EURECOM
Sophia Antipolis

Research activity report 2016

Data Science
Digital Security
Communication systems

EURECOM
CAMPUS SOPHIATECH

GRADUATE SCHOOL AND RESEARCH CENTER IN INFORMATION SCIENCES


EURECOM
Sophia Antipolis

INTRODUCTION BY ULRICH FINGER, DIRECTOR

2016 was marked by changes in the research and teaching areas, particularly the reorganization of our Research & Development Departments.

Further to our 2015 strategic planning, EURECOM moved to restructure our research in anticipation of the ongoing digital transformation of our industry and our society. Specifically, Data Science is clearly a promising research field in which EURECOM has invested considerably these last years. Making it a full-fledged department was therefore a natural step. With this new department, and the closing of the Multimedia Department, we have refocused the Digital Security Department activities and integrated several security elements related to imaging and audio processing. We also expanded those of the Communication Systems Department to fixed networks.

Our belief is that our departments must represent a critical mass in their respective domains and be limited to a few scientific domains while still being able to design systems.

The three departments are as follows:

- Communication Systems Department
- Data Science Department
- Digital Security Department

With this new organization, EURECOM is able to develop a new dynamic for its research while adapting its activities to the new areas of interest suggested by our industrial partners. One of our challenges is to fill the gap between fundamental and applied research to a more business-oriented research in our partner companies. It is not surprising then that over 50% of our joint contracts with industry are with our members.

The first results are already visible: an AXA chair in Machine Learning, the creation of an international alliance for the 4G/5G software radio networks, border control systems through secure image processing, and much more.

This is paired with many publications in top conferences, best paper awards, and many successes in competitive projects, all of which we present in this research overview.

TABLE OF CONTENTS

DIGITAL SECURITY DEPARTMENT

INTRODUCTION

As a result of the recent reorganization of research at EURECOM, all existing activities in the area of computer and communications security have been brought together in the newly created Digital Security Department. By joining forces between the system security and applied cryptography teams from the former networking and security department and the signal processing and biometrics team from the former multimedia department, the new digital security department now features a rare blend of technical and scientific expertise with a near-comprehensive coverage of topic areas in digital security along three directions:

- **System and Software Security**: the basic expertise that lies in malware analysis and detection is further extended to computer forensics and the study of fraud and cybercrime; the main application fields are embedded system security, web security, security for wireless communications and networks.
- **Security Protocols and Applied Cryptography**: the main expertise is in applied cryptography with a strong focus on the design of security protocols and privacy-preserving mechanisms together with secure software design and model-driven engineering for security.
- **Biometrics and Digital Media**: relying on a strong background in digital signal processing, research in biometrics focuses on feature extraction and statistical pattern classification technologies to automate the recognition of individuals using estimates of their biometric characteristics; the main activities are in image forensics, 2D/3D face recognition, speaker diarization and recognition, and soft biometrics.

The department's three research groups that are organized along each of these directions are involved in a number of ongoing or newly accepted H2020 projects, ANR projects and a variety of bilateral contracts with industrial partners. These activities cover a broad variety of topics including emerging subjects such as industrial security, IoT security, security and privacy for cloud computing and Big Data, and privacy and surveillance.

Industrial Security and IoT: relying on the success story of two major milestones such as Avatar, an emulation framework that enables complex dynamic analysis of embedded devices, and firmware.re, a large scale static analysis platform that came up with the finding of a large number of vulnerabilities in more than 30K firmware images, the department is set out for a multi-year project in the field of industrial security whereby the main target of the analysis is industrial automation platforms based on programmable logic controllers and the forthcoming distributed environments using IoT. Apart from vulnerability analysis, IoT will be the target of future research in applied cryptography through the new CHIST-ERA project on user-centric privacy & security in the IoT.

Security & Privacy for the Cloud and Big Data: along with the main focus on the design of customized cryptographic mechanisms for searchable encryption, proof of retrievability, verifiability, part of the activity for cloud security and privacy is on system security aspects such as the vulnerability analysis for virtualization systems. Cloud security and privacy research is conducted as part of two H2020 projects and several bilateral collaborations. Another European project addresses privacy-preserving data collection techniques. An H2020 project proposal aiming at joint approaches for privacy and data analysis for Big Data and incident response techniques based on machine learning are part of the new roadmap.

Biometrics for Surveillance and Privacy: As part of the current drive in biometrics for detecting spoofing attacks, our department is involved in several activities aiming at novel anti-spoofing techniques for audio- and image-based identification, ranging from several collaborative European projects, to partnership with main industries, through the organization of or active participation in various field testing activities such as the ASVSpoof challenge or the Kinect Face Database. Along the applications of this research in surveillance and automated border control, the department participates in a new H2020 project and an ongoing ITEA project whereby new biometrics methods such as soft biometrics are developed in order to guarantee the privacy of individuals while enabling global detection capabilities.

In the sequel of this report the research activities of our department are presented along three directions:

- System and Software Security
- Security Protocols and Applied Cryptography
- Biometrics and Digital Media.

1. SYSTEM AND SOFTWARE SECURITY

(Davide Balzarotti, Aurélien Francillon)

The research conducted by the software and system security group in 2016 can be summarized around four main topics:

1. Software and Operating System Security
2. Spam, Phishing, and Frauds
3. Embedded Systems and Security at the Hardware/Software Interface
4. Web Security

1.1 Software and Operating System Security

A substantial amount of our research focuses on the analysis of new attacks and on the design of defense and detection techniques to protect software and operating systems from compromise attempts.

Our first work in this area in 2016 looked at Control Flow Integrity (CFI) approaches used to protect binary executables. CFI is a very promising defense technique against code-reuse attacks and, while proposals to use hardware features to support CFI already exist, there is still a growing demand for an architectural CFI support on commodity hardware. To tackle this problem, in a paper we presented at the RAID conference [MUE16], we demonstrate that the Transactional Synchronization Extensions (TSX) recently introduced by Intel in the x86-64 instruction set can be used to support CFI. The main idea of our approach is to map control flow transitions into transactions. This way, violations of the intended control flow graphs would then trigger transactional aborts, which constitutes the core of our TSX-based CFI solution. To prove the feasibility of our technique, we designed and implemented two coarse-grained CFI proof-of-concept implementations using the new TSX features. In particular, we show how hardware-supported transactions can be used to enforce both loose CFI (which does not need to extract the control flow graph in advance) and strict CFI (which requires pre-computed labels to achieve a better precision). All our solutions are based on a compile-time instrumentation. We evaluate the effectiveness and overhead of our implementations to demonstrate that a TSX-based implementation contains useful concepts for architectural control flow integrity support.

In parallel to this approach, we also investigated a different technique to protect software against exploitation. In this second work, we focused on detecting (using a custom security hypervisor) signs of heap spraying in an arbitrary process memory. Spraying is a common payload delivery technique used by attackers to execute arbitrary code in presence of Address Space Layout Randomisation (ASLR). Our solution, called Graffiti and presented at the Usenix Security Symposium [CRI16], is an efficient hypervisor-based memory analysis framework for the detection and prevention of spraying attacks. Compared with previous solutions, our system is the first to offer an efficient, complete, extensible, and OS independent protection against all spraying techniques known to date. We developed a prototype open source framework based on our approach, and we thoroughly evaluated it against all known variations of spraying attacks on two operating systems: Linux and Microsoft Windows. Our tool can be applied out of the box to protect any application, and its overhead can be tuned according to the application behavior and to the desired level of protection.

We then turned our attention to the protection of the kernel of an operating system against stealthy rootkit techniques. In fact, modern rootkits have moved their focus on the exploitation of dynamic memory structures, which allows them to tamper with the behavior of the system without modifying or injecting any additional code. In our work [GRA16] we investigated the feasibility to mount a new class of Direct Kernel Object Manipulation (DKOM) attacks that we call Evolutionary DKOM (E-DKOM). The goal of this attack is to alter the way some data structures “evolve” over time. As case study, we designed and implemented an instance of Evolutionary DKOM attack that targets the OS scheduler for both userspace programs and kernel threads. Moreover, we discuss the

implementation of a hypervisor-based data protection system that mimics the behavior of an OS component (in our case the scheduling system) and detect any unauthorized modification.

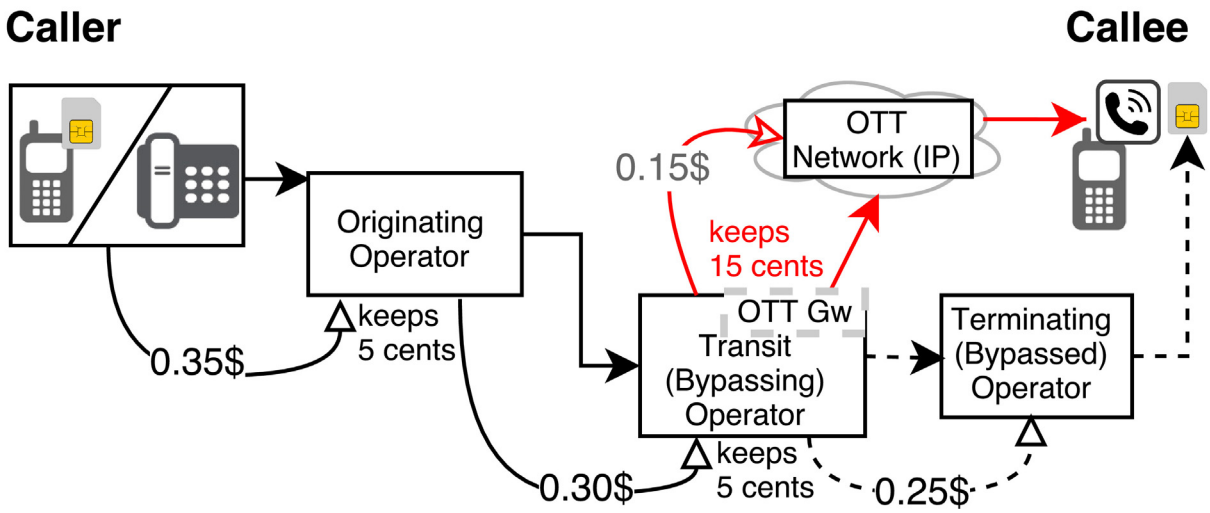
So far we described three different studies that had as main objective the protection of userspace applications and operating systems against attacks. But we also worked on improving the analysis of such attacks, in particular in the case in which they adopt code reuse attacks based on return oriented programming (ROP), which are becoming more and more prevalent every year. ROP-based attacks started as a way to circumvent operating systems protections against injected code, but they are now also used as a technique to keep the malicious code hidden from detection and analysis systems. This means that while in the past ROP chains were short and simple (and therefore did not require any dedicated tool for their analysis), we recently started to observe very complex algorithms – such as a complete rootkit – implemented entirely as a sequence of ROP gadgets. In a paper published at the ASIACCS conference [GRA16b], we presented a set of techniques to analyze complex code reuse attacks. First, we identified and discussed the main challenges that complicate the reverse engineer of code implemented using ROP. Second, we proposed an emulation-based framework to dissect, reconstruct, and simplify ROP chains. Finally, we tested our tool on the most complex example available to date: a ROP rootkit containing four separate chains, two of them dynamically generated at runtime.

