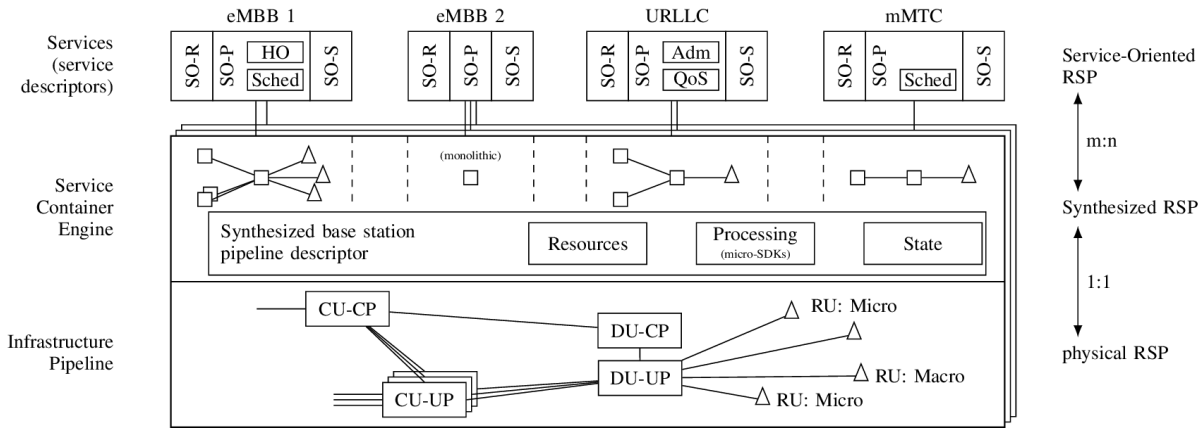


Figure 5: RAN service engine design

Network Slicing being one of the pillars of 5G, it allows running several virtual instances of mobile network on top of a shared infrastructure to support the requirement of 5G network services in terms of bandwidth and latency. Network Slicing builds service-tailored virtual networks to accommodate 5G services. In [KseFra20], we propose a novel solution to slice the Multi-access Edge Computing (MEC) in the context of 5G, while in [MesBer20], we propose a new architecture to instantiate 5G network slice on top of a Datacenter. In [FatLie20], we survey the algorithmic part related to network slicing. In [BakBou20], we introduce a new mechanism based on Machine Learning to predict slice users' CQI and hence optimize the resources for

network slices at the RAN. In [BriKse20], we devise an algorithm relying on federated learning to predict service-oriented network slice performances. Finally, in [TouBer20], we have introduced a multi-domain orchestrator that can be used to deploy network slices on different administrative domains.

2.2.2 Network Virtualization and Orchestration

2.2.2.1 KUBE5G: A CLOUD-NATIVE 5G SERVICE PLATFORM

With the proliferation of use-cases envisioned to be supported by 5G networks, the focus is not only on the performance, but also on the service agility and elasticity. Cloud native principal is a methodology of designing lightweight, isolated context, and deployable at scale applications that natively exploit the features of cloud. However, supporting telco applications (e.g., 4G/5G services) in the cloud brings up many challenges, such as the coexistence of physical and virtual functions, (near) real-time resource provisioning, service continuity, and strict latency and data rates. Kube5G, as a realizable agile service platform in support of 4G/5G cloud native applications, introduces a novel approach in building and packaging a cloud-native compliant telco network function (NF) in a form of nested well-defined layers, as presented in the Figure 6.

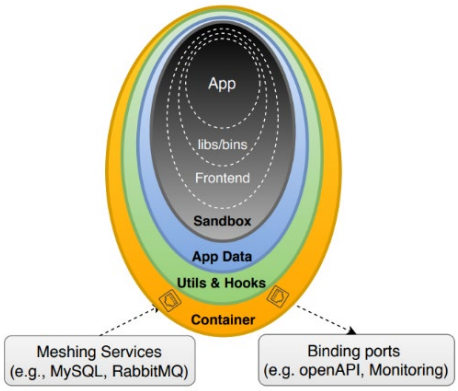


Figure 6: The nested layererd composition of network functions in Kube5G

In addition, Kube5G supports workflow for continuous integration and development (CI/CD) operations in support of multi-version network functions (physical, virtual and cloud functions) as its structure is given in Figure 7. The CI/CD provides new functions/features to the end users in a quick and reliable way. As illustrated in Figure 8, Kube5G architecture is composed of:

- Cloud Infrastructure that can be bare-metal or any type of cloud private/public or even hybrid.
- Software defined resources for the objective of the abstraction of the available resources and exposing them as virtual resources to the applications.
- Infrastructure and application Orchestration the for orchestration the containerized application and the available resources in cloud environment
- Runtime execution that can support not only cloud native applications (CNF), but also the vitalized applications that are designed in a non-cloud native way (VNF)
- The CNF and VNF running on the top to provide the intended services, while they are supported by API and the sake of communication among services, and well as the access to these application from outside of the environment (e.g., remote interaction with the applications)
- A set of extensions for the sake of making Kube5G flexible and extensible for new features.

Kube5G uses Kubernetes as its application lifecycle management system and deploys a Kube5G Operator (which is a good example of Kube5G extensions) on the top of Operator framework to facilitate the deployment of NFs. To this end, Kube5G-Operator implements and automates both basic (e.g. installation, configuration, and monitoring) and advanced (e.g. re-configuration, update, backup, failover, restore) operations in a form of a portable Kubernetes-native applications running inside a Kubernetes cluster. In addition, it is capable of managing complex 5G services (e.g., flexible network deployment), which is traditionally managed by administrators and their operational scripts (e.g., Ansible). These efforts alongside a set of

experiments to validate the structure and compare its performance with the previous designs are given in [AroNikGlobeP20, AroNikGlobeD20, AroNikNoms20].

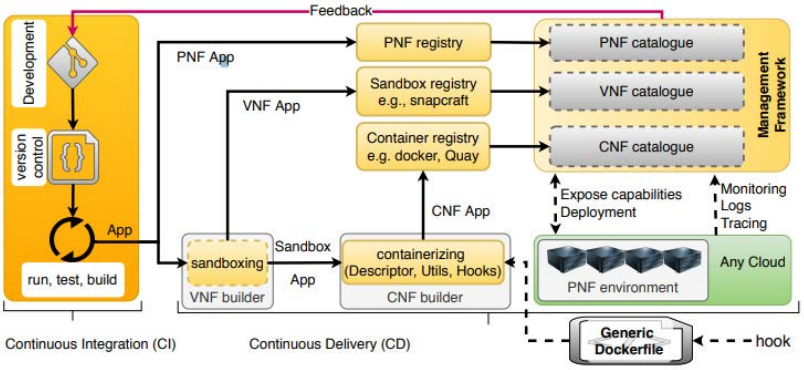


Figure 7 : The CI/CD workflowOf Kube5G

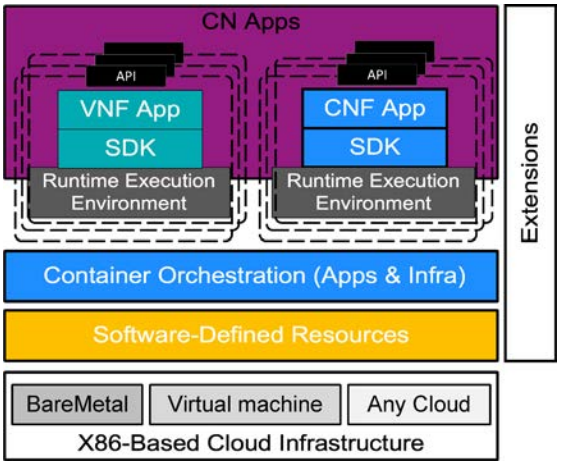


Figure 8: Kube5G Architecture

2.2.3 Multi-Access Edge Computing (MEC)

Multi-access Edge Computing (MEC) is enabling a new generation of services that operate close to end-users aiming at reducing the end-to-end latency. Edge computing allows the deployment of two types of services: (i) services that require low-latency access to user plane traffic; (ii) context-aware services that adapt the delivered service according to the users' environment. In [KsenFra20], we propose a new model to extend MEC services to support IoT protocols and particularly LORA. In [BriFra20], we introduce a placement algorithm of MEC applications on top of a federated Edge cloud infrastructure, considering different criteria such as MEC service availability, CPU resources, and latency. In [FatLie20], we introduce a novel algorithm to predict vehicle-mobility and adapt MEC resources accordingly in the context of vehicular networks. In the same spirit, we propose a MEC orchestration system to manage MEC applications in the context of connected vehicles. In [KseBri20], we propose a novel application to fight COVID-19 by taking advantage of MEC low latency to host an application to detect if persons are respecting social distancing.

2.2.4 Unmanned Aerial Vehicle (UAV)

Unmanned Aerial Vehicle (UAV) applications and services have gained a huge deployment and adoption in different fields, such as the military domain (Defense or reconnaissance) and the civilian domain (Healthcare, surveillance, and transport). UAVs are used for dangerous tasks, where humans cannot intervene. Flying drones are mainly used for their small size and ability to move into narrow places. UAV operations are generally critical and require, during operations, a control link with the drones, which should be reliable with very low latency. In [TalKse20], we propose a new architecture to enable UAV services in 5G and beyond mobile systems. The objective is to manage and run UAV services as a network slice in 5G. In

[SimKse20], we introduce a novel architecture that aims to integrate 5G in U-Space to allow a safe flight of drones using 5G connectivity. In [SimBou20], we introduce a new algorithm that plans drone flight by avoiding MEC relocations and ensures that drones are always connected to the best MEC server to ensure a safe flight. Finally, in [BriKsen20], we survey the challenges related to the usage of federated learning in the context of UAV.

2.3 Short-Packet Channel Coding Strategies for 5G NR Control Channel Coverage Enhancement

In the context of the EURECOM-Qualcomm-France Brevets-IMT contract work on channel coding strategies was carried out during 2020 and resulted in the filing of three patent applications and two 3GPP study item contributions.

The current inventions [Knopp2020a,Knopp2020b,Knopp2020c] propose novel binary linear coding schemes designed for short-packet scenarios where channel uncertainty due to wideband signal fading is present. These were proposed for the Release 17 Coverage Enhancement Study Item [Knopp2020d, Knopp2020e]. The proposed strategies are optimized for short block lengths and compatible with 3GPP resource grids, in particular those for 5G New Radio. These constructions aim to increase the sensitivity of the gNodeB receivers in order to provide enhanced coverage for PUCCH and PUSCH Uplink channels. In the case of control channels, we

1. Propose orthogonal constructions for 3GPP configurations which provide the maximum performance without concern for the amount of signaling dimensions
2. Propose new binary codes for BPSK and QPSK signaling optimized for the NON COHERENT METRIC which perform as close as possible to orthogonal constructions using less signaling dimensions.
3. Propose concatenated schemes using short block length coding optimized for the NON COHERENT METRIC with traditional powerful iterative coding (3GPP Polar code)

In the case of user-plane data channels, we propose concatenated schemes using short block length coding optimized for the NON COHERENT METRIC with traditional powerful iterative coding (3GPP Polar code, 3GPP Turbo code, 3GPP LDPC code) and consider very low-spectral efficiency.

Three new patent filings are expected for early 2021.

2.4 Ultra Reliable Low Latency Communications

2.4.1 Feedback Enhancements for Semi-Persistent Downlink Transmissions in Ultra-Reliable Low-Latency Communication

In downlink semi-persistent scheduling (SPS) transmission, the moment of feedback is pre-configured when SPS resources are activated. However, if the feedback resource is in DL slot of TDD configuration, this feedback is dropped. It is harmful to DL SPS transmission, especially for URLLC.

[LeSal20a] provides a scheme to guarantee the transmission of HARQ feedback and potential data retransmission when there is UL-DL slot/sub-slot conflict at the indicated feedback resource. The scheme comprises a dynamic indication of feedback resource without using an associated downlink control information (DCI). An acknowledgement (ACK)-only feedback protocol is proposed to best suit the scenario in question. The combination of dynamic indication of feedback resource and ACK-only feedback structure guarantees higher reliability and brings flexibility to the transmission of HARQ feedback as confirmed by the simulations.

2.4.2 Channel Access Enhancements in Unlicensed Spectrum for NR URLLC Transmissions

Increase of mobile traffic and bandwidth requirements for new applications have resulted in a shortage of licensed spectrum so that even URLLC services are deemed to be running over the unlicensed spectrum. This may require a significant re-design of channel access, transmission and reception procedures over the unlicensed spectrum.

[LeSal20b] provides an analytical analysis of current channel access procedures, listen before talk (LBT). LBT states in random duration channel access are modeled through Markov chains which help characterize the closed form expressions for average channel access time. These expressions are evaluated using the parameters from currently standardized channel access priority classes. The evaluation shows the total inability of several current channel access classes to support URLLC services over the unlicensed spectrum even under low load conditions. The insights gained from this analysis lead the proposal of new channel access priority classes where new special classes are introduced for URLLC services.

2.4.3 Enhancing URLLC Uplink Configured-Grant Transmissions

In uplink configured grant transmission, the repetitions are not allowed to transmit outside the hybrid automatic repeat request (HARQ) process containing the first repetition. It might cause a smaller number of transmitted repetitions than configuration that is harmful to the performance of URLLC.

[LeSal20c] uses the reserved resources where the UEs can transmit the repetitions outside the original HARQ process until the configured number is reached. The scheme is developed further when the gNB is equipped with a successive interference cancellation (SIC) receiver so it can decode multiple repetitions of the different UEs in the same reserved resource.

2.5 Real-Time Performance Evaluation of Relative Calibration on the OAI 5G testbed

In 2019 we described how we implemented and evaluated a distributed MIMO system based on channel reciprocity on the OpenAirInterface (OAI) C-RAN testbed deployed at Eurecom. One of the issues in this evaluation was that it was not possible to evaluate the absolute calibration error of the system as we did not know the ground truth of the calibration coefficients.

In [MagKal20] overcome these deficiencies by using a simulation model of the OpenAirInterface (OAI) C-RAN testbed deployed at Eurecom. We present the performance evaluation of a fast calibration (FC) scheme based on antenna grouping. We generate an indoor LOS radio wave propagation model using ray tracing technique and perform distributed channel reciprocity calibration which is required to exploit uplink (UL) channel estimates to infer the precoder performed on the downlink (DL) channel. We consider a channel based on the geometry of the area where the RRUs are distributed and compare the different choices to form the RRU groups. We validate that the simulation results match the experimental results from our previous work.

2.6 Experimental Mimo Systems Based on OpenAirInterface

2.6.1 Exploiting Channel Reciprocity in Massive MIMO

Massive MIMO is one of the key enabling technologies for in 5G and will be pursued for 6G. Our work in this topic has been focused on building the OpenAirInterface Massive MIMO testbed, which was one of the world's first LTE compliant base stations equipped with large antenna array. Previously; Eurecom had proposed a generalized framework for antenna calibration that permits the grouping of antennas. Time Division Duplexing (TDD) Massive MIMO relies on channel reciprocity to derive the channel state information at the transmit side (CSIT) from the uplink (UL) channel estimates. This reciprocity is impacted by the transmit and receive front ends which are non-reciprocal and hence a calibration is required to obtain the downlink (DL) channels from the UL channel estimates. In [GopSlo20], we provide a naïve uplink/downlink duality framework in which the uplink/downlink calibrated channel reciprocity is exploited to calibrate uplink optimized beamforming receive filters for use in the downlink.

2.7 Heterogeneous Network Analysis

Combining heterogeneous access networks gives more flexibility to offer service continuity, better Quality of Experience (QoE), service coverage enhancement, and transmission diversity. Challenging issues have to be addressed on the architectural elements and protocols to take advantage of this heterogeneity. Approaches include the combination of abstraction layers, cross layer design and advanced management layers. The following projects illustrate the various approaches.