Finally, to conclude our work on the Software and Operating System security area, we looked at better way to analyze malicious files protected by runtime packing – a technique employed by malware authors in order to conceal (e.g., encrypt) the malicious code and recover it at run-time. In particular, some run-time packers only decrypt individual regions of code on demand, re-encrypting them again when they are not running. This technique is known as shifting decode frames and it can greatly complicate malware analysis. The first solution that comes to mind to analyze these samples is to apply multi-path exploration to trigger the unpacking of all the code regions. Unfortunately, multi-path exploration is known to have several limitations, such as its limited scalability for the analysis of real-world binaries. To mitigate this problem, in a paper presented at the DIMVA conference [UGA16] we proposed a set of domain-specific optimizations and heuristics to guide multi-path exploration and improve its efficiency and reliability for unpacking binaries protected with shifting decode frames.

Research team
Davide Balzarotti
Marius Muench
Fabio Pagani

1.2 Spam, Phishing, and Frauds

In 2016 our group continued its activity in the analysis of online scams and telecommunication frauds. In particular, we focused our effort on three separate problems: an emerging fraud on the telephony networks, a measurement study of a number of less-known spam mitigation techniques, and the first complete study of phishers behavior.



To begin with, we studied a recent form of interconnect telecom fraud: the Over-The-Top (OTT) bypass fraud [SAH16]. In OTT bypass, a normal phone call is diverted over IP to a voice chat application on a smartphone, instead of being terminated over the normal telecom infrastructure. This rerouting (or hijack) is performed by an international transit operator in coordination with the OTT service provider, but without explicit authorization from the caller, callee and their operators. By doing so, they collect a large share of the call charge and induce a significant loss of revenue to the bypassed operators. Moreover, this practice degrades the quality of service without providing any benefits for the users. We studied the possible techniques to detect and measure this fraud and evaluate the real impact of OTT bypass on a small European country. For this, we performed more than 15,000 test calls during 8 months and conducted a user study with more than 8,000 users. In our measurements, we observed up to 83% of calls being subject to OTT bypass. Additionally, we found that OTT bypass degrades the quality of service, and sometimes collide with other fraud schemes, exacerbating the quality issues. Our user study shows that OTT bypass and its effects are also poorly understood by users.

After this new attack on telephony networks, we looked at more traditional threats – starting with spam. Spam has been largely studied in the past years from different perspectives but, unfortunately, it is still an open problem and a lucrative and active business for criminals and bot herders. While several countermeasures have been proposed and deployed in the past decade, their impact and effectiveness is not always clear. In particular, on top of the most common content- and sender-based anti-spam techniques, two minor approaches are popular among system administrators to cope with this annoying problem: greylisting and nolisting. These techniques exploit known features of the Simple Mail Transfer Protocol (SMTP) protocol that are not often respected by spam bots. This assumption makes these two countermeasures really simple to adopt and, at least in theory, quite effective. To verify if this is really the case, in 2016 we conducted the first comprehensive study of nolisting and greylisting, in which we analyzed these spam countermeasures from different perspectives. In a paper published at the DSN conference [PAG16], we measured the world-wide deployment of these solutions and their effectiveness against a real dataset of malware samples responsible to generate over 70% of the global spam traffic. We also measured the impact of these two defensive mechanisms on the delivery of normal emails. Our study provides a

unique and valuable perspective on two of the most innovative and atypical anti-spam systems. Our findings may guide system administrators and security experts to better assess their anti-spam infrastructure and shed some light on myths about greylisting and nolisting.

Finally, we moved our attention to another evergreen in the online scam ecosystem: phising. Phishing is a form of online identity theft that deceives unaware users into disclosing their confidential information. While significant effort has been devoted to the mitigation of phishing attacks, much less is known about the entire life-cycle of these attacks in the wild, which constitutes, however, a main step toward devising comprehensive antiphishing techniques. In a paper that received the best European student award at CCS [HAN16], we presented a novel approach to sandbox live phishing kits that completely protects the privacy of victims. By using this technique, we performed a comprehensive real-world assessment of phishing attacks, their mechanisms, and the behavior of the criminals, their victims, and the security community involved in the process – based on data collected over a period of five months. Our infrastructure allowed us to draw the first comprehensive picture of a phishing attack, from the time in which the attacker installs and tests the phishing pages on a compromised host, until the last interaction with real victims and with security researchers. Our study presented accurate measurements of the duration and effectiveness of this popular threat, and discussed many new and interesting aspects we observed by monitoring hundreds of phishing campaigns.

Research team
Davide Balzarotti
Aurélien Francillon
Xiao Han
Fabio Pagani
Marve Sahin

1.3 Embedded Systems and Security at the Hardware/Software Interface

This year we continued our activities in embedded systems security with work lage scale analysis of real world devices web interfaces, design of secure devices as well as on a reflection about trusted computing.

Embedded devices are becoming more widespread, interconnected, and web-enabled than ever. Recent studies showed that embedded devices are far from being secure and, however, many embedded systems rely on web interfaces for user interaction or administration. Web security is still difficult and therefore the web interfaces of embedded systems represent a considerable attack surface. In 2016, we performed the first fully automated frame work that applies dynamic firmware analysis techniques to achieve, in a scalable manner, automated vulnerability discovery within embedded firmware images[COS16]. We applied our framework to study the security of embedded web interfaces running in Commercial Off-The-Shelf (COTS) embedded devices, such as

routers, DSL/cable modems, VoIP phones, IP/CCTV cameras. We introduce a methodology and implement a scalable framework for discovery of vulnerabilities in embedded web interfaces regardless of the devices' vendor, type, or architecture. To reach this goal, we perform full system emulation to achieve the execution of firmware images in a software-only environment, i.e., without involving any physical embedded devices. Then, we automatically analyze the web interfaces within the firmware using both static and dynamic analysis tools. We also present some interesting case-studies and discuss the main challenges associated with the dynamic analysis of firmware images and their web interfaces and network services. We make observations that shed light on an important aspect of embedded devices which was not previously studied at a large scale.

Apart from analyzing existing devices we also continued our work on the design of devices with strong security guarantees. Many recent works offer practical hardware-assisted security solutions that take advantage of a tight cooperation between hardware and software to provide system-level security guarantees. Formally and consistently proving the efficiency of these solutions raises challenges since software and hardware verification approaches generally rely on different representations. In [LUG16] we explore this topic and we first sketches an ideal security verification solution naturally handling both hardware and software components. Next, we propose an evaluation of formal verification methods that have already been proposed for mixed hardware/software systems, with regards to the ideal method. At last, we presents a conceptual approach to this ideal method relying on ProVerif, and apply this approach to our dynamic root of trust system (SMART).

A lot of research efforts have been put into constructing secure systems, but experience has shown that, while there are many products which have a good level of security, others are really insecure. Some are security devices: security is at the core of their purpose; while other are not. We nevertheless often rely on the their security in our daily life and their failure can have serious consequences. In an invited paper at the DATE conference [FRA16], we discuss why we are in this situation and what we can do to improve the situation. In particular, we defend the thesis that more transparency and more openness in embedded systems hardware and software will foster a more secure ecosystem.

First, there is an economic problem. Besides being a difficult problem to solve correctly, security is most of the times expensive. Second, trust is something that is not blindly granted but that is earned by verifying it. Currently, trusted computing mechanisms often rely on unconditional trust on the systems manufacturer. However, users have too few ways to verify that the systems are trustworthy other than blindly trust the manufacturer. We should design systems where the users, i.e., the devices owners, can decide whom and what to trust. We call this Design For User Trust, where users are in control of the system. Finally, one can only trust a system fully if he can inspect it. Unfortunately, the first security measures that are implemented in embedded systems often prevent such an independent analysis (e.g., deactivation of a debug port, secure boot, encrypted file system, obfuscation). But such measures are more hiding the problems (making it difficult to discover software vulnerabilities) than solving it. They are often useful in securing a system (slowing down an attacker) but should not jeopardize our ability to analyze them. We call this Design For Security Testing. We conclude that more research is needed to make it easier to build secure systems, in particular, in the areas of concrete architectures for Design For User Trust and Design For Security Testing.

Research team
Aurélien Francillon
Florian Lugou
Andrei Costin

1.4 Web Security

This year, our research in the area of web security produced three independent works – covering different aspect of the security lifecycle: protection of web applications, detection of compromised pages, and analysis of new emerging threats.

With the advent of Web 2.0, many users started to maintain personal web pages to show information about themselves, their businesses, or to run simple e-commerce applications. This transition has been facilitated by a large number of frameworks and applications that can be easily installed and customized. Unfortunately, attackers have taken advantage of the widespread use of these technologies – for example by crafting special search engines queries to fingerprint an application framework and automatically locate possible targets. This approach, usually called Google Dorking, is at the core of many automated exploitation bots.
In a paper published at DIMVA [TOF16] we tackle this problem in three steps. We first performed a large-scale study of existing dorks, to understand their typology and the information attackers use to identify their target applications. We then proposed a defense technique to render URL-based dorks ineffective. Finally we studied the effectiveness of building dorks by using only combinations of generic words, and we proposed a simple but effective way to protect web applications against this type of fingerprinting.

Another characteristic of the recent Web is the fact that more and more applications rely on server-side requests (SSRs) to fetch resources (such as images or even entire webpages) from user-provided URLs. As for many other web-related technologies, developers were very quick to adopt SSRs, even before their consequences for security were fully understood. In fact, while SSRs are simple to add from an engineering point of view, in this paper we show that, if not properly implemented, this technology can have several subtle consequences for security, posing severe threats to service providers, their users, and the Internet community as a whole.
To shed some light on the risks of this communication pattern, we analyzed in details this phenomenon and presented the first extensive study of the security implication of SSRs in a paper published at the RAID conference [PEL16]. In our work, we propose a classification and four new attack scenarios that describe different ways in which SSRs can be abused to perform malicious activities. We also presented an automated scanner we developed to probe web applications to identify possible SSR misuses. Using our tool, we tested 68 popular web applications and find that the majority can be abused to perform malicious activities, ranging from server-side code execution to amplification DoS attacks.

Our last work in the area of web security focuses on post-mortem detection of compromised web pages using specially-designed Indicators of Compromise (IOCs). IOCs are forensic artifacts that are used as signs that a system has been compromised by an attack or that it has been infected with a particular malicious software. In a work published at the WWW conference [CAT16] we propose for the first time an automated technique to extract and validate IOCs for web applications, by analyzing the information collected by a high-interaction honeypot.

Our approach has several advantages compared with traditional techniques used to detect malicious websites. First of all, not all the compromised web pages are malicious or harmful for the user. Some may be defaced to advertise product or services, and some may be part of affiliate programs to redirect users toward (more or less legitimate) online shopping websites. In any case, it is important to detect these pages to inform their owners and to alert the users on the fact that the content of the page has been compromised and cannot be trusted.

Also in the case of more traditional drive-by-download pages, the use of IOCs allows for a prompt detection and correlation of infected pages, even before they may be blocked by more traditional URLs blacklists. Our experiments show that our system is able to automatically generate web indicators of compromise that have been used by attackers for several months (and sometimes years) in the wild without being detected. So far, these apparently harmless scripts were able to stay under the radar of the existing detection methodologies – despite being hosted for a long time on public web sites.

Research team
Davide Balzarotti
Onur Catakoglu

1.5 Papers published in 2016 in the area of System and Software Security

[SAH16] Merve Sahin, Aurélien Francillon
Over-The-Top Bypass: Study of a Recent Telephony Fraud
Proceedings of the 23rd ACM conference on Computer and communications security (CCS)

[HAN16] Xiao Han, Nizar Kheir, Davide Balzarotti
PhishEye: Live Monitoring of Sandboxed Phishing Kits
Proceedings of the 23rd ACM conference on Computer and communications security (CCS)

[MUE16] Marius Muench, Fabio Pagani, Y. Shoshitaishvili, C. Kruegel, G. Vigna, D. Balzarotti
Taming Transactions: Towards Hardware-Assisted Control Flow Integrity Using Transactional Memory
Symposium on Research in Attacks, Intrusion, and Defenses (RAID) , Paris

[PEL16] Giancarlo Pellegrino, Onur Catakoglu, Davide Balzarotti, Christian Rossow
Uses and Abuses of Server-Side Requests
Symposium on Research in Attacks, Intrusion, and Defenses (RAID) , Paris

[CRI16] Stefano Cristalli, Mattia Pagnozzi, Mariano Graziano, Andrea Lanzi, Davide Balzarotti
Micro-Virtualization Memory Tracing to Detect and Prevent Spraying Attacks
Proceedings of the 25rd USENIX Security Symposium (USENIX Security)

[TOF16] Flavio Toffalini, Maurizio Abba, Damiano Carra, Davide Balzarotti
Google Dorks: Analysis, Creation, and new Defenses
Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA) , San Sebastian

[GRA16] Mariano Graziano, Lorenzo Flore, Andrea Lanzi, Davide Balzarotti
Subverting Operating System Properties through Evolutionary DKOM Attacks
Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA) , San Sebastian

[UGA16] Xabier Ugarte-Pedrero, Davide Balzarotti, Igor Santos, Pablo G. Bringas
RAMBO: Run-time packer Analysis with Multiple Branch Observation
Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA) , San Sebastian

[PAG16] F. Pagani, M. De Astis, M. Graziano, A. Lanzi, D. Balzarotti
Measuring the Role of Greylisting and Nolisting in Fighting Spam
International Conference on Dependable Systems and Networks (DSN) , Toulouse, France

[GRA16b] Mariano Graziano, Davide Balzarotti, Alain Zidouemba
ROPMEMU: A Framework for the Analysis of Complex Code-Reuse Attacks
ACM Symposium on Information, Computer and Communications Security (ASIACCS) , Xi'an

[COS16] Andrei Costin, Apostolis Zarras, Aurélien Francillon
Automated Dynamic Firmware Analysis at Scale: A Case Study on Embedded Web Interfaces
ACM Symposium on Information, Computer and Communications Security (ASIACCS), Xi'an
[CAT16] Onur Catakoglu, Marco Balduzzi, Davide Balzarotti
Automatic Extraction of Indicators of Compromise for Web Applications
25th International World Wide Web Conference (WWW) , Montreal, Canada

[FRA16] Aurélien Francillon
Trust, But Verify: Why and how to establish trust in embedded devices (invited paper)
Proceedings of Design, Automation and Test in Europe (DATE), Dresden, Germany

[LUG16] Florian Lugou, Ludovic Apvrille, Aurelien Francillon
SMASHUP: a toolchain for unified verification of hardware/software co-designs
Journal of Cryptographic Engineering

2. SECURITY PROTOCOLS AND APPLIED CRYPTOGRAPHY

(Refik Molva, Yves Roudier)

This group's focus in 2016 has been on the following themes:

- Model-driven engineering for cyber-physical systems
- Security certification for distributed software
- Secure and privacy-preserving databases
- Verifiable cloud computing
- Verifiable storage with data reduction
- Searchable encryption

2.1 Model-driven engineering for cyber-physical systems

This research topic aims to investigate methods for producing sound security architecture designs for cyber-physical systems (CPSs). Such systems are increasingly pervasive and critical today with the advent of connected automotive systems and autonomous cars, smart cities, industrial processes, or smart grids for instance. The threats experienced by or envisioned for these systems and their complexity require a more systematic specification of security requirements and the design of a comprehensive and consistent set of security mechanisms and cryptographic protocols. They also motivate the automation of the production of adapted software that would be generated secure by design.

Our activity in this field continued to focus mainly on the design and development of model-driven engineering methodologies and tools to support the abovementioned activities. We have continued our work on SysML-Sec in collaboration with Telecom ParisTech [ALR16], and notably extended the relationship between security and safety. We are also validating the methods developed in the framework of unmanned robotic vehicles notably with respect to the specification, design, development, and verification of the security architecture of embedded subsystems and components, and of security protocols.

We are further investigating the use of Software Product Lines in order to capture the security expertise for protocol security testing, for instance for TLS, in collaboration with the MODALIS team at I3S - University of Nice (UNS) in the framework of the UCN@Sophia Labex program.

Research Team

Yves Roudier (P.I.)
Mehdi Ahizoune (Ph.D. Student)

2.2 Security certification for distributed software

In contrast with the model-driven engineering vision of automatic or assisted secure code derivation and generation, we are also investigating approaches for assessing and certifying the security of distributed software directly produced by developers, especially in the context of mobile and service-oriented software. Secure programming has emerged as a set of rules of thumb, architectural patterns, and programming practices that must be mastered by software developers so as to prevent or mitigate the introduction of vulnerabilities (e.g. buffer overflows) into application code.

We are formalizing the algorithms developed in an ad-hoc manner in so-called code analyzers or scanners that validate the correct application of those guidelines [ZRS+16]. Our approach blends information flow and control flow analysis techniques, which rely on the construction and (mostly static) analysis of data structures like

program dependence graphs (PDGs) out of the source or binary code, together with model checking techniques, that formally capture behavior verifications. Our technique aims at extracting a labeled transition system (LTS) out of the PDG over which we verify the correctness of properties using predicates expressed in the MCL logic. As such, this approach is extremely interesting for implementing a more systematic application vetting process in mobile and service software marketplaces. We are currently implementing an analyzer for Java programs based on the Joana framework to experimentally validate our proposal.

Research Team
Yves Roudier (P.I.)
Zeineb Zhioua (Ph.D. Student)

2.3 Secure and privacy-preserving databases

We are investigating private access control in databases. While this problem shares some similarities with privacy preserving data collection and aggregation, it introduces different challenges, notably related to the presence of a database schema that further leaks private information and increases the risk of data inference based on seemingly harmless requests.

We are notably addressing the question of secure and privacy-preserving databases in the framework of the eBOB project (from the French “Investissement d’Avenir” program), which aims at developing a BigData infrastructure supporting a purchase-to-pay solution through the introduction of analytics-as-a-service to predict the best timeframe for purchases. We are more specifically focusing on No-SQL databases in this context. One problem that has to be addressed is the need to enable analytics requests to run despite the database protection, which involves encryption. We have designed a solution compatible with the eBOB infrastructure and that essentially involves computing predictions over partially blinded records. We are also investigating to what extent model-based security testing might be applied in this framework.

Research Team
Yves Roudier (P.I.)
Mehdi Ahizoune (Ph.D. Student)

2.4 Verifiable cloud computing

Within the ongoing TREDISEC project, we are investigating the problem of verifiable computation which consists of assuring the correctness of the computation outsourced to the untrusted cloud server. In this regard, we introduce two cryptographic protocols that allow a lightweight client to delegate the evaluation of high-degree univariate polynomials and the multiplication of large matrices and further verify the correctness of these operations. By exploiting the mathematical properties of polynomials and matrices, the proposed solutions are more efficient than existing solutions. While the first scheme draws upon the basic properties of Euclidean division of polynomials, the second one ensures an efficient verification thanks to the use of vector projections. Besides efficiency, our protocols give way to public delegatability and public verifiability which empower any third party to query and verify the validity of the outsourced computation. Both of the solutions are provably secure under well-studied assumptions.

Research Team
Monir Azraoui
Kaoutar Elkhiyaoui

Refik Molva
Melek Önen

2.5 Verifiable storage with data reduction

One of the objectives of the EU TREDISEC project is to design new technologies to assure the integrity of the cloud storage operation. In the presence of potentially malicious cloud providers, users should receive some guarantees on the storage of data. Further to our design of a new proof of retrievability (PoR) solution named StealthGuard in 2015, we propose some new PoR primitives that are also compatible with data reduction techniques. To optimize their space savings, cloud providers call for deduplication techniques which consist of keeping a single copy of redundant data. Unfortunately, existing PoR solutions come at odds with data deduplication as the integrity values resulting from the encoding of the data before its storage are generated using a secret key that is only known by the owner of the data. Therefore the PoR encoding of a given file by two different customers results in two different outputs. With the aim of consolidating PoR with file-based deduplication, we propose the message-locked PoR approach whereby the PoR effect on duplicate data becomes identical and depends on the value of the data segment only. The solution is based on a newly proposed server-aided message-locked key generation technique which compared to existing server-aided schemes, relaxes the trust model and ensures that neither the cloud server nor the key server can guess message-locked keys.