2.7.1 Incentivized D2D-Based Offloading and Optimal Relay Selection

A popular approach to offload data and create, in some sense, a "virtual" heterogeneous network, is to treat some UEs as relays, for users with bad signal, such as cell edge users (CUE). The performance of the CUEs can be improved if they transmit their data via suitable relay UEs (RUEs) exploiting device-to-device (D2D) communication. The critical aspect of the relaying concept is to offer convenient incentives for the RUEs to motivate them to act as relays. The contribution of this paper is twofold. To this end, in recent work we proposed a new incentive mechanism for the RUEs that can exploit certain amount of resources allocated to the CUE. Depending on the preferences of users, the CUEs/RUEs can benefit from relaying in terms of capacity enhancement, reduction of energy consumption or both. In this respect, we provided a detailed

analysis on how and when relaying is of benefit for both sides. We also proposed a low-complexity greedy relay selection algorithm incorporating the incentive mechanism that increases capacity up to 32.1% and/or reduces energy consumption by up to 36.1% when compared to state-of-the-art schemes. Finally, we proved that this greedy approach gives close-to-optimal performance. This work has been extended to a journal, where we revisit the model to allow for multiple CUEs per RUE, as well as to consider machine-learning based solutions as input to predict SNR and candidate CUE/RUE pairs [MacSpyBec20].

2.8 Caching and Recommendation Co-Design

Caching content at small cells has been suggested as a win-win opportunity to both improve access latency, and alleviate congestion on the backhaul and core networks. Recently, we considered a caching setup which can be aided by a recommendation system, providing recommendations that are relevant to the user (from a traditional recommendation system standpoint) but also favor low(er) network access cost. More specifically, in [GiaGioSpy21], we consider a user who makes a random (with known statistics) amount of requests, and is browsing videos in a multimedia application according to a PageRank-like access model. For this type of sequential decision problems, a natural formulation is Markov Decision Problems. To this end, we appropriately setup the cost to favor the cached content, but also constrain the policy to offer relevant content to the user. Interestingly, our formulation has an extra trick, which has to do with the decoupling of the action space in the inner minimization loops of the Policy Iteration algorithm. This allows us to scale our solutions for any size of the recommendation batch offered to the user. Furthermore, using MDP style of algorithms leads to a state decomposition which allows for parallel minimization steps. In the results section, algorithmic-wise we show that (a): our approach convincingly beats vanilla MDPs, which become intractable, (b): we outperform recent convex optimization approaches in convergence time, and network gain-wise, we show that myopic state-of-the-art efforts that ignore the length of the user session are highly suboptimal.

Going a step beyond, in [TsiSpy20] we consider the joint caching and recommendation problem based on the fact that Content Providers like Netflix and Google own their in-house network of caches (and therefore can control both caching and user recommendations decisions). We define a framework where the user experience on streaming platforms is expressed as a balanced sum of the quality of service (e.g., streaming rate) as a function of the caching decisions and of the quality of recommendations as a function of the recommendations that the user receives. We then formulate an optimization problem that maximizes the aggregate user experience. Based on the properties of the problem, we provide a polynomial-time algorithm that approximates the optimal solution within a constant factor. Finally, our numerical evaluations show that, under realistic scenarios, our proposed algorithm outperforms baseline policies and algorithms proposed in the literature.

In [CosSpy20a] we also consider joint caching and recommendation, but in the context of “soft cache hits”, earlier introduced by our group. In this work, we first show that the optimization problem generated when introducing a limited number of recommendations as design variables into the caching problem is NP-hard. We then show that using a primal decomposition of the joint problem we can design an algorithm with constant approximation to the optimal solution, and confirm the superior performance of this algorithm over existing heuristics in real-world datasets.

In the follow-up work [CosSpy20b] we study how the complex network characteristics between the individual contents of a catalog such as that of Netflix or Spotify affect the performance of algorithms simultaneously designing the caching and recommendations. In particular, we formalize the model of a content catalog as a graph and identify key "content graph" and communication network setup parameters that have an impact on the performance of the caching and recommendation algorithms. We run extensive experiments in synthetic data isolating each parameter to test its effect independently of the others, and we validate our findings on real-world data. Moreover, we explore the possibility of using the observed dependencies between the catalog structure and the algorithms' outcome to do policy performance prediction, with very encouraging results.

Finally, in [SerKasPin20] we perform an experimental study to confirm the assumptions made in the above works, as well as to corroborate the performance advantages observed in simulated setups. Specifically, we have create a YouTube proxy that re-arranges the default recommendation list for a user, to favor some (assumed) low cost content, and records the click behavior as well as collects explicit feedback over a 5-content viewing session.

2.8.1 Caching Policies for Delay Minimization in Small Cell Networks with Joint Transmissions

Heterogeneous small-cell topologies have been considered as a promising solution to cope with the traffic increase in future 5G/6G cellular networks. A key feature of this kind of networks is the high base station (BS) density, i.e., user equipments (UEs) are often in range of multiple BSs simultaneously. This characteristic enables the deployment of Coordinated Multi-Point (CoMP) techniques. Specifically, BSs may cooperate to jointly transmit content to UEs, achieving higher transmission rates. On top of such architecture, assuming that BSs are equipped with storage capabilities, we design caching schemes aiming to provide UEs with the best quality of experience (QoE) possible. In [RicNegSpy20], we consider that QoE is captured by the average delay a UE experiences in order to retrieve a content. We study two different setups: (i) a static, centralized two-phases cache system, where in phase 1 files popularities are estimated and phase 2 comprises the content placement itself; and (ii) a dynamic cache system where caches individually update their content on-the-fly as new requests arrive. For setup (i), we provide a greedy algorithm with 1/2-approximation rate for the case where files have stationary popularities and network transmission conditions are homogeneous. For setup (ii), we provide online, distributed algorithms that probabilistically update cache states based on the delay gain offered by a given file placement. The first policy, qLRU-Δd, is asymptotically optimal under IRM request process when the insertion probability tends to 0. The second policy, 2LRU-Δd, is empirically shown to be more reactive, so it has better performance handling real request processes. All proposed caching schemes were successfully validated and their performance evaluated through numerical simulations.

3. RESOURCE OPTIMIZATION AND CROSS-LAYER DESIGN

3.1 Cognitive Radio Systems

DECENTRALIZING MULTI-OPERATOR COGNITIVE RADIO RESOURCE ALLOCATION

We address the problem of resource allocation (RA) for spectrum underlay in a cognitive radio (CR) communication system with multiple secondary operators sharing resource with an incumbent prima-ry operator. The multiple secondary operator RA problem is well known to be especially challenging because of the inter-operator coupling constraints arising in the optimization problem, which render impractical inter-operator information exchange necessary. In this work, we consider a satellite set-ting for multi-operator CR. In the CR maturation regime, i.e., the period in which the secondary sub-scriber density is growing yet remains much below that of incumbent users, we show that in fact the inter-operator mutual constraints can be neglected, thus making distributed (across secondary oper-ators) optimization possible. Furthermore, we establish analytically that the mutual constraints as-ymptotically vanish with the primary user density [TohGes20].

We also address the problem of resource allocation (RA) in a cognitive radio (CR) communication system with multiple secondary operators sharing spectrum with an incumbent primary operator. The key challenge of the RA problem is the inter-operator coordination arising in the optimization problem so that the aggregated interference at the primary users (PUs) does not exceed the target threshold. While this problem is easily solvable if a centralized unit could access information of all secondary operators, it becomes challenging in a realistic scenario. In this work, considering a satellite setting, we alleviate this problem by proposing two approaches to reduce the information exchange level among the secondary operators. In the first approach, we formulate an RA scheme based on a partial information sharing method which enables distributed optimization across secondary operators. In the second approach, instead of exchanging secondary users (SUs) information, the operators only exchange their contributions of the interference-level and RA is performed locally across secondary operators. These two approaches, for the first time in this context, provide a trade-off between performance and level of inter-operator information exchange. Through the numerical simulations, we explain this trade-off and illustrate the penalty resulting from partial information exchange [TohGes21].

3.2 Caching in Wireless Networks

CACHE-AIDED OPTIMIZATION IN ADVANCED NETWORKS

The work in [SerLamSpy20] presents a new way of exploiting non-uniform file popularity in caching networks. Focusing on the interference channel with cache-enabled transmitters and receivers, we show how non-uniform file popularity can be used to accelerate the impact of transmitter-side data redundancy in coded caching. This approach is motivated by the recent discovery that under realistic file-size constraints, having content appear in multiple transmitters can boost multiplicatively the speed-up factor attributed to coded caching. We formulate the problem through an optimization algorithm, which seeks to optimize the number of transmitters each file is cached at, as a function of that file's popularity. Part of the optimization effort involves a biconvex problem; such problems are traditionally solved by heuristic Alternate Convex Search methods that generally do not guarantee the global optimum. To avoid this, we follow a more involved path that includes the design of a new search algorithm that exploits the properties of the caching problem itself. The overall optimization algorithm provably achieves the globally optimal solution, and does so with a complexity that scales as a polynomial function of the logarithm of the size of the file catalog. In the end, the optimal transmitter-side cache placement yields multiplicative speedup factors over traditional multi-transmitter coded caching algorithms.

The work in [MalSerPar20]-[MalSerPar20b] establishes the exact optimized performance limits of stochastic coded caching when users share a bounded number of cache states, and when the association between users and caches, is random. Under the premise that more balanced user-to-cache associations perform better than unbalanced ones, our work provides a statistical analysis of the average performance of such networks, identifying in closed form, the exact optimal average delivery time. To insightfully capture this delay, we derive easy to compute closed-form analytical bounds that prove tight in the limit of a large number of cache states. To alleviate this adverse effect of cache-load imbalance, we consider various load-balancing methods, and show that employing proximity-bounded load balancing with an ability to choose from h neighboring caches, the aforementioned scaling reduces dramatically.

4. LOCALIZATION

4.1 User Localization Techniques

SENSOR SELECTION FOR MODEL-FREE SOURCE LOCALIZATION

The ability for a wireless network to precisely localize the radio nodes composing it is a great tool towards system optimization and is increasingly seen as a basic service requirement. In the past, model-free algorithms such as weighted centroid localization (WCL) have proved popular, especially in the context of sensor networks, due to their simplicity and robustness to temporal changes in wireless propagation properties. However, WCL algorithms are biased since they implicitly require a uniform sensor distribution around the source in all directions. In this work, we demonstrate that instead of employing all the sensors that result in a possibly unbalanced sensing pattern, it is better to reduce the number of sensors such that the subset of selected sensors symmetrically distributes around the source, which in principle would need to know the source location in advance. Here, we develop a sensor selection algorithm which manages that goal while blindly. Using less than half of the sensors, a 30% reduction in localization error is demonstrated from our numerical experiments [TohChe20].

4.2 UAV-Aided Localization

The work in [EsrGan20] considers the problem of localizing outdoor ground radio users with the help of an unmanned aerial vehicle (UAV) on the basis of received signal strength (RSS) measurements in an urban environment. We assume that the propagation model parameters are not known a priori, and depending on the UAV location, the UAV-user link can experience either line-of-sight (LoS) or non-line-of-sight (NLoS) propagation condition. We assume that a 3D map of the environment is available which the UAV can exploit in the localization process. Based on the proposed map-aided estimator, we devise an optimal UAV trajectory to accelerate the learning process under a limited mission time. To do so, we borrow tools such as Fisher information from the theory of optimal experiment design. Our map-aided estimator achieves superior localization accuracy compared to the map-unaware methods, and our simulations show that optimized UAV trajectory achieves superior learning performance compared to random trajectories.

5. COMMUNICATION THEORY AND TECHNIQUES

5.1 New Waveforms for Beyond5G

AFFINE FREQUENCY DIVISION MULTIPLEXING (AFDM): A NEW WAVEFORM FOR DOUBLY DISPERSIVE CHANNELS

Future broadband systems are envisioned to support a wide spectrum of services and applications, including wireless communications in high mobility environments. Various conventional multicarrier techniques, such as orthogonal frequency division multiplexing (OFDM) and single-carrier frequency division multiple access (SC-FDMA), have been deployed in existing standardized communication systems. These schemes are shown to achieve satisfactory or even optimal performance in time-invariant frequency selective channels. However, in high mobility scenarios, orthogonality among subcarriers is destroyed due to large Doppler frequency shifts. Therefore, novel beyond OFDM waveforms are required for reliable and efficient transmission in high-mobility scenarios. We have introduced Affine Frequency Division Multiplexing (AFDM), a new chirp-based multicarrier scheme that is based on the discrete affine Fourier transform (DAFT). DAFT, a discrete version of linear canonical transform, is a generalization of discrete Fourier transform with two tunable parameters that can be adapted to the delay-Doppler profile of channel to better cope with doubly dispersive channels. Deriving the AFDM input-output relation, we have shown that by properly tuning its chirp parameters the delay-Doppler representation of each path is achieved and all paths will be separated. Moreover, in [BeKsaKou21] we have analyzed the diversity order of AFDM and have shown analytically that AFDM achieves full diversity in two-path channels. For channels with more than two paths, we have provided conditions that guarantee full diversity order and have shown by simulations that AFDM is diversity optimal in most cases for general linear-time variant channels.

5.2 Multi User MIMO Techniques

D2D-AIDED MULTI-ANTENNA MULTICASTING

Multicast communications, in which a transmitter wishes to convey a common message to multiple receivers, arise naturally in various wireless network scenarios. Specifically, multicast services are relevant for many challenging applications ranging from wireless edge caching, where popular contents are cached during off-peak hours and subsequently distributed, to the broadcasting of safety messages in vehicular networks. In parallel, device-to-device (D2D) communications have recently gained momentum on account of emerging applications such as multicasting, machine-tomachine communication, and cellular-offloading, and are expected to be included in the upcoming fifth-generation (5G) wireless system as a new paradigm for enhancing the network's performance.

It is well known that the multicast capacity is limited by the user equipments (UEs) in poor channel conditions and vanishes when the number of UEs increases for a fixed number of base station (BS) antennas. To overcome this issue, some enhanced multicasting schemes leveraging UE cooperation enabled by D2D links have been proposed to tackle the vanishing behavior of the multicast capacity when the UE population becomes large. In this context, the UEs who are able to decode the common message sent by the BS act as opportunistic relays and cooperate in a subsequent phase to retransmit the information to the remaining ones. This framework, has been considered so far only for the case of single-antenna BS. In [MurAtz19], we point out that endowing the BS with precoding capabilities radically transforms the above scheme both in nature and performance, since we can leverage the precoding gain brought by multiple antennas in addition to the D2D links.

Furthermore, in [MurAtz19b] we propose a two-phase protocol where the precoding optimization at the multi-antenna BS relies only on the channel statistics. This is a key difference with respect to our previous work in [MurAtz19], which assumes perfect channel state information (CSI) at the BS. An anticipated advantage of this approach is the great reduction in the overheard necessary to acquire costly instantaneous CSI, which is a major limitation in scenarios with large number of UEs.

5.3 D2D-Aided Communications

5.3.1 D2D-Aided Multi-Antenna Multicasting Under Generalized CSIT

Multicast communications, in which a transmitter wishes to convey a common message to multiple receivers, arise naturally in various wireless network scenarios. Specifically, multicast services are relevant for many challenging applications ranging from wireless edge caching, where popular contents are cached during off-peak hours and subsequently distributed, to the broadcasting of safety messages in vehicular networks. In parallel, device-to-device (D2D) communications have recently gained momentum on account of emerging applications such as multicasting, machine-to machine communication, and cellular-offloading, and are expected to be included in the upcoming 5G wireless system as a new paradigm for enhancing the network's performance.

It is well known that the multicast capacity is limited by the UEs in poor channel conditions and vanishes when the number of UEs increases for a fixed number of BS antennas. To overcome this issue, some enhanced multicasting schemes leveraging UE cooperation enabled by D2D links have been proposed to tackle the vanishing behavior of the multicast capacity when the UE population becomes large. In this context, the UEs who are able to decode the common message sent by the BS act as opportunistic relays and cooperate in a subsequent phase to retransmit the information to the remaining ones. This framework, has been considered so far only for the case of single-antenna BS. In our early work, we point out that endowing the BS with precoding capabilities radically transforms the above scheme both in nature and performance, since we can leverage the precoding gain brought by multiple antennas in addition to the D2D links.