Research Team
Kaoutar Elkhiyaoui
Refik Molva
Melek Önen
Dimitrios Vasilopoulos

2.6 Searchable encryption

As a follow up to our work in the UCN EU project, we are investigating the problem of searchable encryption in the multi-user setting. We observe that the security models originally proposed by Curtmola et al. do not reflect the problem of collusion between the cloud and a subset of users. Similarly to recent studies which presented a new type of attacks called leakage-abuse attacks against existing single-user searchable encryption techniques, we identify a new vulnerability that is inherent in the design of most of existing multi-user searchable encryption solutions and has practical consequences on the security of existing software. Most of MUSE techniques leak the position at which the search query matched the encrypted data which enables the malicious cloud provider colluding with few users to discover the content of data these users were not authorized to have access to. Besides the problem of security, we are also investigating new MUSE solutions that would decrease the communication overhead when querying some words that are not very frequent in documents. In current solutions, the querier receives as many responses as the number of documents stored at the cloud (even those who do not contain the queried keyword).

Within the ongoing CLARUS EU project, we have designed a new searchable encryption solution combining the seminal work by Curtmola et al. with Cuckoo hashes which helps in improving the performance of the search operation. This solution is fully implemented and is integrated to the CLARUS platform. The framework is compatible with SQL databases. Performance results are encouraging.

Research Team
Monir Azraoui

Kaoutar Elkhiyaoui
Refik Molva
Melek Önen
Cédric Van Rompay

2.7 Publications in 2016 in the area of Security Protocols and Applied Cryptography

[ALR16] L. Apvrille, L. Li, Y. Roudier, “Model-Driven Engineering for Designing Safe and Secure Embedded Systems”, Proceedings of the IEEE workshop on Architecture Centric Virtual Integration (ACVI’16), p. 4-7, April 2016, Venice, Italy.

[CGS+16] E. Cayirci, A. Garaga, A. Santana de Oliveira, Y. Roudier. A Risk Assessment Model for Selecting Cloud Service Providers. To appear in Journal of Cloud Computing.

[ZRS+16] Z. Zhioua, Y. Roudier, S. Short, R. Boulifa Ameur, “Security Guidelines: Requirements Engineering for Code Quality”, in Proceedings of the 3rd International Workshop on Evolving Security and Privacy Requirements Engineering (ESPRE’2016), September 12th, 2016, Beijing, China, co-located with the 24th IEEE International Requirements Engineering Conference.

[BGR15] W. Benghabrit, H. Grall, J. C. Royer, M. Sellami, M. Azraoui, K. Elkhiyaoui, M. Önen, S. Oliveira de Anderson, K. Bernsmed, From regulatory obligations to enforceable accountability policies in the Cloud, Chapter in "Cloud Computing and Services Sciences", Springer International Publishing Switzerland, 2015, ISBN: 978-3-319-25414-2

[PP16] P. Puzio, Deduplication of encrypted data in cloud computing, Ph.D. Thesis Telecom ParisTech, February 2016.
[VOEM16] D. Vasilopoulos, M. Önen, K. Elkhiyaoui, R. Molva Message-locked proofs of retrievability with secure deduplication, in CCSW 2016, 8th ACM Cloud Computing Security Workshop in Conjunction with CCS 2016, 28 October 2016, Vienna, Austria.

[AEOM16] M. Azraoui, K. Elkhiyaoui, M. Önen, R. Molva, Efficient techniques for publicly verifiable delegation of computation, in ASIACCS 2016, ASIACCS 2016, 11th ACM Asia Conference on Computer and Communications Security, May 30-June 3, 2016, Xi'ian, China.

[MA16] M. Azraoui, Verifiability and accountability in the Cloud, Ph.D. Thesis Telecom ParisTech, June 2016.

[GOK16] B. Gallego-Nicasio Crespo, M. Önen, G. Karame, TREDISEC: Towards realizing a truly secure and trustworthy cloud, in ERCIM News, Special theme: Tackling Big Data in the Life Sciences, (104),January 2016.

3. BIOMETRICS AND DIGITAL MEDIA

(Jean-Luc Dugelay, Nicholas Evans)

In 2016, research within the biometrics and digital media group focused on four main topics:

1. (Visual) Soft biometrics
2. Video surveillance
3. Iris/Voice Biometric authentication and anti-spoofing
4. Speaker diarization and acoustic context recognition

The group published 17 papers in 2016 in the most prestigious conferences and journals in the field including Patter Recognition Letters, ICPR, INTERSPEECH and ICASSP. The group also attracted three awards and was first-placed in two internationally competitive evaluations.

3.1 (Visual) Soft biometrics

Within the context of soft biometrics (mainly age and gender) we investigate both most recent techniques such as deep learning and new visual sensors such as plenoptic cameras. The group also proposed a new framework for iris recognition for smartphones.

In collaboration with Orange labs., we worked on apparent age estimation using deep learning techniques [ANT16]. Starting from a pre-trained version of the VGG-16 convolutional neural network (CNN) for face recognition, we train it on the huge IMDB-Wiki dataset for biological age estimation and then fine-tune it for apparent age estimation using the relatively small competition dataset. We show that the precise age estimation of children is the cornerstone of the competition. Therefore, we integrate a separate children VGG-16 network for apparent age estimation of children between 0 and 12 years old in our final solution. The children network is fine-tuned from the general one. We employ different age encoding strategies for training general and children networks: the soft one (label distribution encoding) for the general network and the strict one (0/1 classification encoding) for the children network. Finally, we highlight the importance of the state-of-the-art face detection and face alignment for the final apparent age estimation. Our resulting solution wins the 1st place in the competition "ChaLearn LAP and FotW Challenge and Workshop @ CVPR2016 "" significantly outperforming the runner-up.

Despite being extensively studied in the literature, the problem of gender recognition from face images remains difficult when dealing with unconstrained images in a cross-dataset protocol. We proposed in [ANT16b] a convolutional neural network ensemble model to improve the state-of-the-art accuracy of gender recognition from face images on one of the most challenging face image datasets today, LFW (Labeled Faces in the Wild). We find that convolutional neural networks need significantly less training data to obtain the state-of-the-art performance than previously proposed methods. Furthermore, our ensemble model is deliberately designed in a way that both its memory requirements and running time are minimized. This allows us to envision a potential usage of the constructed model in embedded devices or in a cloud platform for an intensive use on massive image databases.

In collaboration with UAM (Universidad Autónoma de Madrid), we worked on image-based gender estimation from body and face across distances [GON16]. The approach addresses challenging settings such as low resolution-images, as well as settings when faces are occluded. Specifically the face-based features include local binary patterns (LBP) and scale-invariant feature transform (SIFT) features, projected into a PCA space. The features of the novel body based algorithm proposed in this work include continuous shape information extracted from body silhouettes and texture information retained by HOG descriptors. Support Vector Machines (SVMs) are used for classification for body and face features. We conduct experiments on images extracted from video-sequences of the Multi-Biometric Tunnel database, emphasizing on three distance-settings: close, medium and

far, ranging from full body exposure (far setting) to head and shoulders exposure (close setting). The experiments suggest that while face-based gender estimation performs best in the close-distance-setting, body-based gender estimation performs best when a large part of the body is visible. Finally we present two score-level-fusion schemes of face and body-based features, outperforming the two individual modalities in most cases.

Research team
Jean-Luc Dugelay
Grigory Antipov
Ester Gonzalez-Sosa
Chiara Galdi

3.2 Video surveillance

In video surveillance semantic traits estimation as gender and age has always been debated topic because of the uncontrolled environment: while light or pose variations have been largely studied, defocused images are still rarely investigated. Recently the emergence of new technologies, as plenoptic cameras, yields to deal with these problems analysing multi-focus images. Thanks to a microlens array arranged between the sensor and the main lens, light field cameras are able to record not only the RGB values but also the information related to the direction of light rays: the additional data make possible rendering the image with different focal plane after the acquisition. For our experiments reported in [CHI16], we used the GUC Light Field Face Database that includes pictures from the First Generation Lytro camera. Taking advantage of light field images, we explore the influence of defocusing on gender recognition and age estimation problems. Evaluations are computed on up-to-date and competitive technologies based on deep learning algorithms. After studying the relationship between focus and gender recognition and focus and age estimation, we compare the results obtained by images defocused by Lytro software with images blurred by more standard filters in order to explore the difference between defocusing and blurring effects. In addition we investigate the impact of deblurring on defocused images with the goal to better understand the different impacts of defocusing and standard blurring on gender and age estimation.

.

With the proliferation of digital visual data in diverse domains (video surveillance, social networks, medias, etc.), privacy concerns increase. Obscuring faces in images and videos is one option to preserve privacy while keeping a certain level of quality and intelligibility of the video. Most popular filters are blackener (black masking), pixelization and blurring. Even if it appears efficient at first sight, in terms of human perception, we demonstrate in [RUC16] that as soon as the category and the strength of the filter used to obscure faces can be (automatically) identified, there exist in the literature ad-hoc powerful approaches enable to partially cancel the impact of such filters with regards to automatic face recognition. Hence, evaluation is expressed in terms of face recognition rate associated with clean, obscured and de-obscured face images.

Not only is the increasing number of cameras, but also the resolution of visual sensors and the performance of video processing algorithms. We proposed in 2016 a new visual filter [RUC16b] that includes a good trade-off between privacy and intelligibility. It ensures that people are unrecognizable while keeping the scene understandable in terms of events which allows machines to detect abnormal behaviour. The algorithm operates in the DCT domain to be compliant with the popular JPEG and MPEG codecs. For each sensitive area of the picture (i.e. area where privacy needs to be protected), the proposed algorithm uses the low-frequency coefficients of the DCT to display a privacy preserved image of the region and the high-frequency coefficients to hide most of the original information. Finally, our process allows authorized users to nearly reverse the process thanks to the hidden information.