Specifically, we propose a two-phase protocol capitalizing on both the precoding at the multi-antenna BS and on D2D communications which is applicable under generalized assumptions on the CSIT, namely i) Perfect CSIT ii) statistical CSIT and iii) topological CSIT. Hence, the above scenarios from i) to iii) correspond to settings with decreasing requirements on the information at the BS side. While the configuration i) is relevant to the case of moderate (or finite) number of UE and a low-mobility scenario, the configuration iii) is suited to a high-mobility and/or large number of UE regime arising for instance in a vehicular network, where the BS multi-antenna beam pattern ought to be designed on the basis of a city map and road traffic distribution. The configuration ii) is an intermediate case between i) and iii). Such work has been submitted to the IEEE TWC in February 2021 [MurAtz21].

5.4 Metasurface-assisted Wireless Networks

5.4.1 RISMA: Reconfigurable Intelligent Surfaces Enabling Beamforning for IoT Access

Re-configurable Intelligent Surfaces (RISs), which apply controllable transformations into impinging radio waves without leveraging on power amplifiers, pose a niche of opportunities for the optimization of wireless systems at a low cost. RIS technologies are indeed gaining a lot of momentum because of their ability to turn the stochastic nature of the wireless environment---fundamentally passive---into a programmable channel that plays an active role on the way in which signals propagate.

A RIS is essentially a continuous meta-surface that can be modeled as a grid of discrete unit cells spaced at sub-wavelength distance that can alter their electromagnetic response, such as phase, amplitude, polarization and frequency in a programmable manner. For instance, they can be tuned such that signals bouncing off a RIS are combined constructively to increase signal quality at the intended receiver or destructively to avoid leaking signals to undesired receivers. Despite being a continuous surface, such cells are conveniently modelled as equivalent antennas. To this aim, we built an optimization framework and propose an algorithm to jointly optimize both the precoder of the transmitter and the RIS parameters in terms of induced phase shifts and amplitude attenuation produce substantial gains in sum-rate performance. Compared to existing works our method, denoted as RISMA, has the key advantage to obtains simple and tractable expressions of the optimization variables. As a result, our proposed RISMA approach guarantees high performance and scalability in the case of massive access scenarios where a large number of UEs needs to be served. Moreover, we adapt RISMA to accommodate for practical constraints when using low-resolution RISs that can only induce a discrete set of phase shifts to the incoming signal rather than any real value from a range. To this purpose we propose Lo-RISMA and show its effectiveness even when the low-resolution RIS has only few bits of precision. Such method has been presented in our paper which was accepted by the IEEE JSAC special issue on massive access for 5G and beyond in Aug. 2020 [MurSci20].

5.5 Multi-User MIMO Systems

5.5.1 Rate Balancing in Multi-User MIMO Systems

One important criterion in designing wireless networks is ensuring fairness. With respect to applications in communication networks, fairness is closely related to min-max or max-min optimization problems, also referred to as balancing problems. In [GhaSlo20], we study the max-min user rate optimization problem for multicell multiuser MIMO systems. Due to the complexity of the problem, we exploit the rate Mean Squared Error (MSE) relation, formulate the balancing operation as constraints leading to Lagrangians in optimization duality, allowing to transform rate balancing into weighted MSE minimization with Perron Frobenius theory. We consider the problem for per cell power constraints and reformulate the respective Lagrange multipliers as a single weighted power constraint in which the weighting can be optimized to lead to the satisfaction with equality of all power constraints.

5.6 Massive MIMO Systems

5.6.1 Hybrid Beamforming for Massive MIMO

The French FUI project MASS-START aims at the development of a 5G open source solution for the Massive MIMO (MaMIMO). The main targets involve the development of signal processing algorithms for 5G cellular systems. Communication in the millimeter wave (mmWave) frequency bands is seen as a potential high speed enabling technology for 5G systems. mmWave frequencies allow the use of large number of antennas at the Base Station (BS) to increase the spectral efficiency of the system. Due to the large bandwidths and therefore high costs of the digital processing chains, hybrid analogue digital architectures promise a reasonable trade-off between performance and complexity. The aim of the project is to be aligned as much as possible with the 5G NR standard.

In a hybrid system, the number of radio frequency (RF) chains are less than the number of antennas and the analog beamformer (BF) network is implemented using phase shifters, which leads to multiple signal beams in particular angular directions. Low dimensional digital BF is implemented at the baseband, which multiplexes the multi-user signals across these multiple beams. However, the unit modulus constraint on the analog phasors make the sum rate maximization problem highly non-convex. Hence, most of the state of the art designs on multi-user MaMIMO systems focus on a sub-optimal approach to analog phasor design. From our own previous work based on weighted sum mean squared error (WSMSE) solution, it is noted that the iterative algorithm for the update of the phasors of the analog beamformer does converge to different local optimum, which depends on the initialization, due to which the performance compared to a fully digital BF degrades. In 2018, we developed efficient analog phasor design based on deterministic annealing (DA, originally proposed in machine learning literature) for narrowband systems. In [KurSlo20a], we extend this approach for a full-duplex wideband orthogonal frequency division multiplexing (OFDM) system, which is indeed the driving technology behind next generation wireless systems and make much more practical to consider compared to the narrowband system design.

MASSIVE MIMO BEAMFORMING DESIGN: A NEW LARGE SYSTEM ANALYSIS

The large system analysis becomes an important topic to consider since Monte-Carlo simulations involving large numbers of antennas and user equipments (UEs) become cumbersome in a MaMIMO system. For the large system analysis (LSA) of massive MIMO (MaMIMO), we considered a stochastic geometry inspired randomization of the user covariance subspaces [KurSlo20b]. We derive the BFs under partial channel state information at the Tx (CSIT) using an upper bound of the expected weighted sum rate (EWSR), referred to as Expected Signal and Interference Power weighted sum rate (ESIP-WSR). Gaussian (posterior) partial CSIT can optimally combine channel estimate and channel covariance information. In a previous work from our group, we have extended the LSA of MaMIMO to a scenario with users having different channel covariance matrices and BF techniques with partial CSIT. However, due to the abundance of different covariance matrices, the resulting deterministic analysis does not allow for much insight. The multi-antenna stochastic geometry aspect introduced in [KurSlo20b] reduces such LSA analysis back to the simplicity of the case of multiple of identity covariance matrices. Major results can be summarized as follows: 1) We evaluate the ergodic sum rate performance for least squares (LS), linear minimum mean squared error (LMMSE) and subspace projection channel estimators with ESIP-WSR which is tight in (a certain) massive MISO limit. Numerical results suggest that there is substantial gain by exploiting the channel covariance information compared to just using the LS estimates. 2) The analysis presented in the paper provides accurate (also validated in the simulations) spectral efficiency (SE) expressions under realistic channel estimation quality which are useful at any operating signal-to-noise-ratio (SNR). Moreover, these SE expressions are very simple and provide analytical insights

into the system behavior and depends only on few system parameters such as channel power across multipath, Tx power, rank of the channel covariance matrix and number of Tx antennas. 3) We furthermore provide certain illustrative examples, which are special cases of the ESIP-WSR BF such as perfect channel CSIT case, and covariance only CSIT (CoCSIT) scenario where only the channel covariance information is known at the BS. We show that we can obtain analytical expressions for the implicit equations, which need to be solved as part of the large system analysis, which in fact provide analytical insights into the system behavior. 4) We compute the (low and high) SNR offsets for the sum rate of different BFs under various channel estimates. The resulting analysis being conducted for constant channel estimation error (a case of finite rate feedback channels in frequency division duplex (FDD) systems) and for estimation error being scaling inversely proportional to the SNR.

MASSIVE MIMO PILOT CONTAMINATION: A RATE SPLITTING APPROACH

The spectral efficiency (SE) of Massive MIMO (MaMIMO) systems is affected by low quality channel estimates. Rate-Splitting (RS) has recently gained some interest in multiuser multiple antenna systems as an effective means to mitigate the multi-user interference due to imperfect channel state information. Our work [KurSlo20c] investigates the benefits of RS in the downlink of a single-cell MaMIMO system when all the users use the same pilot sequence for channel estimation. Novel expressions for the SE achieved in the downlink by a single-layer RS strategy (that relies on a single successive interference cancellation at each user side) are derived and used to design precoding schemes and power allocation strategies for common and private messages. A maximum ratio (MR) precoding scheme is used for private streams while a precoder based on a weighted combination of the channel estimates of all UEs is adopted for the common stream. A novel algorithm is proposed to allocate the power among the common and private streams. The interference leakage aware water-filling solution is derived using difference of convex functions approach. Numerical results are used to show that the proposed RS solution achieves higher spectral efficiency that conventional MaMIMO with maximum ratio precoding.

MASSIVE MIMO CHANNEL ESTIMATION USING SPARSE SIGNAL PROCESSING

This section presents the channel estimation algorithm proposed by our team “Learned ChEster” (which stands for learned channel estimator) for the ML5G-PHY [mmWave MIMO channel estimation] challenge conducted by ITU. The problem formulation here considers a mmWave channel estimation under a hybrid MIMO architecture at the base station and user side. We were given datasets at different SNRs (mostly in the low SNR regime) and for different number of pilot subcarriers. Our proposed solution [SubKur20d] won the third prize in the challenge. Here is brief overview of our solution. We obtain an initial channel estimate using multi-level greedy search procedure (MLGS) with a coarsely quantized beamspace dictionary. We adopt the simultaneously weighted orthogonal matching pursuit (SW-OMP) algorithm as our base algorithm to form an initial estimate of the channel. As the sparsifying dictionary is unknown a priori, we use row-truncated discrete Fourier transform matrices as the array steering matrices. Hence, we obtain S virtual beamspace channel taps in the first step. Further, noise variance is estimated using the residual output from MLGS stage. MLGS provides reasonable AoDs and AoA estimates, but the virtual beamspace quantized sparsifying dictionary used may not be able to obtain the exact AoD and AoA that lie in the off grid regions of the dictionary. To combat this, we adopt a statistical inference procedure (which is sparse Bayesian learning – SBL) to obtain near accurate AoDs and AoAs which improves the NMSE performance of our proposed algorithm. By analyzing the training dataset, we observed that the channel is sparse both in the virtual beamspace and delay domains. We exploited the beamspace sparsity, and obtained the frequency domain channel estimates using MLGS and SBL. In this final step, we exploit the delay domain sparsity also, and denoise the channel in the delay domain that reduces the mean square error between the original and estimated channels further.

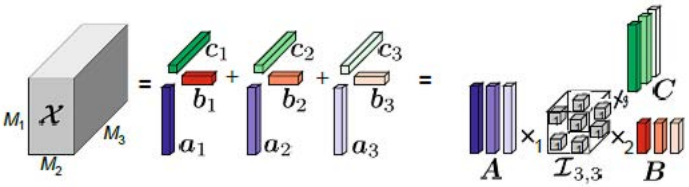


Illustration of Tensor Decomposition

In [KurSlo20e], we look at joint sparse signal estimation and dictionary learning using approximate Bayesian inference schemes. In our previous works, we had inferred that mean field approximation algorithms does not provide accurate posterior variance estimation. So an trade off exists between complexity and performance for different variational Bayesian inference based schemes such as mean field (MF), belief propagation (BP) and expectation propagation (EP). Inspired from the multi-dimensional signal model in massive/mmWave MIMO OFDM systems, the resulting channel can be written as the Kronecker product of several factor matrices (which are representative of antenna array response matrices, path delay response, doppler response etc). In this work, using a combination of BP-MF-EP, we propose an efficient (in terms of complexity reduction and accurate reconstruction performance) dictionary learning and sparse signal estimation algorithm. Further, in [KurSlo20f], we look at more structured factor matrices (which form the measurement matrix). In fact, we observe that original approximate message passing based algorithms (xAMP) need not converge for any measurement matrix. However, our proposed generalized swept AMP (SwAMP-SBL) based algorithm (sequential updates in AMP) has better convergence properties compared to generalized AMP algorithms in the literature. The advantage of SwAMP-SBL is that it does not require any damping factor (as in GAMP and this factor cannot be found in closed form) to force convergence.

Another prominent theoretical contribution on SBL was about the posterior variance prediction of the sparse state vector computed by BP or xAMP algorithms. In the literature, it is shown that the posterior mean squared error (MSE) predicted by xAMP converges to true posterior MMSE. However, state of the art ignores the convergence of per component posterior MSE values. We fill in this gap in our work [KurSlo20g], where we show that the posterior MSE predicted by BP or xAMP based SBL algorithms converge to the true MMSE values (in the case of i.i.d entries in the measurement matrix). For this purpose, we utilize the random matrix theory based large system analysis results from our group in 2012. Moreover, it is worth mentioning that the posterior variance analysis for non-i.i.d sparse state vector is done for the first time in the literature for xAMP based algorithms.

5.6.2 Optical Beamforming for Massive MIMO

Optical fiber has revolutionized the communication systems as it offers high bandwidth, immunity to electromagnetic interference and tapping, flexibility and resistance to corrosive materials. Initially, optic fibre's main purpose was to allow long-distance communication, which was very challenging wirelessly due to attenuation, but later it enabled numerous applications such as imaging, sensing fiber lasers, etc. Another initially unforeseen but nowadays essential is the interaction of optics and microwaves, which has numerous applications for radar, communication systems, warfare systems and instrumentation. This area is well known as microwave photonics and allows photonic generation and distribution, processing and monitoring of microwave signals, and photonic-assisted analog to digital conversion. Recently, the application of optic components to replace the RF circuitry for hybrid beamforming (BF) has caught the attention. Optically beamformed phased-array antennas offer many advantages over their electrical counterparts such as small size, low weight, no susceptibility to electromagnetic interference and wide instantaneous bandwidth and squint-free array steering (true-time delay (TDD)). These two last features are very interesting for broadband wireless access (BWA) networks, especially those operating in the millimeter-wave band, which are envisaged to fulfil the growing demand for bandwidth requirements of the current and the future multimedia services and applications. It has many advantages in terms of bandwidth, circuit complexity, size, weight, and processing speed over conventional techniques. The advantage of having TDD reduce the beam squint phenomena and therefore can improve the performance of a communication network considerably.

In the context of Huawei project OPTIBEAM, we have presented two deliverables on optical beamforming for the next generation of wireless communications. In the first deliverable, firstly, we illustrated the characteristics of photodiodes photodetectors and nonlinear materials, which serves as an interaction medium between the optical to the mmWave/THz domain. Then the state-of-the-art about the TTD and phase control techniques, which are typically used to implement an optical beamforming network is presented. Finally, numerous optical beamforming architectures are presented, and their advantages are extensively discussed. Also, experimentally validated optical beamforming designs are presented, which help to understand the optical beamforming performance gain in real-world applications. In the second deliverable, a survey on the most prominent architectures available in the state-of-the-art and a few architectures on Hybrid beamforming with the analog beamformer implemented optically which the multi-beam capability are discussed. Then, interesting research directions are identified, vastly improving hybrid beamforming designs' performance and getting the maximum achievable performance very close to a fully digital beamforming design.

5.7 Massive MIMO Signal Processing Techniques

5.7.1 Self-Interference Cancellation and Beamforming for Full Duplex Communications

Full-duplex (FD) has to potential to double the performance of a wireless communication system as it allows simultaneous transmission and reception in the same frequency band. It avoids using two independent channels for bi-directional communication, by allowing more flexibility in spectrum utilization, improving data security, and reducing the air interface latency and delay issues. Self-Interference (SI) is a significant challenge to deal with to achieve FD operation, which could be around 110 dB compared to the received signal of interest. However, continuous advancement in the SI cancellation (SIC) techniques has made the FD operation feasible. SIC schemes split the workload into the passive, analog and digital domain. The most challenging stage of SIC is the analog SIC stage, for which extra hardware is required. The analogue-to-digital-converters (ADCs) have only limited dynamic range (LDR), and if the analog SIC stage fails to mitigate the SI sufficiently, it leads to saturation of the converters. Saturation noise is well-known to be the most challenging noise, which can significantly limit the performance of a FD system. Also, the other non-ideal circuity in the transmit and receive chains limit ideal SIC. Therefore, for correct performance analysis of a FD system, the effect of RF circuitry by using the LDR model must be considered.

SIGNAL RECONSTRUCTION IN FD SYSTEMS

As the SI is severe, some SIC is required at the antenna level or in the analog domain before analog to digital conversion (ADC) in the receiver chain because the ADC has only a limited dynamic range. Especially in classical receiver design in which all signals are attempted to be represented without ADC saturation. In [SheSlo20] we consider deliberately scaling up the received signal, provoking ADC saturation due to the SI signal. This leads to missing samples which we propose to reconstruct under the assumptions that the receive signal of interest is a low pass bandlimited signal with known spectrum (mask), oversampling and (perfect) digital SIC after the ADC. The missing samples are estimated by fixed lag Kalman smoothing. More upscaling leads to fewer available samples but with less quantization noise. Hence an optimum compromise arises. We provide an approximate resulting MSE analysis based on large random matrix theory, replacing randomly selected Fourier transform vectors by vectors of i.i.d. variables. Simulation results show the improvements in reconstruction Signal toNoise Ratio (RSNR) and the optimal compromise behavior.

BEAMFORMING IN FD SYSTEMS

In [KurSlo20a], we look at a more robust beamforming design under partial CSIT. We propose an alternating minorization based approach for optimizing an upper bound of the expected weighted sum rate (EWSR), which is tight in a certain massive MIMO limit. The minorization approach also involves user stream power optimization, which implicitly determines the number of supportable streams for a user. This follows from the interference leakage aware based power optimization (per stream). It is worth mentioning that derivation of the BFs using the alternating minorization approach in [KurSlo20a] is not trivial compared to the EWSR based designs for half-duplex system. This is due to the SI power components in the rate expressions. This indeed required us to assume that higher-order terms involved in the LDR noise components are negligible. To the best of our knowledge, this is the first work that looks at a multi-stage/HBF design under partial CSIT using a more practical LDR model. The simulation results in [KurSlo20a] show that using an analog combiner stage at Rx (which operates before the Rx side LDR noise) has better sum rate performance compared to using a two stage BF at Tx side for SI nulling. Also, it demonstrates the huge improvement in spectral efficiency compared to the state of the art designs which ignore the LDR noise effect for the BF design.

It has been shown that the non-ideal behaviour of the transmit and receive chains limits non-ideal SIC and therefore, the performance reported in most of the research papers is not obtainable in a practical FD system. The non-ideal behaviour of the transmit chains dominates the LDR noise variance, especially the power amplifier. Its non-linear behaviour leads to inaccurate SI channel estimation and limits the maximum achievable of a FD system compared to the traditional half-duplex system. However, the performance can be improved by setting the per-antenna power constraints together with the sum-power constraints. Compared to the traditional sum-power constraints, the joint sum-power constraints consider the total transmission power limit and the physical power consumption limits imposed on each power amplifier. The per-antenna power constraints can also be flexibly fixed to limit the non-ideal behaviour of the transmit chains to reduce the transmit LDR noise variance. As a consequence, we can improve the uplink rate by sacrificing the downlink rate slightly. In practice, an optimal tradeoff between the downlink and the uplink rate should be investigated before flexibly fixing the per-antenna power

constraints. In [SheSlo21a], we propose the first-ever digital beamforming design for a MIMO FD communication under the joint sum-power and per antenna power constraints. In [SheSlo21b], the concept is extended to the hybrid beamforming of the FD MIMO devices operating in the millimeter wave band. In [SheSlo21b], we present two novel hybrid beamforming algorithms for a single cell mmWave Massive MIMO FD system with multi-antenna half-duplex users. We jointly design the hybrid beamformer at the FD base station and the digital beamformer for all the uplink users and show the drastic effect of the non-ideal hardware on the maximum achievable performance of a FD system operating in a millimeter-wave band. We report that the FD system can operate even worse at high LDR noise than an HD system.

5.8 Decentralized Cooperative Networks

BOUNDS AND STRATEGIES IN COOPERATIVE NETWORKS

Network-wide interference management via cooperation among geographically distributed transmitters (TXs) is a well-established paradigm, also known as multi-cell multiple-input multiple-output (MIMO) or network MIMO, with the potential of overcoming current cellular systems limitations. However, its feasibility has been questioned for several practical scenarios where more stringent backhaul and/or feedback constraints severely limit accurate CSIT distribution across the network, giving rise to a so-called distributed CSIT configuration. Despite its relevance, research on the distributed CSIT assumption is far from being mature, and most of the past works are based on asymptotic analysis.

In [MirKobGesKer20], we establish the capacity region of MIMO fading multiple-access channels where (distributed) CSIT is acquired via over-the-uplink feedback links of different rates. Linear precoding is proven to be capacity-achieving, with the caveat that an unconventional number of data streams may be needed. As a byproduct, we show that the use of additional data streams allows so solve the otherwise intractable capacity computation problem via standard convex optimization tools. Building on the above result, [MirKobGes20] discusses the problem of joint feeback and precoding design in the presence of asymmetric feedback rate constraints. Furthermore, [MirGes20] applies similar techniques for interference managamente in a network with 2 TXs and 2 RXs, confirming in the finite SNR regime the benefits of accurate distributed precoding design already unveiled by asymptotic SNR analysis.

5.9 Coded Caching in Wireless Networks

Within the context of coded caching, our work in [LamEli20] reveals the interesting connection between having multiple transmitters and having heterogeneity in the cache sizes of the receivers. Our work effectively shows that having multiple transmit antennas – while providing full multiplexing gains – can also simultaneously completely remove the performance penalties that are typically associated to cache-size unevenness. Focusing on the multiple- input single-output Broadcast Channel, the work first identifies the performance limits of the extreme case where cache-aided users coincide with users that do not have caches, and then expands the analysis to the case where both user groups are cache-aided but with heterogeneous cache-sizes. In the first case, the main contribution is a new algorithm that employs perfect matchings on a bipartite graph to offer full multiplexing as well as full coded-caching gains to both cache-aided as well as cache-less users. An interesting conclusion is that, starting from a single- stream centralized coded caching setting with normalized cache size g, then adding L antennas allows for the addition of up to approximately L/g extra cache-less users, at no added delay costs.

In [JouLamEli20] we consider cache-aided wireless communication scenarios where each user requests both a file from an a-priori generated cacheable library (referred to as `content'), and an uncacheable `non-content' message generated at the start of the communication session. We focus our investigation on single-transmitter wireless networks with cache-aided receivers, where the wireless channel is modelled by a degraded Gaussian broadcast channel (GBC). For this setting, we study the (normalized) delay-rate trade-off, which characterizes the content delivery time and non-content communication rates that can be achieved simultaneously. We propose a scheme based on the separation principle, which isolates the coded caching problem from the physical layer transmission problem, and prove its information-theoretic order optimality up to a multiplicative factor of 2.01. A key insight emerging from our scheme is that substantial amounts of non-content traffic can be communicated while maintaining the minimum content delivery time.

In our work [ZhaBazEli21]-[ZhaBazEli21], we address the worst-user bottleneck of coded caching, which is known to diminish any caching gains due to the fundamental requirement that the multicast transmission rate should be limited by that of the worst channel among the served users. We consider the quasi-static Rayleigh fading Broadcast Channel, for which we first

show that the coded caching gain of the XOR-based standard coded-caching scheme completely vanishes in the low-SNR regime. Yet, we show that this collapse is not intrinsic to coded caching by presenting a novel scheme that can completely recover the caching gains. The scheme exploits an aspect that has remained unexploited: the shared side information brought about by the file size constraint. The worst-user effect is dramatically ameliorated because it is replaced by the worst-group-of-users effect, where the users within a group have the same side information and the grouping is decided before the channel or the demands are known.

In our work in [BazEli21], we resolve the cache-aided variant of the famous ``PN'' problem, in the regime where the number of users providing perfect CSI is smaller or equal to the number of transmit antennas. In particular, we derive the optimal rate-memory trade-off under the assumption of uncoded placement, and characterize the same trade-off within a factor of 2.01 for general placement. The result proves that the ``PN'' impact remains similar even in the presence of side information, but also that the optimal trade-off is not achievable through serving independently the two sets of users.

Multi-antenna cache-aided wireless networks have been known to suffer from a severe feedback bottleneck, where achieving the maximal Degrees-of-Freedom (DoF) performance required feedback from all served users. These costs matched the caching gains and thus scaled with the number of users. In the context of the L-antenna MISO broadcast channel with K receivers having normalized cache size g, our work in [LamBazEli20] pairs a fundamentally novel algorithm together with a new information-theoretic converse, and identifies the optimal tradeoff between feedback costs and DoF performance, by showing that having CSIT from only C<L served users, implies an optimal one-shot linear DoF of C+K*g. As a side consequence of this, we also now understand that the well known DoF performance L+K*g, is in fact exactly optimal. In practice, the above means that we are now able to disentangle caching gains from feedback costs, thus achieving unbounded caching gains at the mere feedback cost of the multiplexing gain. This further solidifies the role of caching in boosting multi-antenna systems; caching now can provide unbounded DoF gains over multi-antenna downlink systems, at no additional feedback costs. The above results are extended to also include the corresponding multiple transmitter scenario with caches at both ends.

Coded caching in large MIMO systems: The work in [ParEliLam20] considers the cache-aided multiple-input single-output broadcast channel (MISO BC) where an L-antenna transmitter serves K receiving users, each assisted by one of Q <= K caches with normalized capacity g. Our work provides a novel coded caching scheme that achieves the exact best known, near optimal, DoF L+Kg, and does so even if L > Kg, thus covering an important hole in identifying the optimal performance for the multi-antenna shared-cache problem. Therefore, our work reveals that shared-cache systems with many transmit antennas can also enjoy both full multiplexing gains (L) as well as full caching gains (Kg) despite the sharing of the caches. A side benefit of this scheme is its applicability in multi-antenna settings with dedicated users caches, where it can offer the advantage of reducing the subpacketization without sacrificing the DoF performance.

Then in [SalParSha20] we present an entirely novel caching scheme, termed cyclic multi-antenna coded caching, whose unique structure allows for the resolution of the above bottlenecks in the crucial regime of many transmit antennas. For this regime, where the multiplexing gain can exceed the coding gain, our new algorithm is the first to achieve the exact optimal DoF with a subpacketization complexity that scales only linearly with the number of users, and the first to benefit from a multicasting structure that allows exploiting uplink-downlink duality in order to yield ultra-fast optimized beamformers. In the end, our novel solution provides excellent performance for networks with finite SNR, finite file sizes, and many users.

Caching over uneven channels: In [LamZhaSim20] we identify the fundamental limits of cache-aided coded multicasting in the presence of the well-known 'worst-user' bottleneck. This stems from the presence of receiving users with uneven channel capacities, which often forces the rate of transmission of each multicasting message to be reduced to that of the slowest user. This bottleneck, which can be detrimental in general wireless broadcast settings, motivates the analysis of coded caching over a standard Single-Input-Single-Output (SISO) Broadcast Channel (BC) with K cache-aided receivers, each with a generally different channel capacity. For this setting, we design a communication algorithm that is based on superposition coding that capitalizes on the realization that the user with the worst channel may not be the real bottleneck of communication. We then proceed to provide a converse that shows the algorithm to be near optimal, identifying the fundamental limits of this setting within a multiplicative factor of 4. Interestingly, the result reveals that, even if several users are experiencing channels with reduced capacity, the system can achieve the same optimal delivery time that would be achievable if all users enjoyed maximal capacity.

5.10 UAV-aided Wireless Mesh Networks

UAV RELAYS FOR MESH NETWORKS

Mesh networks are known to provide enhanced and robust coverage by leveraging the multi-hop relaying and self-organization capabilities. Despite these advantages, in deployment scenarios where some nodes are severely obstructed from others, overall network connectivity may still be hampered. In this work [EsrGanGes20], we investigate the use of an unmanned aerial vehicle (UAV) serving as a smart relay to improve the connectivity in a wireless mesh network. It is the first contribution of its kind in the context of mesh networks where an UAV autonomously navigates itself to maximize the mesh connectivity based on the positioning algorithm that exploits the radio measurements collected in the network. We also validate the performanc of the developed algorithm in a real-life experimental setup.

AUTONOMOUS UAV PATH PLANNING WITH REINFORCEMENT LEARNING

Autonomous deployment of unmanned aerial vehicles (UAVs) supporting next-generation communication networks requires efficient trajectory planning methods. In [BayThe20], we propose a new end-to-end reinforcement learning (RL) approach to UAV-enabled data collection from Internet of Things (IoT) devices in an urban environment. An autonomous drone is tasked with gathering data from distributed sensor nodes subject to limited flying time and obstacle avoidance. By intelligently preprocessing and exploiting a multi-layer map of the environment fed through convolutional network layers to the agent, we show that our proposed network architecture enables the agent to make movement decisions for a variety of scenario parameters that balance the data collection goal with flight time efficiency and safety constraints. In contrast to previous work, the trained agent can adapt to a change in the location of IoT devices, maximum UAV flight time, or other scenario parameters, without requiring any retraining. In [BayThe20b], we expand our initial results to the case of multi-UAV cooperation with multi-agent reinforcement learning (MARL). By using dual global-local map processing, we also show that our proposed method can be extended to large map and state spaces. In [TheBay20] and [TheBay20b], we show that a classical robotics problem, coverage path planning, where the UAV's goal is to cover all points inside an area of interest, can be formulated in an analogue way to IoT data harvesting and solved by the same method.

6. LEARNING FOR WIRELESS NETWORK

6.1 Wireless-aided learning

WIRELESS DISTRIBUTED EDGE LEARNING

We consider distributed machine learning at the wireless edge, where a parameter server (PS) builds a global model with the help of multiple wireless edge devices that perform computations on local dataset partitions. Edge devices transmit the result of their computations (updates of current global model) to the server using a fixed rate and orthogonal multiple access over an error prone wireless channel. In case of a transmission error, the undelivered packet is retransmitted until successfully decoded at the receiver. In [SoKou20a], [SoKou20b] we aimed at answering the following question: how many edge devices do we need in a distributed edge learning system operating under latency constraints? If an edge device is added to the system, computing load will be shared; thus, the computation time per device will decrease. In contrast, an additional edge device will have to be allocated wireless resources to transmit its intermediate results. However, since wireless resources are limited and do not usually scale with the number of participating devices, as the number of edge devices increases, available resources per edge device are reduced. Also, in non-orthogonal multiple access, higher interference is generated when more edge devices transmit to the PS simultaneously. As a consequence, increasing the number of edge devices can lead to longer communication time. Leveraging on the fundamental tradeoff between computation and communication in distributed systems, we derived the optimal number of edge devices to minimize the average completion time while guaranteeing convergence and a certain accuracy.

6.2 Learning-aided wireless

In the context of wireless networking, it was recently shown that multiple Deep Neural Networks can be jointly trained to offer a desired collaborative behaviour capable of coping with a broad range of sensing uncertainties. While promising, a major challenge in the implementation of such method is that information noise statistics may differ from agent to agent and, more importantly, that such statistics may not be available at the time of training or may evolve over time, making burdensome retraining necessary. This situation makes it desirable to devise a "universal" machine learning model, which can be trained once for all so as to allow for decentralized cooperation in any future feedback noise environment. In [ZecGes20] we propose an architecture inspired from the well-known Mixture of Experts (MoE) model, which was previously used for non-linear regression and classification tasks in various contexts, such as computer vision and speech recognition. We consider the decentralized power control problem as an example to showcase the validity of the proposed model and to compare it against other power control algorithms. We show the ability of the so called Team-DMoE model to efficiently track time-varying statistical scenarios.