Research team
Jean-Luc Dugelay
Natacha Ruchaud
Valeria Chiesa
Khawla Mallat

3.3 Iris/Voice Biometric authentication and anti-spoofing

In 2016, we proposed a novel approach described in [GAL16] for fast iris recognition on mobile devices. Its key features are: (i) the use of a combination of classifiers exploiting the iris colour and texture information; (ii) its limited computational time, particularly suitable for fast identity checking on mobile devices; (iii) the high parallelism of the code, making this approach also appropriate for identity verification on large database. The proposed method has been submitted to the Mobile Iris CHallenge Evaluation II. The test set employed for the contest evaluation is made available on the contest web page. The latter has been used to assess the performance of the proposed method in terms of Recognition Rate (RR) and Area Under Receiver Operating Characteristic Curve (AUC)
Through a CIFRE PhD studentship in association with NXP Software, the group has also been working to improve the reliability of automatic speaker verification technology by studying the impact of different text content on the security of spoken pass-phrases [VAL16]. The hypothesis is that some phones, and hence pass-phrases, are more speaker discriminative than others. The judicious choice of spoken pass-phrase is then a complementary approach to improve recognition performance. Through a thorough statistical analysis of relative pass-phrase strength, the work showed that weak pass-phrases are consistent across dis-joint sets of speakers. The ultimate goal of the work is to develop an automated approach of detecting and preventing the use of weak spoken pass-phrases for biometric speaker recognition.

Biometric spoofing refers to the circumvention or deception of biometric recognition technology by fraudsters who impersonate the biometric characteristics of another person in order to gain illegitimate access to sensitive or protected resources. Countermeasures aim to deflect spoofing attacks through explicit detection and thus to improve the robustness or security of biometrics technology. In 2016, the group's work has focused on anti-spoofing countermeasures for automatic speaker recognition.

Based upon the hypothesis that there is greater potential for spoofing detection using features inherently different to those used for biometric recognition, the group's work in [TOD16a] introduced an entirely new approach to detect spoofing attacks based on the constant Q transform. This technique, which results in constant Q cepstral coefficients (CQCC), provides a variable spectro-temporal resolution in the analysis of speech signals which is effective in capturing the tell-tale signs of artificial speech and manipulation. Performance was assessed on the standard ASVspoof 2015 database and was shown to outperform the previous best approach by 72% relative. The work was presented at the 2016 Odyssey Speaker and Language Recognition Workshop where the group received the Best Paper Award.

The success of CQCC features for anti-spoofing prompted us to experiment with their use for automatic speaker recognition itself [TOD16b]. Initial results were promising though inferior to what could be achieved with conventional features. When coupled with a new articulation rate (ARTE) filter, a new approach closely resembling traditional relative spectral (RASTA) processing, performance exceeds that achieved with Mel-frequency cepstral coefficients. Fusion experiments, which combine independent classifier scores produced using the two different features sets, lead to even better performance, with relative improvements of between 40 and 60% across two standard databases.

Our next step involved the combination of automatic speaker verification with spoofing countermeasures in an integrated approach [SAH16]; most previous work to develop countermeasures tackle spoofing detection independently from speaker verification while it is the effect on the latter which is of greatest interest. This work, a collaboration between partners of the EU H2020 OCTAVE project, combined two different automatic speaker verification systems and six different spoofing countermeasures in cascaded and parallel integration frameworks. Results for the cascaded approach were found to be extremely effective, especially in the case of previously unseen spoofing attacks, namely those not encountered during classifier training and optimisation. This is the real use case since, in practice, the nature of spoofing attacks can never be known a priori.

The great majority of work in anti-spoofing for automatic speaker verification was performed using large databases of unconstrained speech. However, most authentication applications where the problem of spoofing is most relevant, dictate at least some form of text constraints. This is predominantly since it is only through a text-dependent operational mode that acceptable performance can be delivered with short duration utterances. The latter is strictly necessary on the grounds of usability; reliable recognition must be delivered with utterances in the order of a few second in duration rather than a few minutes. There is then an opportunity to improve recognition performance by actively verifying the text content. This is different to speech recognition, however, since the task here is not to recognise the spoken text, but simply to verify that it matches an expected, pre-defined text. This work [KIN16], another result from the EU H2020 OCTAVE project and the first study of utterance verification, showed that false acceptances can be almost eliminated, even in the case that an impostor knows the spoken pass-phrase of another user.

The final joint paper stemming from the EU H2020 OCTAVE project [DEL16] extended EURECOM's previous work with the CQT using a more efficient approach in the infinite impulse response constant Q transform (IIR-CQT). Features resulting from the combination of IIR-CQT and cepstral analysis are referred to as infinite impulse response - constant Q, Mel-frequency cepstral coefficients (ICMC). The work assessed the utility of ICMC features for integrated automatic speaker verification and utterance verification. ICMC features were found to improve both utterance verification performance and three different approaches to text-dependent speaker verification. The fusion of the two was found to eliminate false acceptances with only a modest increase in the rate of false rejections.

Research team
Hector Delgado
Jean-Luc Dugelay
Nicholas Evans
Chiara Galdi
Massimiliano Todisco
Giacomo Valenti

3.4 Speaker diarization and acoustic context recognition

In 2016, the group's activity in digital media encompassed speaker diarization and acoustic context recognition. Speaker diarization, also referred to as the task of determining 'who spoke when', involves statistical modelling techniques which are closely related to biometric speaker recognition. Speaker diarization is in some sense a pre-cursor to speaker recognition when there is potential for more than a single active speaker, i.e. in security/surveillance scenarios. Similar modelling and pattern recognition techniques are also employed for the detection or recognition of acoustic contexts and scenes, technology which is also applicable to security/surveillance.

In 2016 some of the group's work in speaker diarization was supported in the form of a direct industrial contract with Intel Mobile Communications. The longer term goal of this work was to develop real-time or on-line speaker diarization technology. Even off-line speaker diarization is extremely challenging. On-line operation compounds the difficulty and usually results in unacceptable error rates. As a result, our work in 2016 [SOL16] investigated a semi-supervised approach to speaker diarization in which speaker models are initialised off-line with manually labelled data. The remainder of the process is then on-line and is performed using an incremental approach to maximum a posteriori speaker model adaptation developed by the group. This work produced the first (semi)-on-line approach to speaker diarization which is capable of supporting practical applications.

Additional work in speaker diarization started in 2016 through the award of a joint-national ANR project ODESSA in cooperation with Idiap and the support of the Swiss National Science Foundation. The goal of this work is to bring developments in state-of-the-art speaker recognition to improve speaker diarization performance for real-time applications involving public security and the fight against terrorism and cyber-criminality. Early work has investigated binary approaches to diarization and its suitability for on-line operation. The group submitted their work [PAT16] to the Albayzin Evaluation in speaker diarization. We achieved the single best results and received the System Award for Speaker Diarization (ex-aecquo).

The group's work in acoustic context recognition is supported through a CIFRE PhD studentship in association with NXP Software. Our earlier work investigated different means of capturing recurrent acoustic structure. This higher-level information was found to be useful in distinguishing between different acoustic scenes. Most approaches use standard algorithms and hand-crafted features. In contrast, our work in 2016 investigated automatic feature learning approaches which have better potential to capture recurrent acoustic structure. This work took the form of deep learning and convolutional neural networks. Feature extraction operates on spectral representations of raw acoustic data, hence the potential to overcome the draw-back of hand-crafted features. This, the group's first attempt to apply deep learning to acoustic context recognition was a success, with our system outperforming a standard Gaussian mixture model baseline system by 5% for 15 different acoustic classes.

With almost no exceptions, all previous work in acoustic context recognition relates to closed set evaluations, namely a fixed number of acoustic classes to which any given test segment must be assigned. This is far from the practical scenario which is, by its very nature, open set. The group's work presented in [BAT16b] was the first to address the open set issue. This work proposed a formulation for open set acoustic context recognition as a detection problem and introduced a new open set classifier with purpose-designed experimental protocols and metrics. Fundamental to the contribution was a means of assessing performance with varying degrees of openness which characterise different applications scenarios. The open set approach is based upon support vector data description (SVDD) classifier. This is used to learn a hypersphere which encapsulates the training data. The new approach is less susceptible to over fitting and is shown to outperform the more traditional approaches to acoustic context recognition developed for less realistic, closed set scenarios.

Research team

Pramod Bachhav
Daniele Battaglino
Hector Delgado
Nicholas Evans
Leela Gudupudi
Giovanni Soldi
Massimiliano Todisco

3.5 Papers published in 2016 in the area of Biometrics and Digital Media

[ANT16] Grigory Antipov, Sid-Ahmed Berrani, Jean-Luc Dugelay
Minimalistic CNN-based ensemble model for gender prediction from face images
Pattern Recognition Letters, 15 January 2016, Vol.70

[RUC16] Natacha Ruchaud, Jean-Luc Dugelay
Automatic face anonymization in visual data: Are we really well protected?
EI 2016, ISAT International Symposium on Electronic Imaging, February 14-18, 2016, San Francisco, USA

[TOD16a] Massimiliano Todisco, Hector Delgado, and Nicholas Evans
A new feature for automatic speaker verification anti-spoofing: Constant Q cepstral coefficients
ODYSSEY 2016, The Speaker and Language Recognition Workshop, June 21-24, 2016, Bilbao, Spain

[SOL16] Giovanni Soldi, Massimiliano Todisco, Hector Delgado, Christophe Beaugeant and Nicholas Evans
Semi-supervised on-line speaker diarization for meeting data with incremental maximum a-posteriori adaptation
ODYSSEY 2016, The Speaker and Language Recognition Workshop, June 21-24, 2016, Bilbao, Spain

[ANT16b] Grigory Antipov, Moez Baccouche, Sid-Ahmed Berrani, Jean-Luc Dugelay
Apparent age estimation from face images combining general and children-specialized deep learning models
CVPRW 2016, 29th IEEE Conference on Computer Vision and Pattern Recognition Workshops, June 26th-July 1st, 2016, Las Vegas, USA.