6.3 Tools for learning

SUBMODULARITY IN ACTION: FROM MACHINE LEARNING TO SIGNAL PROCESSING APPLICATIONS

Submodularity is a discrete domain functional property that can be interpreted as mimicking the role of well-known convexity/concavity properties in the continuous domain. Submodular functions exhibit strong structure that lead to efficient optimization algorithms with provable near-optimality guaran-tees. These characteristics, namely, efficiency and provable performance bounds, are of particular interest for signal processing (SP) and machine learning (ML) practitioners, as a variety of discrete optimization problems are encountered in a wide range of applications. Conventionally, two general approaches exist to solve discrete problems: 1) relaxation into the continuous domain to obtain an approximate solution or 2) the development of a tailored algorithm that applies directly in the discrete domain. In both approaches, worst-case performance guarantees are often hard to establish. Fur-thermore, they are often complex and thus not practical for large-scale problems. In this work, we show how certain scenarios lend themselves to exploiting submodularity for constructing scalable solutions with provable worst-case performance guarantees. We introduce a variety of submodular-friendly applications and elucidate the relation of submodularity to convexity and concavity, which enables efficient optimization. With a mixture of theory and practice, we present different flavors of submodularity accompanying illustrative real-world case studies from modern SP and ML. In all of the cases, optimization algorithms are presented along with hints on how optimality guarantees can be established [TohAmi20].

7. EXPERIMENTAL ACTIVITIES

7.1 OpenAirInterface Software Alliance

The interest in the software is constantly increasing and 2020 was a very good and active year for the OAI software alliance (OSA). We delivered the first version of the 5G NR non-standalone version of OAI and did a successful demonstration of a 5G connection with a commercial off-the-shelf (COTS) phone in summer 2020.

Further 3 new strategic partners Qualcomm, Interdigital, Facebook Connectivity joined the existing consortium of Orange, Fujistu and PAWR. Two partners TCL and Nokia Bell Labs left the rank of strategic partners. Strategic members left: TCL, Nokia Bell Labs.

The OSA also received a record number of donations in 2020, which lead to the creating of the OAI 5G RAN and OAI 5G CN project groups. These groups now actively contribute to the development of OAI on top of the existing CI/CD team.

3 new engineers were hired to support the CI/CD team and 1 engineer was hired on the 5G CN project group. The Alliance has also started to subcontract some of the development work for 5G standalone mode (part of the RAN project group) to Allbesmart, a company based in Portugal.

Another milestone is the change of license of OAI MME and HSS to 3-clause BSD license and integration of these projects into the MAGMA project, which is an open-source project lead by Facebook that gives network operators an open, flexible and extendable mobile core network solution[1].

Last but not least, OSA organized a virtual workshop on November 12th, 2020, which was attended by more than 350 people.

7.2 Open5GLab

Open5GLab is EURECOM's experimental platform activity based on the OpenAirInterface open source software and others.

Open5GLab at EURECOM is one of 3 experimental 5G sites in France in the context of the 5G-EVE ICT-17 project. Construction began in July 2018 and 5G experimentation is now available. The site is interconnected with similar sites in Europe in the 5G-EVE network. It is also one of the test sites for the OPNFV VCO 3.0 (Virtual Central Office) project and as such is interconnected with sites in North Carolina, USA and Montreal, Canada.

Open5gLab provides experimental 5G services including so-called Enhanced Mobile Broadband (eMBB) and massive machine-type communications and is based on fully open-source tools and open-architecture design. It is the main experimental playground for OpenAirInterface (OAI) and Mosaic-5g (M5G) software packages. The site's cluster computing resource makes use of RedHat's OpenShift 4.2 Kubernetes container platform and benefits from technical support from RedHat. The cluster is used for radio-access, core network and mobile-edge functions. Some bare-metal nodes with in-lab 5G-capable radio devices can be used by experimenters and developers and are interconnected with the Kubernetes cluster. External access for onboarding software, collecting measurement data and developing basic software for the site is available for partners using secure-shell access. Interfaces for an external orchestrator (e.g. ONAP) is currently being integrated in the Open5GLab. The nodes of the site are also used by OpenAirInterface Jenkins-based continuous integration / continuous delivery (CI/CD) framework.

Open5GLab's radio infrastructure includes indoor and high-power outdoor radio-units operating in several 4G and 5G bands in the immediate vicinity of the test site, specifically Band 28 (700 MHz), Band n38 (2.6 GHz TDD), Band n78 (3.5 GHz TDD), Band n258 (25 GHz TDD). The outdoor units are interconnected with the switching fabric using 300m fiber (10/25 Gbit/s). The units are a combination of in-house designs and commercial remote radio-units.

[1] https://www.magmacore.org/

Open5GLab provides remotely-controllable 4G and 5G user-equipment, including both off-the-shelf smartphones and cellular IoT modules. This allows experimenters to control and extract measurements from the user-equipment in a running experiment. With the help of regional partners, Open5GLab will provide embedded vehicular user-equipment nodes in 2020. Two drones are also equipped with 4G and soon 5G user-equipment and OAI-based 4G radios for mobile basestation experiments. With the help of EURECOM, software can be on-boarded into the user-equipment devices.

The Open5GLab site provides the means to on-board new network functions to a running 5G infrastructure and test them in both a controlled laboratory setting and in a deployed live network. Experimenters can either make use of existing network functions and basic applications or choose to on-board their own software to test with Open5GLabs infrastructure. This allows scientists to focus on their function of interest and its interconnection with a full network and collect measurements of the new design. In addition, if the network function enhances either the OAI or M5G implementation, it will automatically be reusable by a global network of researchers using OAI/M5G since it will be redistributable with OAI/M5G.

OAI and M5G software packages are used extensively around the globe. The Open5GLab site allows this community to have access to the laboratory used by the main developers of OAI and M5G. It is thus a crucial infrastructure for testing OAI and M5G in both laboratory and outdoor deployments. Since the site is used in OAI's CI/CD framework the entire OAI/M5G communities will be using the site automatically when committing code.

Figure 7: Base Station Equipment on EURECOM's roof

7.2.1 5G NR Implementation in OpenAirInterface

The OpenAirInterface team has been working on an implementation of 5G-NR since 2017 with a focus on the eMBB use case. A first pre-5G demonstration was given at Mobile World Congress February 2018 where we had shown the feasibility of maintaining a throughput of 300Mbps over a 80MHz channel using the new 5G-NR LDPC channel coding. This was achieved by offloading the most computationally expensive task, the LDPC decoder, to an FPGA. Using a software-only implementation of the LDPC decoder, a throughput of 150Mbps can be supported.

Since this initial demo, we have started to follow the 3GPP 5G-NR specifications, whose first version was released in March 2018. This first version uses the non-standalone architecture (NSA) option 3GPP, also called EUTRA-NR dual connectivity (EN-DC). In this option, the 5G cell is operating under the control of a 4G cell, which serves as an anchor to the system and carries all control plane traffic. UEs first need to connect to the 4G network and will receive all the necessary configuration to connect to a 5G cell through RRC signaling on the 4G link. This setup will allow a smooth migration from 4G to 5G.

A first demonstration of OpenAirInterface 5G NSA with a commercial off-the-shelf (COTS) phone was achieved in summer 2020. Since then the implementation has continuously improved and is now able to support a stable uplink and downlink end-to-end traffic.

The system has also been integrated with other radio heads such as the ones from Benetel and AW2S. The Benetel system is the first one to use the O-RAN 7.2 fronthaul split, which is today becoming the industry standard for open RAN deployments. The evaluation system from Benetel is powered by OpenAirInterface and is commercially available.

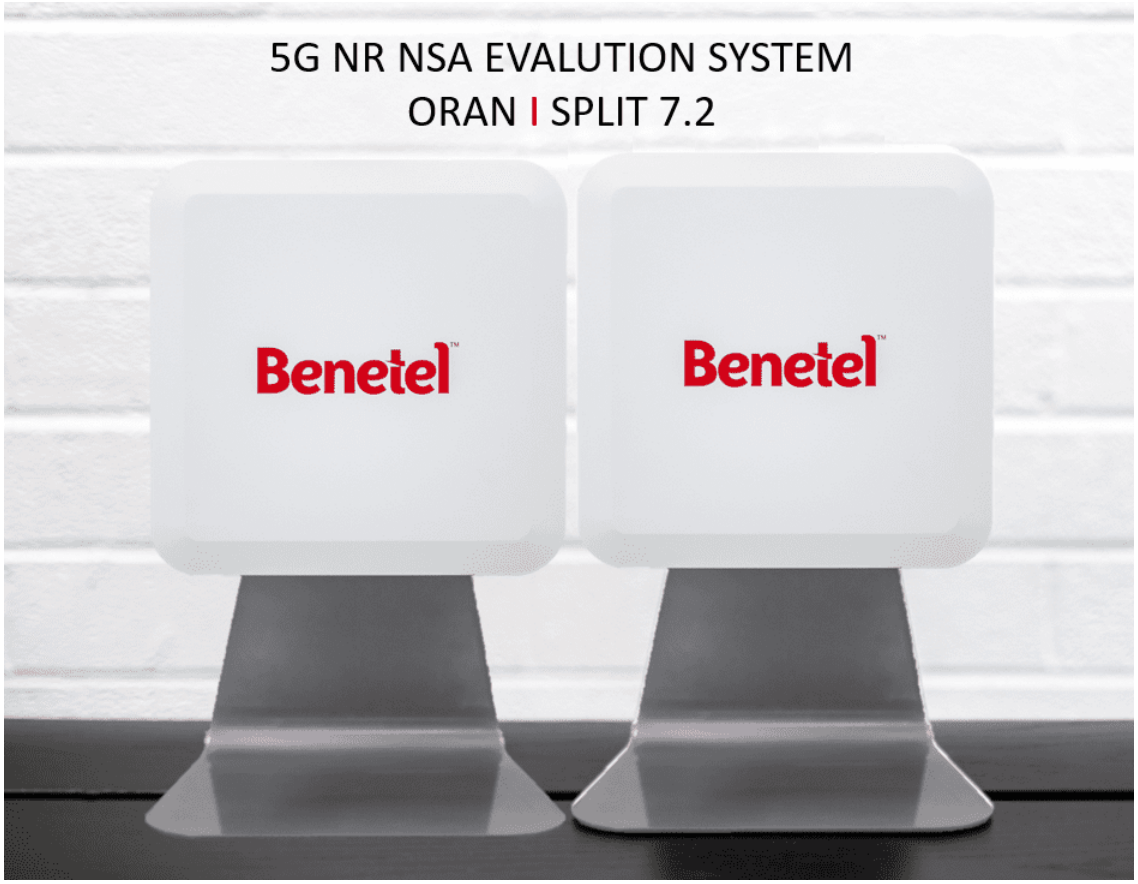


Figure 8: Benetel 5G NR NSA Evaluation system based on O-RAN 72 split

In 2020 we have also started implementing the necessary functions in the 5G gNB to also support standalone mode. By the end of 2020 we already have most of the functionality available, including the 5G core network (see next section). However, interoperability testing has proven difficult due the lack of availability of devices.

7.2.2 5G Core Implementation in OpenAirInterface

As a part of the OpenAirInterface project, the basic components for 5GC are under development. At this stage, we have a functional 5GC consisting of four main components as AMF, SMF, UPF and NRF. For the moment, OAI 5GC supports some basic procedures such as UE registration, de-registration, service request, and the session management procedures including session establishment, modification and release. These functions have been validated with a professional tester, namely dsTest.

For the roadmap, first we will continue to introduce additional features to the existing components such as IPv6 support, handover procedures, paging, etc. We also plan to integrate our 5GC with OAI gNB and test with a real UE. Next, we will implement new entities as UDM, AUSF, NEF, UDR, etc. We will also work on the performance evaluations of our 5GC in terms of number of supported UEs, throughput, network processing delay, etc. By the end of this project, we should have a full stand-alone 5G core implementation and a complete deployment framework for a micro-services based architecture. For more information regarding the current status of OAI 5GC, please refer to our website: https://openairinterface.org/oai-5g-core-network-project/.

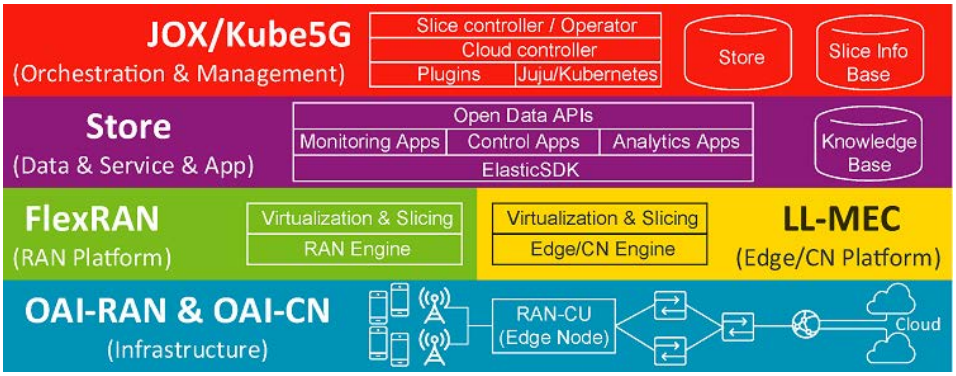
7.3 Mosaic-5G.IO Activities

M5G is an ecosystem of opensource platforms and use-cases for 4G-5G R&D to build a lightweight 5G service delivery platform across reusable software components. Mosaic-5G is formed to develop, promote, and share an ecosystem of open-source platforms and use cases for 5G system research and development leveraging software-defined networking (SDN), network function virtualization (NFV), and multi-access edge computing (MEC) technology enablers [NkCyKa-18].

Mosaic-5G is composed of the following submodules:

1. The network Store: a repository of platform packages, software development kits (SDKs), network control applications and data sets;
2. The JOX juju orchestrator: event-driven orchestrator core with plugins to interact with different network domains;
3. Kube5G – which is a cloud-native agile 4G and 5G service platform which is built on the top of openshift Operator, Kubernetes and docker technologies;
4. The LL-MEC low latency MEC platform: it's an ETSI-aligned MEC platform that can act as software-defined core network controller;
5. The FlexRAN for real-time controller providing flexible and programmable platform to enable software-defined radio access network;
6. The OAI-RAN for monolithic and disaggregated RAN, delivered as a software packages and cloud images
7. The OAI-CN for monolithic and disaggregated CN, delivered as a software packages and cloud images

This repository is designed to serve as a meta repository from where each submodule can be pulled.



The current software is considered as a complementary platform to OpenAirInterface (OAI) for 5G services bridging the gap between radio/core network infrastructure and network virtualization/control. In May 2018, EURECOM's Mosaic5G initiative to build a multi-stakeholder community that will allow collaboration and interaction across the members. EURECOM has already welcomed more than 80 Mosaic5G academia and industrial members, highlighting Davidson Consulting, Interdigital, Wipro & Rutgers' University. More information about Mosaic5G associate members is given in the following link: http://mosaic-5g.io/members/ The current memberships could lead to both potential financial and contribution partnerships at a business level. To this end, EURECOM plans to reinforce the synergy with other open-source initiatives (as OpenAirInterface) and trigger the collaboration and interactions across the community.

7.3.1 Demonstrations and Events

The following demonstration and events were carried out:

- NOMS Demo, 5G cloud-native: Network management and automation, April 2020.
- NOMS Demo, Efficient multi-service RAN slice management and orchestration, April 2020.
- MobiCom Demo, Service-oriented intelligent and extensible RAN, September 2020.
- Globecom Demo, Towards 5G cloud-native: Automation and flexible network aggregation, December 2020.
- Mosaic5G, 1-Day Workshop, "Agile 5G Network Services and Applications", on Thursday 2020-12-03, Link to the Workshop Program

The workshop event got quite a big audience, up to 260 participants; sessions were quite interactive, and the general feedback was very positive.

All the materials shared during the workshop have been made available publicly on M5G website as well as social media platforms including LinkedIn and Youtube.

Here is a brief description of the activities held online: From its inception in 2016 till now, the Mosaic5G stack has made it possible to deliver 4G and more recently 5G on as-a-service basis leveraging technology enablers such as SDN, NFV/cloud-native, and MEC technology, prototyped on the top of the OpenAirInterface 4G/5G software platforms. This workshop will be a great opportunity to share experiences and visions on agile and data-driven 4G/5G network services and applications with the community.

To this end, this workshop aims to inform the community about the latest achievments in the Mosaic5G ecosystem and its roadmap for future development and target use-cases. The workshop will also provide a training sessions and live use-case demonstrations on four platforms, namely FlexRAN, LL-MEC, Kube5G, and Store.

The program of the workshop was divided in two parts, the theoretical and the practical, respectively – the first of which, aimed at informing participants about advances in current explored topics and projects whereas the second of which, aimed at performing some live trainings based on the four main components of the Mosaic5G stack:

1. Network Store
2. Kube5G
3. Flex-RAN
4. LL-MEC

8. NEW COLLABORATIVE PROJECTS STARTED IN 2020

8.1 5G-RECORDS

5G-RECORDS aims to explore the opportunities which new 5G technology components bring to the professional production of audiovisual content, including Programme Making and Special Events (PMSE). 5G-RECORDS targets the development, integration, validation and demonstration of 5G components for professional content production, as part of an overall ecosystem integrating a subset of 5G network functions. The project will take a business-to-business (B2B) perspective, where 5G becomes integrated in media creation workflows including audio-visual infrastructure backbones. The challenge is to use 5G components from previous 5G-PPP projects and earlier R&D investments and further develop them. These components will be evaluated in three end-to-end 5G infrastructures provided by the project. These include the core network (5GC), radio access network (RAN) and end devices. The project aims to use of non-public networks (NPNs) as a way to bring these new components to emerging markets and new market actors. NPNs can be deployed as independent and standalone 5G networks or in conjunction with a public network. The project also aims to address recent emerging remote and distributed production workflows where cloud technologies cooperate with 5G. To ensure the successful demonstration of these use cases, the project brings together a set of experienced partners whose expertise covers both 5G and content production value chains. Most of the 5G components will be developed and implemented by specific key partners, being mostly innovative SMEs. The 5G-RECORDS implemented technologies will be based on existing 3GPP Rel-15 and Rel-16 specifications, while also prototyping some emerging Rel-17 capabilities. All technology components are expected to reach by the end of the project a minimum Technology Readiness Level (TRL) of 7. 5G-RECORDS has considered 3 use cases to embrace some of the most challenging scenarios in the framework of professional content production: live audio production, a multi-camera wireless studio and live immersive media production.

5G-RECORDS participants are Universitat Politecnica de Valencia (Project Coordinator), Spain, Union Européenne de Radio Télévision (EBU) , Switzerland, British Broadcasting Corporation (BBC), UK, RAI Radiotelevisione Italiana, Italy, Sennheiser Electronic (Germany), Nokia (Spain), Ericsson (Germany), Telefonica Investigacion y Desarrollo (Spain), Cumucore (Finland), Accelleran (Belgium), RED Technologies (France), Image Matters (Belgium), LiveU (Israel), FIVECOMM (Spain), EURECOM (France), Universidad Politécnica de Madrid (Spain).

EURECOM leads a workpackage on trials and provides one of the three experimentation sites. This will make use of open5glab infrastructure to test innovative low-latency real-time audio processing applications from Sennheiser. The EURECOM facility will be interconnected with radio-access network components from Acceleran and Core Network components from Cumucore.

8.2 LeadingEdge

The LeadingEdge project will deliver a novel and holistic framework to efficiently cope with unresolved challenges in edge computing ecosystems, regarding dynamic resource provisioning to multiple coexisting services amidst unknown service- and system-level dynamics. The project approach is three-faceted; it will optimize intra-service resource provisioning, inter-service resource coordination, and user perceived quality of experience (QoE).

8.3 Affordable5G

Affordable5G aims at creating a 5G network that will deliver a complete and affordable solution covering the needs of private and enterprise networks through technical innovation that span across all parts of 5G network, leveraging cell densification, RU/DU/CU split, hardware acceleration, edge computing and core network virtualisation, seamlessly combined with the adoption of open source RAN, MEC and MANO solutions, for cloud-native, micro-service based deployments. To achieve its innovative and ambitious goal, the consortium brings together ten European SMEs, supported by MVNOs, system integrators and research institutes, grasping the opportunity to enhance their products, according to each company's roadmap, while fostering collaboration among them. Affordable5G will offer a first-class opportunity to European SMEs to become frontrunners in the global 5G competition, facilitating them in their commercialisation paths.

8.4 SAFE4Rail3 (H2020 JU SHIFT2RAIL)

Safe4RAIL-3 is the follow up project from Safe4RAIL-2 and aims at bringing the research and innovations from it to prototypes. Similarly, to Safe4RAIL-2, the OAI platform has been chosen for prototyping LTE V2X communication adapted to railway requirements.

8.5 IntellIoT (H2020 NG-IOT)

IntellIoT is a Pan-European project focusing on the development of integrated, distributed, human-centered and trustworthy IoT frameworks applicable to healthcare, manufacturing and agriculture. Enabling technologies such as 5G, cybersecurity, distributed technology, Augmented Reality and tactile internet, IntellIoT also champions end-user trust, adequate security and privacy by design.

The OAI platform will bring the 5G and Beyond 5G technology and innovation supporting decentralized data-driven services for three use cases: Smart Factory, Smart Agriculture and Smart Health.

9. ONGOING PROJECTS AND PROJECTS FINISHED IN 2020

9.1 ERC Projects

ERC Consolidator Project (DUALITY)

DUALITY is 5 year ERC funded project which started May 2017. This work explores the theoretical foundations of transforming memory into data rates, and to explore their practical ramifications in wireless communication networks.

Motivated by the long-lasting open challenge to invent a communication technology that scales with the network size, we have recently discovered early indications of how preemptive use of distributed data-storage at the receiving communication nodes (well before transmission), can offer unprecedented throughput gains by surprisingly bypassing the dreaded bottleneck of real-time channel-feedback. For an exploratory downlink configuration, we unearthed a hidden duality between feedback and preemptive use of memory, which managed to doubly-exponentially reduce the needed memory size, and consequently offered unbounded throughput gains compared to all existing solutions with the same resources. This was surprising because feedback and memory were thought to be mostly disconnected; one is used on the wireless PHY layer, the other on the wired MAC.

This development prompts our key scientific challenge which is to pursue the mathematical convergence between feedback-information-theory and preemptive distributed data-storage, and to then design ultra-fast memory-aided communication algorithms, that pass a battery of tests for real-life validation.

This is a structurally new approach, which promises to reveal deep links between feedback information theory and memory, for a variety of envisioned wireless-network architectures of exceptional promise. In doing so, our new proposed theory stands to identify the basic principles of how a splash of memory can surgically alter the informational structure of these networks, rendering them faster, simpler and more efficient. In the end, this study has the potential to directly translate the continuously increasing data-storage capabilities, into gains of wireless network capacity, and to ultimately avert the looming network-overload caused by these same indefinite increases of data volumes.

DUALITY combines several fields of theoretical expertise:

- Caching
- Information and communication theory
- Combinatorial designs
- Discrete mathematics
- Analysis
- Large deviations, etc.

The results of DUALITY are reported earlier in the section dedicated to communication theory under the decentralized coordination setting.

9.2 French and European Collaborative Projects

9.2.1 ECOLOGICAL-BITS-AND-FLOPS (ANR Jeune Chercheur)

"COMPLEXITY AND BIDIRECTIONAL INFORMATION THEORY: COMPLEXITY-FEEDBACK-PERFORMANCE LIMITS AND A NEW CLASS OF ECOLOGICAL INFORMATION NETWORKS"

This project, which finished in 2020, aimed to tackle the bottleneck of computational complexity corresponding to the need for algorithms that require extreme computing resources, and the bottleneck of feedback corresponding to the need for equally idealistic feedback mechanisms that must disseminate massive amounts of overhead information about the fluctuating states of each link in the network.

9.2.2 MASS-START (FUI)

The project ran from 2017 to 2020 as a collaboration between EURECOM, TCL, Orange, Syrtem, SDRF, and the Université Cote d'Azur. The project's goal was to develop 5G massive MIMO systems for both sub-6GHz (frequency range 1 –FR1) and mm-wave (frequency range 2 – FR2). Different hardware platforms and antenna modules were developed for both frequency ranges. Eurecom's role in the project is mainly the 5G-NR software development and massive MIMO algorithm development.

From EURECOM's point of view the project produces a number of high quality publications [KuSl20a] - [KuSl20e] as well as contributed significantly to the development of the FR2 features of OpenAirInterface. These include the basic physical layer features such as the subcarrier spacing of 120 kHz (compared to 30 kHz in FR1) as well as the beamforming and beam-management techniques. In the final demonstration we showed these features together by using a setup based on National Instruments USRP plus a mm-wave radio head from Interdigital. We validated the OpenAirInterface implementation by using this setup for both gNB and UE. In 2021 we will replace the OAI UE with a commercial off-the-shelf UE to validate the interoperability with commercial equipment.

9.2.3 5GENESIS (H2020 ICT-17)

In 2020, the activities of the project have progressed towards providing the final version of the 5G experimentation facility which targets different vertical use cases and KPIs of interest, across the 5 platforms participating in the project. This progress covers aspects such as: finalizing the common facility architecture and the automation/experimentation framework, continue integrating/updating the individual components across the different platforms according to the architecture, provide a portable demonstrator for dissemination activities.

EURECOM has participated in these activities by providing the software and partially the hardware for the OAI 5G RAN (gNB and UE) and 4G Core network components, the required support to the platforms in order to automate launching of these components, as well as monitoring tools developed for use with OAI to keep track of the RAN procedures and evaluate the performance. In this context, Eurecom developed first the so-called "noS1 mode" which allowed to perform some testing with IP traffic injected over 5G radio using the OAI gNB and UE components and without any connection to the core network. This setup has been used from Malaga, Athens and portable demonstrator platforms. As a next step, Eurecom has started collaborating with Malaga platform for the integration of the OAI 5g NSA setup using the OAI gNB with commercial (COTS) UEs and core network. This integration work will be finalized at the beginning of 2021 and then it is envisioned that it will be used within the portable demonstrator as well.

OAI Components/ 5GENESIS platforms	*Malaga*	*Athens and portable demonstrator*	*Limassol*
noS1 setup with OAI 5G gNB and OAI 5G UE components	x	x	
OAI 4G Core Network		x	x
Ongoing integration of 5G NSA setup with OAI gNB and 5G COTS UE	x		

TABLE 1 SUMMARY OF OAI COMPONENTS USED WITHIN DIFFERENT PLATFORMS OF 5GENESIS

9.2.4 GEOLOC (FUI)

Geolocation estimation in (Narrowband) IoT systems. This project kicked off in January 2018 but was put on hold in the spring because some SMEs did not satisfy some rules on company finances. An opportunity was given to replace these partners with other acceptable and compatible partners, which was explored and successfully reached. The renewed consortium had a kick-off meeting of the restarted project in Jaunary 2020.

9.2.5 EFIGI

EFIGi aims at innovating in the area of new compression formats through contributions to the MPEG standardization process and through real-world characterization of algorithms on platforms, notably the infrastructure based on OpenAirInterface at EURECOM. Part of the tests will include evaluation of the efficiency of the implementation of the novel algorithms in a rate-distortion sense in comparison to state-of-the-art approaches (such as HEVC) and to include energy efficiency constraints. The project also contributes to the Gree MPEG project within the MPEG community. EFIGI was completed in December 2020.

9.2.6 5GENESIS (H2020 ICT-17)

In 2019, the first integration cycle of the project has been completed successfully, by providing five platforms that have managed to integrate major components from the reference architecture and have clearly defined the steps towards composing the full 5GENESIS facility during the next two years of the project.

EURECOM's role in this project is to provide the 5G-NR UE implementation assuring the interoperability with commercial equipment. However, some of the platforms are also relying on the infrastructure components of OpenAirInterface (eNB, gNB, EPC) as their 5G solution and have already integrated them in their platforms.

9.2.7 5G-EVE

5G-EVE offers experimental platforms to vertical industries and to all 5GPPP Phase3 projects facilities to validate their network KPIs and their services.

Important representatives of these vertical industries are directly involved as partners of 5G-EVE exactly to influence the design of the end-to-end 5G services, and to provide an early assessment. The 5G-EVE end-to-end facility consists of the interconnection of four 5G-site-facilities (France, Spain, Italy, Greece), which have been selected because of their considerable previous work with vertical industries and standardisation bodies, on top of their 5G technology competences. 5G-EVE aims at creating synergies between a significant number of facilities that will ensure sustainability and impact in terms of exploitation.

The 5G-EVE facility will enable experiments with:

- heterogeneous access, including NR, licensed/unlicensed spectrum, advanced spectrum management;
- Mobile Edge Computing, backhaul, core/service technologies;
- means for site-interworking and multi-site/domain/technology slicing/orchestration.

5G-EVE will be initially compliant with 3GPP Rel. 15 and, later on, with Rel. 16.

Industrial verticals will be facilitated in the specification/analysis of experiments through:

- intent-based, and other high-level, interfaces;
- means for advanced 5G testing,

(i.e., for KPI analysis, technology bench-marking, performance diagnosis: a VNF pool, including open source and proprietary, radio/network/service, components will be developed and made available.)

5G-EVE will impact standards, and has the potential and strategy for ensuring the sustainability of the facility beyond the project lifetime, therefore becoming a cornerstone of the 5G PPP programme and beyond.

9.2.8 5G-Victori (H2020 5GPPP Phase 3)

5G-VICTORI will conduct large scale trials for advanced vertical use case verification focusing on Transportation, Energy, Media and Factories of the Future and cross-vertical use cases. It leverages 5G network technologies developed in 5G-PPP phase 1 and 2 projects and exploits extensively existing facilities interconnecting main sites of all ICT-17 infrastructures i.e. 5G-VINNI, 5GENESIS and 5G-EVE and the 5G UK test-bed in a PanEuropean Infrastructure. The main software platforms used will be OpenAirInterface and Mosaic-5G stack.

9.2.9 5G-CROCO (H2020 5GPPP Phase 3)

5G CRoss-border COntrol (5G-CROCO) aims at trialling 5G technologies in the Metz-Merzig-Luxembourg cross-border corridor, traversing the borders between France, Germany and Luxembourg. The objective is to validate advanced 5G features in the cross-border context, such as New Radio (NR), Mobile Edge Computing (MEC), distributed computing, predictive Quality of Service (QoS), Software Defined Networking (SDN), Network Slicing, and improved positioning systems, all combined together, to guarantee that innovative use cases for CCAM can be enabled. EURECOM will extend its IoT platform and the geoservice developed in CONCORDA in a MEC architecture, with the particular challenge to support URLL cross-border, cross-operator information exchange.

9.2.10 SAFE4Rail2 (H2020 JU SHIFT2RAIL)

Safe4RAIL-2 and its activities will be based on the development of technological pillars aimed at improving the capabilities of the Train Control and Management System (TCMS). In particular, Safe4RAIL-2 will development LTE devices suitable for the deployment of a Wireless Train Backbone (WLTB), along with analyses including wireless consist communications (notably LTE-V2X), virtual coupling and 5G technology. The OAI platform has been chosen as prototyping platform for Safe4RAIL.

9.2.11 EMPOWER (H2020 ICT-21)

EMPOWER has the ambition to accelerate the joint development between the EU and the US of advanced wireless platforms targeting the new connectivity frontiers beyond 5G. EMPOWER targets the creation of a joint EU-US advanced wireless ecosystem for (i) bridging the relevant EU-US Wireless communities and stakeholders, such as scientific researchers, platform engineers, standardization experts, regulators, and product incubators; and (ii) developing a strategic EU-US collaboration agenda and supporting its execution ahead of worldwide competition for beyond 5G connectivity standards. EMPOWER foresees twinning with the best researchers and practitioners involved in projects funded by USA, especially with entities participating in the PAWR programme (https://www.advancedwireless.org/). EMPOWER will provide instruments for inducing collaboration between ongoing and forthcoming 5G and beyond initiatives targeting at wireless networks experimentation on both ends of the Atlantic. Through the EMPOWER instruments we aim to create an efficient means for stimulating the mobility of ideas and people between European and similar American experimental wireless platform initiatives. We also aim at encouraging stronger collaboration between fundamental and experimental wireless researchers by making access to platform tools and data exchange simpler. EMPOWER instruments will also provide a wealth of information for global and regional standards and regulatory organizations (e.g. ITU-R, ETSI) and industry fora (e.g. NGMN). An important output of EMPOWER will also be in the form of recommendations on technologies and experimentation methodologies for future wireless experimentation objectives. This will assist in providing coordination between EU (FP9) and US NSF programmes for future individual and joint calls.