[RUC 16b] Natacha Ruchaud and Jean-Luc Dugelay
Privacy protecting, intelligibility preserving video surveillance
ICME 2016, IEEE International Conference on Multimedia and Expo, July 11-15, 2016, Seattle, USA

[CHI16] Valeria Chiesa and Jean-Luc Dugelay
Impact of multi-focused images on recognition of soft biometric traits
OPTICS + PHOTONICS 2016, SPIE Optical Engineering + Applications, 28 August-1 September 2016, San Diego, USA

[BAT16a] Daniele Battaglino, Ludovic Lepauloux and Nicholas Evans
Acoustic scene classification using convolutional neural networks
DCASE 2016, Workshop on Detection and Classification of Acoustic Scenes and Events, September 3rd, 2016, Budapest, Hungary

[TOD16b] Massimiliano Todisco, Héctor Delgado and Nicholas Evans
Articulation rate filtering of CQCC features for automatic speaker verification
INTERSPEECH 2016, Annual Conference of the International Speech Communication Association, September 8-12, 2016, San Francisco, USA

[VAL16] Giacomo Valenti, Adrien Daniel and Nicholas Evans
On the Influence of text content on pass-phrase strength for short-duration text-dependent automatic speaker authentication
INTERSPEECH 2016, Annual Conference of the International Speech Communication Association, September 8-12, 2016, San Francisco, USA

[SAH16] Md Sahidullah, Héctor Delgado, Massimiliano Todisco, Hong Yu, Tomi Kinnunen, Nicholas Evans and Zheng-Hua Tan

Integrated spoofing countermeasures and automatic speaker verification: an evaluation on ASVspoof 2015
INTERSPEECH 2016, Annual Conference of the International Speech Communication Association, September 8-12, 2016, San Francisco, USA

[KIN16] Tomi Kinnunen, Md. Sahidullah, Ivan Kukanov, Héctor Delgado, Massimiliano Todisco, Achintya Sarkar, Nicolai Thomsen, Ville Hautamaki, Nicholas Evans and Zheng-Hua Tan
Utterance verification for text-dependent speaker recognition: a comparative assessment using the RedDots corpus
INTERSPEECH 2016, Annual Conference of the International Speech Communication Association, September 8-12, 2016, San Francisco, USA

[BAT16b] Daniele Battaglino, Ludovick Lepauloux and Nicholas Evans
The open-set problem in acoustic scene classification
IWAENC 2016, 15th International Workshop on Acoustic Signal Enhancement, September 13-16, 2016, Xi'an, China

[PAT16] Jose Patino, Héctor Delgado, Nicholas Evans and Xavier Anguera
EURECOM submission to the Albayzin 2016 speaker diarization evaluation
IBERSPEECH 2016, November 23-25, 2016, Lisbon, Portugal
Best Albayzin Evaluation System Award - Speaker Diarization evaluation (ex-aecquo)

[GON16] E. Gonzalez-Sosa, A. Dantcheva, R. Vera-Rodriguez, Jean-Luc Dugelay, F. Bremond, J. Fierrez
Image-based gender estimation from body and face across distances
ICPR 2016, 23rd International Conference on Pattern Recognition, December 4-8, 2016, Cancun, Mexico

[GAL16] Chiara Galdi, Jean-Luc Dugelay
Fusing iris colour and texture information for fast iris recognition on mobile devices
ICPR 2016, 23rd International Conference on Pattern Recognition, 4-8 December 2016, Cancun, Mexico

[DEL16] Héctor Delgado, Massimiliano Todisco, Md. Sahidullah, Achintya K. Sarkar, Nicholas Evans,Tomi Kinnunen and Zheng-Hua Tan
Further optimisations of constant Q cepstral processing for integrated utterance verification and text-dependent speaker verification
SLT 2016, IEEE Workshop on Spoken Language Technology, December 13-16, 2016, San Juan, Puerto Rico

4. STAFF 2016

Professors

Davide Balzarotti

Jean-Luc Dugelay

Nicholas Evans

Refik Molva (Dept. Head)

Assistant Professors

Aurélien Francillon

Yves Roudier

Post-doctoral Fellows

Monir Azraoui

Hector Delgado

Kaoutar Elkhiyaoui

Chiara Galdi

Melek Önen

Massimilano Todisco

Engineers

Jonas Zaddach

Visiting Scientists

Ester Gonzalez-Sosa (UAM)

Alessandro Lulli (U. of Pisa)

Fabio Pagani (U. Milano)

Flavio Toffalini (U. of Verona)

Frank Hartung (U. of Aachen)

Magne Johnsen (NTNU)

Ph.D Students

Jagdish Achara

Mehdi Ahizoune

Grigory Antipov

Pramod Bachhav

Daniele Battaglino

Valeria Chiesa

Andrei Costin

Onur Çatakoglu

Son-Hai Ha

Xiao Han

Leela Gudupudi

Julien Keuffer

Florian Lugou

Marius Münch

Jose Patino

Duy Hung Phan

Pasquale Puzio

Cédric Van Rompay

Natacha Ruchaud

Merve Sahin

Giovanni Soldi

Giacomo Valenti

Dimitrios Vasilopoulos

Zeineb Zhioua

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, and addressing numerous applied problems. The study of data analysis comes with its own challenges, such as the development of methods, algorithms and ultimately computer programs for making reliable inferences from high dimensional and heterogeneous data. As a consequence, the Data Science research program is centered around the disciplines to semantically integrate and enrich data, to model and understand data, to design and analyze scalable computational approaches to machine learning, and to build systems that allow storing and processing vast amounts of Data. Ultimately, our research enables new and improved applications to emerge, in a multitude of domains.

The main research lines underpinning our academic and industrial projects, which we develop in the remainder of this report, involve the development of a solid foundation of systems and theoretical tools to interact with, manipulate and model large amounts of data:

- Large-scale data mining and fusion
- Information extraction and knowledge base population
- Machine learning, deep learning and statistical modeling
- Game theory, adversarial learning and economics models of data
- Distributed systems and algorithms

The department activities are fueled by numerous research and industrial projects, which we outline next:

- H2020 projects: IOSTACK
- FP7 projects: BigFoot, CloudSpaces
- FUI projects: eBob, NexGenTV
- EIT Digital project: 3cixty
- ANR projects: DOREMUS, GAFES, ASRAEL, T-ERC CONNECTED
- Industrial projects: BMW, ISMB, Orange Labs, Nokia Bell Labs, Data Transparency Lab grant

In the following, we report a list of the outcome of last year's remarkable and highly visible achievements:

- HDR, Patrick Loiseau
- AXA Chair Grant, Maurizio Filippone
- H2020 European Star Award, Pietro Michiardi
- Winner of the Semantic Sentiment Analysis Challenge @ ESWC 2016
- 1 Best Paper Award at ITC 2016
- 3 PhD defenses: Dr. Duy-Hung Phan, Dr. José Luis Redondo Garcia, Dr. Xiaohu Wu
- Workshops organizations: LiME 2016, SemStats 2016
- Chair role in conferences: B. Huet TPC Chair ACM MM 2016, P. Loiseau Steering Committee chair NetEcon Workshop
- Sabbatical: B. Huet @ IBM T.J. Watson, P. Loiseau @ MPI

1. DATA MINING AND INFORMATION EXTRACTION

In this Section, we detail our research work broadly dealing with information extraction from multimedia content. This includes extracting and disambiguating named entities or predicting the sentiment from any type of text, identifying persons in TV programs, analyzing images shared on social media during popular festivals, building recommendation systems and exploratory search engines, or even generating automatically movie trailer. Our research produces concrete software that compete and yield remarkable results in international benchmarks such as TRECVID, SemEval and the NEEL and OKE challenge.

1.1 Knowledge extraction in web media: at the frontier of NLP, Machine Learning and Semantics

Nowadays, the Web offers a vast amount of media content using multiple modalities: videos, pictures, sounds and texts. The main purpose of web media publication is, however, to be human readable. It is a difficult problem to extract knowledge contained in media content, and even more difficult to make the difference between what is useful, and what is not, as this is generally application and user dependent. Our research is focused on textual content in its broad diversity: newspaper articles, newswire, microposts shared on social platforms, encyclopedia articles, blog posts, audio transcription, video subtitles, etc. The other focus of this research is to properly disambiguate what we can extract from this textual content against one or multiple knowledge bases that have either generic content (e.g. DBpedia, Wikidata) or specialized content in some vertical domains (e.g. Geonames, Musicbrainz). Such a diversity calls for a knowledge extraction method that must be adaptive to these different criteria. This research is at the intersection of NLP (Natural Language Processing), machine learning and semantic technologies.

We are tackling three research challenges for performing adaptive knowledge extraction and disambiguation: i) extracting and typing entities from different kind of textual content, ii) handling different knowledge bases and studying the role of a generic indexing mechanism to guide the disambiguation process, and iii) setting automatically an entity linking process to make it self-adaptive to multiple criteria. Regarding the first challenge, we are aiming at extracting entities that must be agnostic to the kind of text and handle multiple entity types. The second challenge is focused on an agnostic process that handles multiple knowledge bases and how to optimally index the content in order to best disambiguate entities. The third challenge aims to develop an entity linking system that can be self-adapted to these different criteria.

We are developing ADEL (ADaptive Entity Linking), http://adel.EURECOM.fr/api/. We made the following scientific contributions: i) improve the NER extraction method by changing the way we use a NER system, ii) add a coreference resolution method among the existing extraction methods (POS, NER and dictionary), iii) develop a method to measure the efficiency of an index for an entity linking task over a given dataset, iv) improve the architecture of our entity linking system in order to make it self-adaptive, v) conduct a thorough study of current benchmark datasets for the entity linking task and vi) improve the linking process by re-ranking a list of candidate links according to a context.

ADEL has participated in numerous international benchmark competitions and in particular, won the OKE challenge at ESWC 2016 and was selected as the strongest baseline that no system could beat at the NEEL challenge at WWW 2016.

Research Team
Raphael Troncy (PI)
Giuseppe Rizzo (associate researcher in ISMB)
Julien Plu (PhD)

1.2 SentiME: Semantic Sentiment Analysis in Reviews and Short Messages

Sentiment analysis is defined as "*The process of computationally identifying and categorizing opinions expressed in a piece of text, especially in order to determine whether the writer's attitude towards a particular topic, product, etc. is positive, negative, or neutral".* Sentiment analysis is attracting a lot of research interest: identifying trends, extracting a sentiment for a specific hashtag on Twitter or understanding better people's opinion about a specific product sold online, restaurant or hotel that a traveler aims to book has immediate impact for determining marketing strategy or improving customer support. For example, mining and understanding the polarity of reviews is crucially important for future customers that seek opinions and sentiments to support their decision buying process.

We first performed a thorough replicate study of the top performing systems in the yearly SemEval Twitter Sentiment Analysis task, where participants must classify the message polarity of tweets among three classes (positive, neutral, negative). We discover differences among the results obtained by the top systems that have been officially published and the ones we are able to compute. Learning from the studies being made on the systems, we have developed SentiME, an ensemble system composed of five state-of-the-art sentiment classifiers. SentiME first trains the different classifiers using the Bootstrap Aggregating Algorithm using a set of linguistic (such as ngrams), and semantics (such as dictionary-based of polarity values for emojis) features. The classification results are then aggregated using a linear function that averages the classification distributions of the different classifiers.

SentiME has also been evaluated over the SemEval 2015 and 2016 test set, properly trained with the corresponding SemEval training data. We show that SentiME would outperform the best ranked system of the challenge in both years, including in 2016 where the top 3 systems were based on deep learning approaches. We also adapt SentiME to participate in the ESWC challenge where the goal is to extract the sentiment polarity of 1 million Amazon product reviews. When properly trained, SentiME reaches a F-measure of 88.05% over the test set for the detection of positive and negative polarity, which ranks our approach as the first system among the ones competing in this challenge.

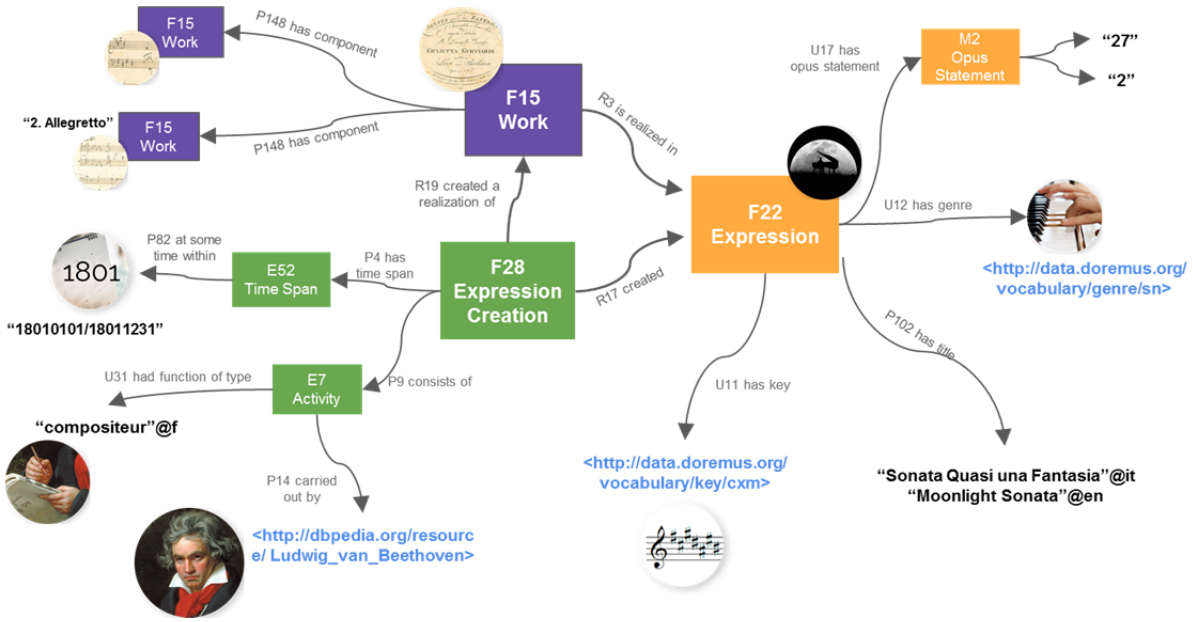
Research Team
Raphael Troncy (PI)
Efstratios Sygkounas (EURECOM student: semester project + internship)
Giuseppe Rizzo (associate researcher in ISMB)
Enrico Palumbo (PhD)

1.3 Knowledge-based Music Recommendation: Models, Algorithms and Exploratory Search

The DOREMUS project aims at bringing together the music catalogs of three major French cultural institutions: Bibliothèque Nationale de France, Radio France and Philharmonie de Paris. While they all contain music metadata, these catalogs exhibit key differences: they use different data models encoded in heterogeneous formats; they respond to specific editorial needs and adopt different viewpoints; they were built over long periods of time by numerous contributors, leading sometimes to potential inconsistency. However, they still share common objectives: collect the music related memory produced in our society as precisely as possible, and make it available to the widest range of users in a great variety of use cases. Our challenge is to unify these catalogs without losing their richness nor their particularities. Our focus is on the description of classical and traditional music works as well as their interpretations (events).

We advocate the usage of linked data technologies, and in particular RDF, as the common data model for representing and interlinking music catalogs coming from various cultural institutions. More specifically, we rely

on the FRBRoo model, and we have proposed several extensions composed of classes and properties specific to the music domain (works, recordings, scores, performances, etc.) in the DOREMUS ontology. We have developed numerous multilingual controlled vocabularies represented in SKOS. We have developed tools to convert and integrate the legacy databases from those cultural institutions (MARC2RDF).



Once the datasets have been integrated, we proposed OVERTURE, a web application enabling to explore the interlinked catalogs of those major music caalogs. OVERTURE is an exploratory search engine prototype that enables to browse through the reconciled collection of bibliographical records of classical music and to highlight the various interpretations of a work, its derivative, its performance casting as well as other rich metadata.

Beyond the exploration of the catalogs, we also aim to develop novel recommendation algorithms that fully take into the richness of the semantic graph describing musical works. We have developed AR&TREx (ARtists and TRacks Explorer), a tool that enables the user to ask for recommendations based on a seed (song) to both the Spotify and the Last.fm recommendation APIs.

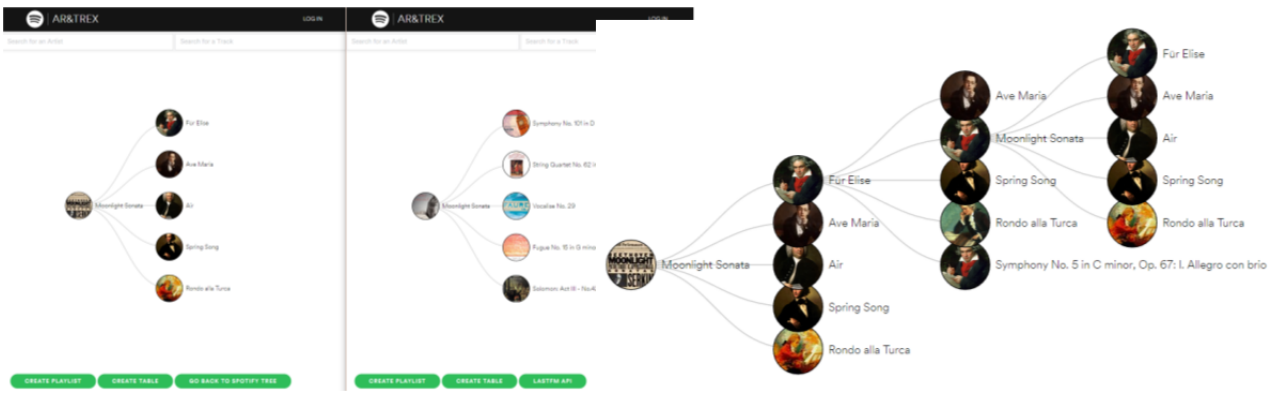


Figure 1: Moonlight Sonata recommendations from Last.Fm and Spotify

Research Team

Raphael Troncy (PI)
Eva Fernandez (EURECOM student: semester project + internship)
Pasquale Lisena (PhD)

1.4 Knowledge Base Generation for Smart Cities

Can we present the city of Nice or Milan in such a way that visitors can digitally plan in advance a visit, and explore what is happening once being in the city with the support of a digital guide?

In the context of Expo Milano 2015, we have answered this question by collecting a vast amount of cultural and tourist data about the city of Milan from the E015 digital ecosystem and social media platforms (such as Facebook, Google, Foursquare or Yelp). The resulting knowledge base, so-called the 3cixty KB, acts as a curated data marketplace which is then used to feed the ExplorMI 360 multi-device application composed of a web application and mobile companion guides available for Android and iOS devices. In 2016, we have developed several new knowledge bases for the cities of London, Nice, Singapore and the Madeira Island and we focused on the scientific challenge of improving our entity deduplication algorithm.

The 3cixty KB contains descriptions of events, places, transportation facilities and social activities, collected from numerous static, near- and real-time local and global data providers. Entities in the 3cixty KB are deduplicated, interlinked and enriched using semantic technologies.

One of the major issues encountered in the generation of knowledge bases is the integration of data coming from a collection of heterogeneous data sources. A key essential task when integrating data instances is entity matching. Entity matching is based on the definition of a similarity measure among entities and on the classification of the entity pair as a match if the similarity exceeds a certain threshold. This parameter introduces a trade-off between the precision and the recall of the algorithm, as higher values of the threshold lead to higher precision and lower recall, and lower values lead to higher recall and lower precision. Our novel contribution is a stacking approach for threshold-based classifiers. It runs several instances of classifiers corresponding to different thresholds and use their predictions as a feature vector for a supervised learner. We show that this approach is able to break the trade-off between the precision and recall of the algorithm, increasing both at the same time and enhancing the overall performance of the algorithm. We also show that this hybrid approach performs better and is less dependent on the amount of available training data with respect to a supervised learning approach that directly uses properties' similarity values. In order to test the generality of the claim, we have run experimental tests using two different threshold-based classifiers on two different data sets. Finally, we show a concrete use case describing the implementation of the proposed approach in the generation of the 3cixty knowledge base.

Research Team
Raphael Troncy (PI)
Enrico Palumbo (PhD)
Julien Plu (PhD)
Giuseppe Rizzo (associate researcher in ISMB)

1.5 Person detection and identification for TV programs enrichments

Due to the growing amount of digital multimedia content available, there is a crucial need for analysing, indexing, and search tools. A practical way to index many TV programs is to identify the relevant person that appears in them. Moreover, enriching celebrity (politicians, actors, etc.) videos by offering additional and related information via (Google, Wikipedia, youtube,etc.) is of great interested for viewers.

In the context of the NextGenTV project, we are interested in person detection and identification for TV program enrichments. The proposed system provides the identity of persons (politicians in our case) appearing in a given video shot and then provides additional information about the most “visually” relevant person. The proposed system combines approaches from computer vision for face detection and pose estimation, with those from machine learning for classification and deep features embedding.

The person identification system start with the detection of faces using the dlib library. Real-time pose estimation is then performed (with OpenCV's affine transformation) in order to align face features (eyes, nose, mouth). The face representation methodology is based on the use of the Google’s FaceNet Deep Neural Network, which achieves state-of-art face recognition performance using only 128-bytes per face. Here, the low-dimensional representation is important for efficient recognition. The next step consists in person identification, giving a name to the detected faces, using a supervised machine learning algorithm. An SVM classifier is trained on collection of public personalities gathered using the google image search api and some youtube videos. For NexGenTV, the faces to detect are French politicians.

To perform video enrichment, we propose first to localize the talking person in video shot using visual information. Few works in the state of the arts have focused only on detecting lips activity to detect the talking person. In our work, we modulate the lips motion to localize the talking face in a current TV shot program. Indeed, a talking score is computed to capture the strength of the lip activity for all faces appearing simultaneously. Work is in progress to include the audio information synchronized with the lips motion for a fully multimodal person identification approach.