9.2.12 Toyota Research - EURECOM

EURECOM has a recurring research contract (since 2018) grant from Toyota Research Labs in the US to work on future Knowledge-centric Vehicular Communication and Networks. The objective is to develop the mechanisms for retrieving, exchanging and disseminating knowledge in future networks of automated vehicles and infrastructures.

9.2.13 DUPLEX: Multi-Antenna Full Duplex Radio for Future Wireless Systems

EURECOM is project leader for this ANR project which began in 2016. The project concerns full-duplex communications which describes communication systems where transmission and reception are concurrent on the same time and frequency resources. Which in theory allows an increase in the capacity of communication.

The DUPLEX project objectives are:

- to study the theoretical limits (throughput) of full duplex communication equipment
- to develop antenna techniques, analog and digital processing for the cancellation of the transmitted signal at the receiver
- to develop a full duplex communication equipment (prototype) for the next generation of communications.

The project address two scenarios:

- full-duplex communication between two communication nodes. In this case, the signal cancellation device uses the knowledge of signal, assumed to be known
- full-duplex relaying, in which the relay processes, amplifies and retransmits the received signal in the same band. In this scenario, several cases will be considered depending on the information to be relayed (i.e. decodable or not).

The project is divided into 5 tasks:

- Task 1: scenario refinement.
- Task 2: digital cancellation of self-interference.
- Task 3: analog techniques, including the design and implementation of antennas and circuits.

- Task 4 (prototyping) is dedicated to the integration of all the hardware and software parts developed (in Tasks 2 and 3) in the hardware platform OpenAirInterface (www.openairinterface.org), Furthermore, this task also aims to test and validate the system in laboratory conditions.
- Finally, task 5 (dissemination, communications) aims to promote the scientific results of the project and to disseminate the reusable results for future industrialization.

The DUPLEX project work finished in 2019 but the last deliverables and reports have been finalized in 2020.

9.2.14 Other ongoing collaborative funded projects

GEOLOC (FUI) Geolocation estimation in (Narrowband) IoT systems. This project kicked off in January 2018 but was put on hold in the spring because some SMEs did not satisfy some rules on company finances. An opportunity was given to replace these partners with other acceptable and compatible partners, which was explored and successfully reached. The renewed consortium had a kick-off meeting of the restarted project in Jaunary 2020.

10. PUBLICATIONS

[AroNikGlobeP20] O. Arouk, N. Nikaein, “Kube5G: A cloud-native 5G service platform”, GLOBECOM 2020, IEEE Global Communications Conference, Taipei, Taiwan, 2020.

[AroNikGlobeD20] O. Arouk, N. Nikaein, “Towards 5G Cloud-Native: Automation and Flexible Network Aggregation”, GLOBECOM 2020, IEEE Global Communications Conference, Taipei, Taiwan, 2020.

[AroNikNoms20] O. Arouk and N. Nikaein, "5G Cloud-Native: Network Management & Automation," NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium, Budapest, Hungary, 2020, pp. 1-2, doi: 10.1109/NOMS47738.2020.9110392.

[BazSlo20] Ahamad Bazzi, Dirk Slock, “ Robust Music Estimation Under Array Response Uncertainty” IEEE Int’l Conf. Acoustics, Speech and Signal Processing (ICASSP), 4-8 May 2020, Barcelona, Spain

[BazEli21] A. Bazco-Nogueras and P. Elia, "Rate-Memory Trade-Off for the Cache-Aided MISO Broadcast Channel with Hybrid CSIT," 2020 IEEE International Workshop on Information Theory (ITW), Riva del Garda, Italy, 2021, to appear.

[BakBou20] “Channel stability prediction to optimize signaling overhead in 5G networks using machine learning”, S. Bakri, M. Bouaziz, P. Frangoudis, A. Ksentini, ICC 2020, IEEE International Conference on Communications, 7-11 June 2020, Dublin, Ireland (Virtual conference).

[BayThe20] H. Bayerlein, M. Theile, M. Caccamo, D. Gesbert, “Multi-UAV Path Planning for Wireless Data Harvesting with Deep Reinforcement Learning”, in Proc. IEEE Global Communications Conference (GLOBECOM), Taipei, Taiwan, 7-11 December 2020.

[BayThe20b] H. Bayerlein, M. Theile, M. Caccamo, D. Gesbert, “UAV Path Planning for Wireless Data Harvesting: A Deep Reinforcement Learning Approach”, arXiv:2010.12461 [cs.MA], 2020.

[BeKsa21] A. Bemani, N. Ksairi, and M. Kountouris, “Diversity of Affine Frequency Division Multiplexing over Doubly Dispersive Channels”, IEEE ICC 2021 Workshops.

[BriKse20] “Federated learning for UAVs-enabled wireless networks: Use cases, challenges, and open problems”, B. Brik A. Ksentini; M. Bouaziz, IEEE Access, Vol.8, N°1, December 2020.

[BriFra20] “Service-oriented MEC applications placement in a federated edge cloud architecture”, B. Brik, P. Frangoudis , A. Ksentini, ICC 2020, IEE International Conference on Communications, 7-11 June 2020, Dublin, Ireland (Virtual conference)

[BriKse20] “On predicting service-oriented network slices performances in 5G: A federated learning approach” 45th IEEE Conference on Local Computer Networks, 16-19 November 2020, Sydney, Australia (Virtual Conference)

[CodHae20a] L. Codeca, J. Erdmann, V. Cahill, J. Härri, “SAGA: An activity-based multi-modal mobility scenario generator for SUMO”, SUMO 2020, International SUMO User Conference, Simulating Connected Urban Mobility, October 26-28, 2020, Berlin, Germany

[CodHae20b] L. Codeca, J. Erdmann, V. Cahill, J. Härri, “Comparing Multi-Modal Traffic Assignments in Large-Scale Simulations using the Macroscopic Fundamental Diagram: Mode Shift from Cars to Powered Two Wheeler”, 7th International IEEE Conference on Models and Technologies for Intelligent Transportation Systems, 2021.

[CosSpy20a] M. Costantini, T. Spyropoulos, T. Giannakas, P. Sermpezis, "Approximation guarantees for the joint optimization of caching and recommendation", In ICC 2020-2020 IEEE International Conference on Communications (ICC) (pp. 1-7), June 2020

[CosSpy20b] M. Costantini and T. Spyropoulos, "Impact of Popular Content Relational Structure on Joint Caching and Recommendation Policies", In 2020 18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOPT) (pp. 1-8), June 2020.

[DatHae20a] A. Kousaridas, A. Schimpe, Andreas, S. Euler, X. Vilajosana, M. Fallgren, G. Landi, F. Moscatelli, S. Marmpounakis, F. Vázquez-Gallego, R. Sedar, R. Silva, L. Dizambourg, S. Wendt, M. Muehleisen, K. Eckert, J. Härri, J. Alonso-Zarate, “5G cross-border operation for connected and automated mobility: Challenges and solutions”, Future Internet, 2020, Vol.12, N°5

[DatHae20b] R. Muñoz, F. Vazquez-Gallego, R. Casellas, R. Vilalta, R. Sedar, P. Alemany, R. Martinez, J. Alonso-Zarate, A. Papageorgiou, M. Catalan-Cid, X. Vilajosana, A. Bartoli, D. Guilhot, F. Moscatelli, F. G. Landi, R. Silva, L. Dizambourg, A. Fernandez, S. Datta, J. Härri, M. Muehleisen, “5GCroCo Barcelona trial site for cross-border anticipated cooperative collision avoidance”, EUCNC 2020, European Conference on Networks and Communications, 15-18 June 2020, Dubrovnik, Croatia (Virtual Conference).

[DebJma20] “Algorithmics and modeling aspects of network slicing in 5G and beyonds network: Survey”, F. Debbabi, R. Jmal, L. Fourati, L. Chaari, A. Ksentini, IEEE Access, Vol.8, 7 September 2020.

[DigHae20] G. Di-Giacomo, “Data-Driven Road Hazard Detection for Automated Driving”, Master Thesis, EURECOM, 2020, publication under preparation.

[DunHae20a] D. Deveaux, J. Härri, “On the Traffic Impact of Risk-Aware Autonomous Vehicles Rerouting Strategies”, 3rd Symposium on Management of Future Motorway and Urban Traffic Systems, 2020.

[DunHae20b] D. Deveaux, T. Higuchi, S. Uçar, J. Härri, O. Altintas, “A definition and framework for vehicular knowledge networking”, ArXiV, 29 May 2020, submitted to IEEE Vehicular Technology Magazine (under review).

[DunHae20c] S. Uçar, T. Higuchi, D. Deveaux, J. Härri, O. Altintas, “Vehicular knowledge networking and application to risk reasoning”, D2VNet 2020, ACM MobiHoc Workshop on Cooperative data dissemination in future vehicular networks, in conjunction with MOBIHOC 2020, 2020, Boston, MA, USA.

[DunHae20d] D. Deveaux, T. Higuchi, S. Uçar, J. Härri, O. Altintas, “On the orchestration of federated learning through vehicular knowledge networking”, VNC 2020, IEEE Vehicular Networking Conference, 16-18 December 2020.

[EsrGan20] O. Esrafilian, R. Gangula and D. Gesbert, "3D Map-based Trajectory Design in UAV-aided Wireless Localization Systems," in IEEE Internet of Things Journal, 2020.

[EsrGanGes20] O. Esrafilian, R. Gangula and D. Gesbert, "Autonomous UAV-aided Mesh Wireless Networks," IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), Toronto, ON, Canada, 2020

[FatLie20] “AutoMEC: LSTM-based user mobility prediction for service management in distributed MEC resources”, U. Fattore, M. Liebsch, B. Brik, A. Ksentini, MSWiM 2020, 23rd International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, 16-20 November 2020, Alicante, Spain.

[GiaGioSpy21] T. Giannakas, A. Giovanidis, T. Spyropoulos, “SOBA: Session optimal MDP-based network friendly recommendations”, accepted in IEEE Conference of Computer Communications (INFOCOM), May 2021

[GiaFra20] “Orchestrating heterogeneous MEC-based applications for connected vehicles”, F. Giannone, P. Frangoudis, A. Ksentini, L. Valcarenghi, Computer Networks, Vol.180, 24 October 2020

[GhaSlo20] Ghamnia Imene, Dirk Slock and Yuan-Wu Yi, “MIMO user rate balancing in multicell networks with per cell power constraints”, In Proc. IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, 2020

[GhoCot20a] Gholami, Roya, Laura Cottatellucci, and Dirk Slock. "Favorable propagation and linear multiuser detection for distributed antenna systems." IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020.

[GhoCot20b] Gholami, Roya, Laura Cottatellucci, and Dirk Slock. "Channel Models, Favorable Propagation and Multistage Linear Detection in Cell-Free Massive MIMO." IEEE International Symposium on Information Theory (ISIT). IEEE, 2020.

[GopSlo20] Gopala, K. ; Slock, D., “Reciprocity-based beamforming techniques for spectrum sharing in MIMO networks”, Book Chapter N°9, of "Spectrum Sharing: The Next Frontier in Wireless Networks", Wiley, April 2020.

[Hae20a] A. Arriola, M. Münzer, A. Geven, J. Pujol, H. Envall, J. Härri, N. Mata, R. Aigner, “Safe4RAIL-2: Advanced architectures and components for the next-generation train control and monitoring system”, TRA 2020, 8th Transport Research Arena, April 27-30, 2020, Helsinki, Finland

[Hae20b] J. Manco, G. Gallud Baños, J. Härri, M. Sepulcre, “Prototyping V2X applications in large-scale scenarios using OpenAirInterface” ,VNC 2020 IEEE Vehicular Networking Conference, 16-18 December 2020.

[Hae20c] A. Molinar, C. Campolo, J. Härri, C-E Rothenberg, A. Vinel, “5G-V2X communications and networking for connected and autonomous vehicles”, Special issue of "Future Internet" (ISSN 1999-5903); Guest Editors: Antonella Molinaro, Claudia Campolo, Jérôme Härri, Christian Esteve Rothenberg, Alexey Vinel, 2020.

[Hae20d] J (Paul) Jeong, Yiwen (Chris) Shen, Tae (Tom) Oh, S. Céspedes, N. Benamar, M. Wetterwald, J. Härri, “A comprehensive survey on vehicular networks for smart roads: A focus on IP-based approaches”, Elsevier Journal on Vehicular Communications, January 2021

[JouLamEli20] H. Joudeh, E. Lampiris, P. Elia and G. Caire, "Fundamental Limits of Wireless Caching Under Mixed Cacheable and Uncacheable Traffic," 2020 IEEE International Symposium on Information Theory (ISIT), Los Angeles, CA, USA, 2020, pp. 1693-1698, doi: 10.1109/ISIT44484.2020.9173941.

[KurSlo20a] Kurisummoottil Thomas Christo, Slock Dirk, “Rate maximization under partial CSIT formulti-stage/hybrid BF under limited dynamic range for OFDM full-duplex systems”, IEEE 91st Vehicular Technology Conference (VTC-Spring), 25-28 May 2020, Antwerp, Belgium.

[KurSlo20b] Kurisummoottil Thomas Christo, Slock Dirk, “Massive MIMO Stochastic Geometry and Analysis of Beamforming with Partial CSIT",2nd 6G Wireless Summit, March 2020, Levi, Lapland, Finland.

[KurSlo20c] Kurisummoottil Thomas Christo, Bruno Clerckx, Luca Sanguinetti, Slock Dirk, “A Rate Splitting Strategy for Mitigating Intra-Cell Pilot Contamination in Massive MIMO”, IEEE International Conference on Communications Workshops (ICC Workshops), 7-11 June 2020, Dublin, Ireland.

[KurSlo20e] Kurisummoottil Thomas Christo, Slock Dirk, “BP-VB-EP Based Static and Dynamic Sparse Bayesian Learning with Kronecker Structured Dictionaries”, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4-8 May 2020, Barcelona, Spain.

[KurSlo20f] Kurisummoottil Thomas Christo, Slock Dirk, “Generalized Swept Approximate Message Passing based Kalman Filtering for Dynamic Sparse Bayesian Learning”, IEEE 28th European Signal Processing Conference (EUSIPCO), 18-21 Jan 2021, Virtual Conference.

[KurSlo20g] Kurisummoottil Thomas Christo, Slock Dirk, “Posterior Variance Predictions in Sparse Bayesian Learning under Approximate Inference Techniques”, IEEE Asilomar Conference on Signals, Systems, and Computers, 1-5 Nov 2020, Virtual Conference.

[KhaHae20a] I. Khan, S. Sesia, J. Härri, “Enhancing ETSI DCC for multi-service vehicular safety communication”, VTC 2020-Fall, IEEE 92nd Vehicular Technology Conference, 2020.

[KhaHae20b] I. Khan, M. Sepulcre, J. Härri, “Cooperative wireless congestion control for multi-service V2X communication”, IV 2020, 31st IEEE Intelligent Vehicles Symposium, 2020.

[KhaHae20c] I. Khan, S. Sesia, J. Härri, “Method and apparatus for improving reactive congestion control algorithms for vehicular communications”, Patent, June 2020.

[Knopp20a] Knopp, Raymond, “DEVICES AND METHODS FOR LINEAR ERROR CORRECTION CODES CONSTRUCTION”, European Patent Application, number EP20305929.0, Aug. 2020

[Knopp20b], Knopp, Raymond, “DEVICE AND METHOD FOR TRANSMITTING CONTROL DATA WITH LOW PEAK-TO-AVERAGE POWER RATIO”, European Patent Application, number EP20306231.0, Oct. 2020

[Knopp20c], Knopp, Raymond, “IMPROVED DEVICE AND METHOD FOR TRANSMITTING CONTROL DATA WITH LOW PEAK-TO-AVERAGE POWER RATIO”, European Patent Application, EP20306230.2, Oct. 2020

[Knopp20d], Knopp, Raymond, “R1-2006880- Limitations of NR Short Block-Length Codes for PUCCH Coverage Enhancement”, 3GPP TSG-RAN WG1 Meeting #102 e-Meeting, August 17th – August 28th, 2020

[Knopp20e], Knopp, Raymond, “R1-2008938/R1-20089451- Low-PAPR Sequence-Based Approaches for PUCCH Coverage Enhancement”, 3GPP TSG-RAN WG1 Meeting #103 e-Meeting, October 26th – Nov 13th, 2020, 2020

[KseBri20] “An edge-based social distancing detection service to mitigate COVID-19 propagation”, A. Ksenitni, B. Brik, IEEE Internet of Things Magazine, Vol.3, N°3, September 2020

[KseFra20] “Toward slicing-enabled multi-access edge computing in 5G”, A. Ksentini, P. Frangoudis, IEEE Network Magazine, Vol.34, N°2, March 2020.

[KsenFra20] “On extending ETSI MEC to support LoRa for efficient IoT application deployment at the edge”, A. Ksentini, P. Frangoudis, IEEE Communications Standards Magazine, Vol.4, N°2, June 2020.

[LamEli20] E. Lampiris and P. Elia, "Full Coded Caching Gains for Cache-Less Users," in IEEE Transactions on Information Theory, vol. 66, no. 12, pp. 7635-7651, Dec. 2020, doi: 10.1109/TIT.2020.3015348.

[LamZhaSim20] E. Lampiris, J. Zhang, O. Simeone, P. Elia, "Fundamental Limits of Wireless Caching under Uneven-Capacity Channels". Accepted for publication in International Zurich Seminar (IZS) 2020.

[LamBazEli20] E. Lampiris, A. Bazco-Nogueras and P. Elia, "Resolving the Feedback Bottleneck of Multi-Antenna Coded Caching," 2020, submitted to IEEE Transactions on Information Theory, arxiv: https://arxiv.org/pdf/1811.03935.pdf

[LeSal20a] Le Trung-Kien, Salim Umer, Kaltenberger Florian, “Feedback Enhancements for Semi-Persistent Downlink Transmissions in Ultra-Reliable Low-Latency Communication", EUCNC, European Conference on Networks and Communications, 2020.

[LeSal20b] Le Trung-Kien, Salim Umer, Kaltenberger Florian, "Channel Access Enhancements in unlicensed spectrum for NR URLLC transmissions", GLOBECOM, 2020.

[LeSal20c] Le Trung-Kien, Salim Umer, Kaltenberger Florian, "Enhancing URLLC Uplink Configured-grant Transmissions", VTC Spring, 2021.

[LeSal20d] Le Trung-Kien, Salim Umer, Kaltenberger Florian, “An overview of physical layer design for Ultra-Reliable Low-Latency Communications in 3GPP Releases 15, 16, and 17” IEEE Access 2020.

[LeSal20e] Le Trung Kien, Salim Umer, Kaltenberger Florian, "Performance enhancement of PUSCH repetition method in wireless communication systems", US patent application No. 63/061979, 2020.

[LeSal20f] Le Trung Kien, Salim Umer, Kaltenberger Florian, "Efficient scheduling in a wireless communications system", US patent application No. 63/061970, 2020.

[LeRos20] Le, Hoang-Minh; Rossi, Jean-Pierre; Slock, Dirk TM, “A geometric interpretation of trilateration for RSS-based localization”, 28th European Signal Processing Conference (EUSIPCO 2020), Jan. 18-22, 2021, Amsterdam, The Netherlands.

[MacBecSpy20] P. Mach, Z. Becvar, T. Spyropoulos, “Incentive-based D2D Relaying in Cellular Networks,” accepted in IEEE Transaction on Communications, Nov. 2020

[MagKal20] Magounaki, Theoni; Kaltenberger, Florian; Knopp, Raymond, “Modeling the distributed MU-MIMO OAI 5G testbed and group-based OTA calibration performance evaluation,” SPAWC 2020, 21st IEEE International Workshop on Signal Processing Advances in Wireless Communications, 26-29 May 2020, Atlanta, GA, USA (Virtual Conference).

[MalSerPar20b] A. Malik, B. Serbetci, E. Parrinello, and P. Elia, "Fundamental Limits of Stochastic Shared Caches Networks". Submitted to IEEE Transactions on Communications. Online version available at https://arxiv.org/abs/2005.13847.

[MalSerPar20] A. Malik, B. Serbetci, E. Parrinello, and P. Elia, "Stochastic Analysis of Coded Multicasting for Shared Caches Networks," 2020 IEEE Global Communications Conference (GLOBECOM), Taipei, Taiwan, December 7-11, 2020.

[MesBer20] “Towards the quest for 5G network slicing”, F. Messaoudi, P. Bertin, A. Ksentini, CCNC 2020, IEEE Consumer Communications & Networking Conference, 10-13 January 2020, Las Vegas, USA.

[MirKobGesKer20] L. Miretti, M. Kobayashi, D. Gesbert, and Paul de Kerret, “Cooperative Multiple-Access Channels with Distributed State Information,” submitted to IEEE Trans. Inf. Theory, preprint (v3): https://arxiv.org/abs/1911.07899

[MirKobGes20] L. Miretti, M. Kobayashi, and D. Gesbert, “Precoding for Cooperative MIMO Channels with Asymmetric Feedback,” in Proc. IEEE International Conference on Communications (ICC), virtual conference, June 2020

[MirGes20] L. Miretti and D. Gesbert, “Interference Mitigation for Cooperative MIMO Channels with Asymmetric Feedback,” in 54th Asilomar Conference on Signals, Systems and Computers, virtual conference, Nov 2029

[MurAtz21] P. Mursia, I. Atzeni, M. Kobayashi, and D. Gesbert, “D2D-Aided Multi-Antenna Multicasting under Generalized CSIT,” vol. 0, pp. 1–30, 2021, [Online]. Available: http://arxiv.org/abs/2102.01624.

[MurSci20] P. Mursia, V. Sciancalepore, A. Garcia-Saavedra, L. Cottatellucci, X. Costa-Perez,and D. Gesbert, “RISMA: Reconfigurable Intelligent Surfaces Enabling Beamformingfor IoT Massive Access,”IEEE J. Sel. Areas Commun. (JSAC) Special Issue on Massive Access for 5G and Beyond, Aug. 2020.

[ParEliLam20] E. Parrinello, P. Elia and E. Lampiris, "Extending the Optimality Range of Multi-Antenna Coded Caching with Shared Caches," 2020 IEEE International Symposium on Information Theory (ISIT), Los Angeles, CA, USA, 2020 pp. 1675-1680.

[RicNegSpy20] G. Ricardo, G. Neglia and T. Spyropoulos, "Caching Policies for Delay Minimization in Small Cell Networks with Joint Transmissions," in Proc. of IEEE International Conference on Communications (ICC), Dublin, Ireland, 2020, pp. 1-6

[SerKasPin20] P. Sermpezis, S. Kastanakis, J. I. Pinheiro, et al., “Towards QoE-Aware Recommendations,” in Proc. of CARS (Context-aware Recommendation Systems) workshop of ACM RecSys (Recommendation Systems) conference, 2020.

[SalLe] Salim Umer, Le Trung Kien, "Configurable Uplink transmissions in a wireless communication system", US patent application No. 63/061982, 2020.

[SalParSha20] M. Salehi, E, Parrinello, S. P. Shariatpanahi, P. Elia and A, Tolli, "Low-Complexity High-Performance Cyclic Caching for Large MISO Systems," Submitted to IEEE Transactions on Communications. Online version available at https://arxiv.org/abs/2009.12231

[SerLamSpy20] B. Serbetci, E. Lampiris, T. Spyropoulos, and P. Elia, "Augmenting Multiple-Transmitter Coded Caching using Popularity Knowledge at the Transmitters," 18th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt 2020), Volos, Greece, June 15-19, 2020.

[SchNik20a] R. Schmidt and N. Nikaein, "Demo: Efficient Multi-Service RAN Slice Management and Orchestration", presented at the NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium (virtual conference), Budapest, Hungary, 2020, pp. 1–2.

[SchNik20b] Robert Schmidt and Navid Nikaein, "Service-oriented intelligent and extensible RAN", presented at the 26th Annual International Conference on Mobile Computing and Networking (MobiCom '20) (virtual conference), London, United Kingdm, 2020, pp. 1–3.

[SchNik21] Robert Schmidt and Navid Nikaein, " RAN Engine: Service-Oriented RAN through Containerized Micro-Services", accepted for IEEE Transactions on Network and Service Management, to be published in March 2021.

[SimBou20] "Supporting UAV services in 5G networks: New high-level architecture integrating 5G with U-space", S. Si-Mohammed, M. Bouaziz, H. Hellaoui, O. Bekkouche, A. Ksentini, T. Taleb, L. Tomaszewski, T. Lutz, G. Srinivasan, T. Järvet, P. Montowtt, IEEE Vehicular Technology Magazine, November 2020

[SimKse20] "UAV mission optimization in 5G: On reducing MEC service relocation", S. Si-Mohammed, A. Ksentini, M. Bouaziz, Y. Challal, A. Balla, GLOBECOM 2020, IEEE Global Communications Conference, 7-11 December 2020, Taipei, Taiwan.

[SheSlo20] Chandan Kumar Sheemar, Dirk Slock, "Receiver Design and AGC Optimization with Self Interference Induced Saturation", In Proc. International Conf. on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, 2020.

[SheSlo21a] Chandan Kumar Sheemar, Dirk Slock, "Beamforming for Bidirectional MIMO Full Duplex Under Joint Sum Power and Per Antenna Power Constraints ", In Proc. International Conf. on Acoustics, Speech and Signal Processing (ICASSP), Toronto, 2021.

[SheSlo21b] Chandan Kumar Sheemar, Dirk Slock, "Hybrid Beamforming for Bidirectional Massive MIMO Full Duplex Under Practical Considerations", IEEE Vehicular Technology Conference (VTC-Spring), Helsinki, Finland.

[SheSlo21c] Chandan Kumar Sheemar, Christo Kurisummoottil Thomas, Dirk Slock "Practical Hybrid Beamforming for Millimeter Wave Massive MIMO Full Duplex with Non Ideal Hardware", Open journal of IEEE ComSoc- Special Issue on Full Duplex Wireless Communications Emerging Technologies Initiatives.

[SoKou20a] J. Song and M. Kountouris, "Optimal number of edge devices in distributed learning over wireless channels", in Proc. IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), 2020.

[SoKou20b] J. Song and M. Kountouris, "Wireless distributed edge learning: How many edge devices do we need?" IEEE Journal on Selected Areas in Communications, Series on Machine Learning for Communications and Networks, December 2020, to appear (DOI: 10.1109/JSAC.2020.3041379)

[Slo20a] Papageorgiou, Georgios K; Voulgaris, Konstantinos; Ntougias, Konstantinos; Ntaikos, Dimitrios K; Butt, Majid M; Galiotto, Carlo; Marchetti, Nicola; Frascolla, Valerio; Annouar, Hicham; Gomes, Alvaro; Morgado, Antonio J; Pesavento, Marius; Ratnarajah, Tharmalingam; Gopala, Kalyana; Kaltenberger, Florian; Slock, Dirk TM; Khan, Faheem A; Papadias, Constantinos B, "Advanced dynamic spectrum 5G mobile networks employing licensed shared access", IEEE Communications Magazine, Vol.58, N°7, July 2020

[Slo20b] Rabih Chrabieh, Peter Bagnall, Serdar Sezginer; Dirk Slock, "Advanced TOA estimation for multipath channels" IEEE/ION Position, Location and Navigation Symposium (PLANS), 20-23 April 2020, Portland, USA

[Slo20c] Papadias, C. B. ; Ratnarajah, T. ; Slock, D., eds, "Spectrum Sharing: The Next Frontier in Wireless Networks", Wiley, April 2020.

[Slo20d] Papadias, C. B. ; Ratnarajah, T. ; Slock, D., “Introduction: From cognitive radio to modern spectrum sharing”, Book Chapter N°1, of "Spectrum Sharing: The Next Frontier in Wireless Networks", Wiley, April 2020.

[SubKur20d] Sai Subramanyam Thoota, Christo Kurisummoottil Thomas, Chandra R Murthy, Rakesh Mundlamuri, Sameera Bharadwaj, Marios Kountouris, “Learned Chester: A Multilevel-Greedy and Bayesian Compressive Channel Estimator for Frequency-Selective Hybrid mmWave MIMO Systems”, ITU AI/ML in 5G Conf., Site Specific mmWave Channel Estimation Challenge, 16 Dec.202.

[TalKse20] “On supporting UAV based services in 5G and beyond mobile systems”, T. Taleb, A. Ksentini, H. Hellaoui, O. Bekkouche, IEEE Network Magazine, October 2020.

[TheBay20] M. Theile, H. Bayerlein, R. Nai, D. Gesbert, M. Caccamo, “UAV Coverage Path Planning under Varying Power Constraints using Deep Reinforcement Learning”, in Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 25-29 October 2020.

[TheBay20b] M. Theile, H. Bayerlein, R. Nai, D. Gesbert, M. Caccamo, “UAV Path Planning using Global and Local Map Information with Deep Reinforcement Learning”, arXiv:2010.06917 [cs.RO], 2020.

[TohChe20] E. Tohidi, J. Chen and D. Gesbert, "Sensor Selection for Model-Free Source Localization: where Less is More," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 4831-4835.

[TohGes20] E. Tohidi, D. Gesbert, A. Bazco-Nogueras and P. d. Kerret, "Decentralizing Multi-Operator Cognitive Radio Resource Allocation: An Asymptotic Analysis," ICC 2020 - 2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 2020, pp. 1-6.

[TohAmi20] E. Tohidi, R. Amiri, M. Coutino, D. Gesbert, G. Leus and A. Karbasi, "Submodularity in Action: From Machine Learning to Signal Processing Applications," in IEEE Signal Processing Magazine, vol. 37, no. 5, pp. 120-133, Sept. 2020.

[TohGes21] E. Tohidi, D. Gesbert and P. Ciblat, "Distributed Resource Allocation Algorithms for Multi-Operator Cognitive Communication Systems," 2020 28th European Signal Processing Conference (EUSIPCO), Amsterdam, 2021, pp. 1737-1741.

[TouBer20] “Towards cross-domain service function chain orchestration”, N. Toumi, O. Bernier, D-E. Meddour, A. Ksentini, GLOBECOM 2020, IEEE Global Communications Conference, 7-11 December 2020, Taipei, Taiwan.

[TsiSpy20] D. Tsigkari, T. Spyropoulos, “User-centric Optimization of Caching and Recommendations in Edge Cache Networks”, Proc. IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), Cork, Ireland, 2020.

[ZhaBazEli21] H. Zhao, A. Bazco-Nogueras and P. Elia, "Wireless Coded Caching Can Overcome the Worst-User Bottleneck by Exploiting Finite File Sizes," 2020, submitted to IEEE Transactions on Wireless Communications.

[ZhaBazEli21b] H. Zhao, A. Bazco-Nogueras and P. Elia, "Resolving the Worst-User Bottleneck of Coded Caching: Exploiting Finite File Sizes," 2020 IEEE International Workshop on Information Theory (ITW), Riva del Garda, Italy, 2021, to appear.

[ZecGes20] Zecchin, Matteo, David Gesbert, and Marios Kountouris. "Team Deep Mixture of Experts for Distributed Power Control." 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). IEEE, 2020.

Research activity report 2020

EURECOM
Sophia Antipolis

450 route des Chappes -CS 50 193
F-06904 Sophia Antipolis cedex

Tel. : +33 (0)4 93 00 81 00
Fax : +33 (0)4 93 00 82 00

www.eurecom.fr