In addition, we proposed a vision-based person counting approach to quantify the amount of visual appearance of each famous person during the video stream. Hence, at the end of the video, using the obtained visual counters, we can conclude about the most viewed person on the TV program. This is of particular importance for political debates where it is common to allow all candidates the same speaking time. Thanks to this work the visible time can also be provided to viewers.

Research Team
Benoit Huet (PI)
Olfa Ben-Ahmed (Post-Doc)

1.6 Video Hyperlinking (TRECVid 2016)

EURECOM has been deeply involved in the TRECVID 2016 Video Hyperlinking (LNK) tasks. We took part in the organization of the task and submitted two distinct set of runs for evaluation (one with Maria Eskevich from Radboud University, The Netherlands, the other with Elena Baralis, Paolo Garza, and Mohammad Reza Kavoosifar from Politecnico di Torino).

Video Hyperlinking task in 2016 kept the main framework of video-to-video search between anchor and target segments. The task resembles a query by example scenario where the example is a video fragment and the result is an ordered list of video fragment related to the query. For the 2016 edition of the task the dataset was changed from the professionally created and curated broadcast content provided by BBC to a collection of semi-professional user-generated videos (crawled from the blip.tv website). In addition, the anchor video segments were chosen to reflect the uploader’s intent and to be of truly multimodal nature, i.e. a combination of both audio and visual streams is crucial for the anchor understanding, processing, and target selection. EURECOM provided the visual concepts detected by a deep model (computed on the ImageNet dataset with 1000 categories) for every shot (video fragment) from the dataset to task participants. The top five concepts are provided for each key-frame along with their scores.

In the system created in collaboration with Radboud University named “EURECOM”, the focus is placed on the extension of text retrieval with visual concept detections. The Terrier Information Retrieval tool is used for initial indexing and retrieval, while the ranked list is further readjusted based on the visual features extracted for both videos and their connection to the audio content, and their contexts that are defined using word2vec terms proximity. we have also extracted the visual features using the same principle (one key-frame per video shot) with the GoogleNet deep network architecture which was shown to provide better accuracy on the ImageNet challenge. Our results show the benefit of improved visual concept detection and showed that the combination of multiple confidence scores incorporating textual and visual context is important and non-trivial to take into account. The figure shows the performance of various systems based on the MAiSP metric. This metric takes into account the user experience of interacting with multimedia content which is more time-consuming and demanding comparing to the scrolling of the textual retrieval results. It shows that our best run achieves 2sd best performance over all participants.

In the hyperlinking system created in collaboration with Politecnico Di Torino entitled “EURECOM-POLITO”, the focus is on the generation of rich queries. We investigated the importance of the temporal granularity with the submission of runs using either fixed temporal windows and runs based on the shot segmentation provided by the organizers. The system combines several features (textual and visual) in order to consider the different facets of the video collection. Specifically, we combined automatically generated transcripts (ASR), visual concepts, and the text extracted by means of an OCR tool. We also exploited WordNet to find synonyms which are used to apply a simple query expansion technique. The four submitted runs aimed at analysing the impact of the considered features on the quality of the retrieved hyperlinks. Unfortunately, the generation of long queries, using synonyms from WordNet, dramatically affected the performance of the system. The results do not provide a significant insight about which of the considered components should be included or excluded in order to achieve the best performance in terms of the metrics employed by the Hyperlinking task organizers.

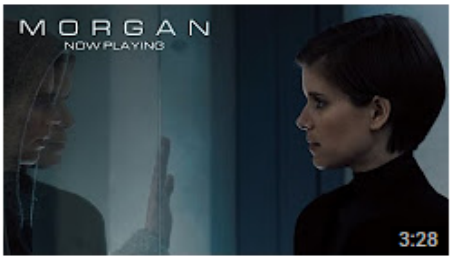
Research Team
Benoit Huet (EURECOM)
Maria Eskevich (Radboud University, The Netherlands)
Elena Baralis (Politecnico di Torino, Italy)
Paolo Garza (Politecnico di Torino, Italy)
Mohammad Reza Kavoosifar (Politecnico di Torino, Italy)

1.7 Automating Movie Trailer Making

The opportunity to work on the making of the trailer of the AI horror movie Morgan (from Luke Scott - 20th Century Fox), using cognitive computing technologies arose through an academic visiting period at IBM TJ Watson in summer 2016. The creation of movie trailer using multimedia analysis algorithm had been studied with limited extent and focused essentially on action blockbusters. Most of the work had concentrated on the extraction of scenes containing the most motion and/or the loudest audio activity. Intuitively, the trailer of specific movie genres, such as horror, drama or Sci-fi, requires tailored feature choices. Yet, it is not obvious what makes a great trailer scene for an AI horror movie and any genre.

With the recent advances in machine learning and representation learning, a training set, composed of about 100 horror movie trailers was constituted. The media collection was cut into segments and scenes using audio and visual segmentation. Each media item was then submitted to a wide range of feature extraction from both the audio and video streams. The extracted features, audio based emotions, visual concepts (people, objects and scenery), visual sentiments, location of the shot, the lighting, etc.. from the 100’s of training trailers were then studied extensively in order to identify what an AI horror trailer is usually made of. What was learned from the

analysis of 100's of trailers was then employed to select 10 key moments from the Morgan movie in order to provide a set of captivating scenes from which the trailer should be made. The shot selection was then handed over to a corporate film maker who arranged them into the final trailer which can be seen on 20th Century Fox's YouTube channel: https://www.youtube.com/watch?v=gJEzuYynaiw



Morgan | IBM Creates First Movie Trailer by AI [HD] | 20th Century FOX

20th Century Fox

4 months ago • 2,845,109 views

Scientists at IBM Research have collaborated with 20th Century Fox to create the first-ever cognitive movie trailer for the movie ...

CC

It usually requires between 2 to 3 weeks for a video editor to create the trailer of a movie. It is a labour-intensive process which requires the editor to view the entire movie and perform a selection of candidate shots. Thanks to the intelligent preselection process devised, the trailer was produced in under 24h after being injected into the experimental Watson APIs. A significant improvement which according to the editor who made the final cut, removed the burden away from the task and allowed to concentrate on the creative part!

Research Team

Winston Hsu (IBM Visiting scientist from National Taiwan University)
Dihraj (Senior Research Scientist, IBM T.J. Watson)
Benoit Huet (IBM Visiting scientist)
John R. Smith (IBM Fellow, Manager, Multimedia and Vision)

1.8 Multimedia Analysis of Festival social data

This line of work is part of the ANR project GAFES (Galerie des Festivals), which aims at extracting information from the social data published by professional and people in the context of festivals. This is an inter-disciplinary project in collaboration with sociologists who specialize in the analysis of cultural events. The project has a collaboration with several festivals, which are quite diverse in content, purpose, public and organization:

- Cannes film festival,
- Avignon theater festival,
- Lyon fête des lumières,
- "Vieilles charrues" music festival,
- TransMusicales music festival.

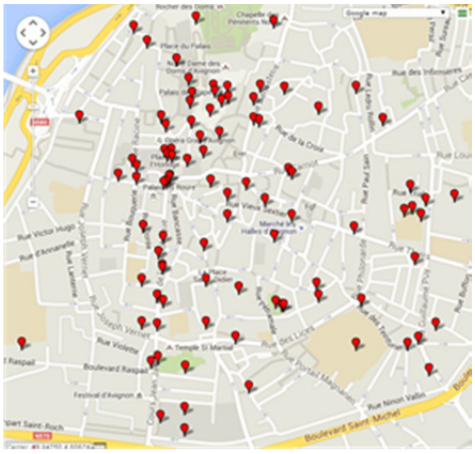
The project has established a number of tools to collect social data which is published both by the general public and by professionals in the context of those festivals, like tweets, news, blogs, photos, videos, etc… EURECOM

has specially worked on the collection of pictures and videos, using Flickr, Twitter, Instagram, Youtube and Google.

Pictures	Twitter	flickr	instagram	Total
Avignon	3188	1927	3752	8867
Cannes	5154	1585	3539	10278
Charrues	1878	1532	8204	11614
Transmusicales	3132	928	914	4974
Lumieres	3245	1584	6805	11634
				47367

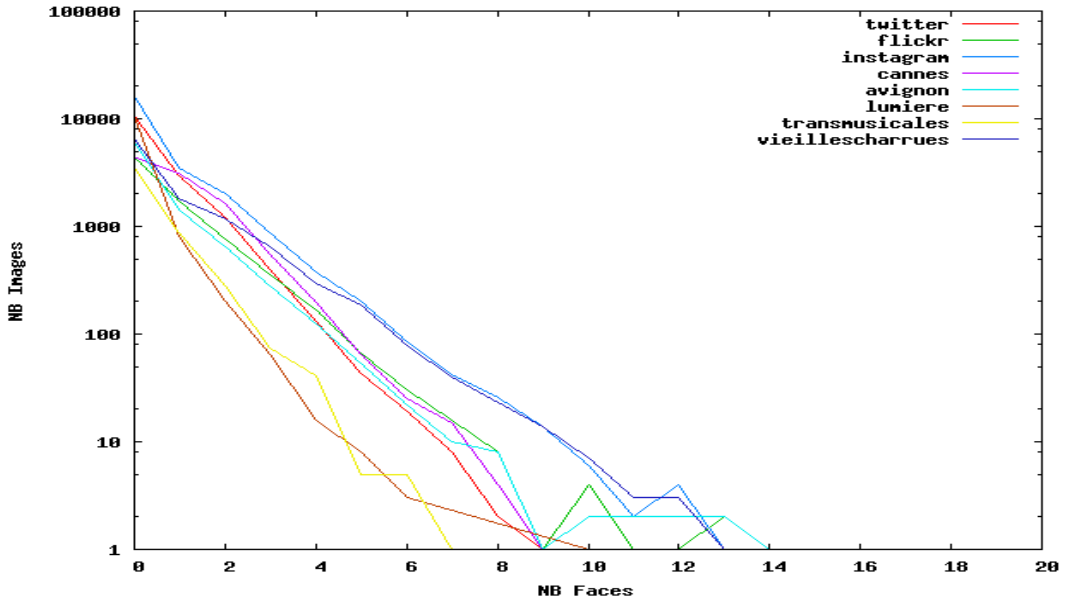
Pictures collected on the festivals

We have applied some analysis processing on these databases. On some images, we could extract the geolocalization from the metadata, and build a composed map of the picture locations.



Geolocalized pictures in Avignon

We have also performed face detection on the pictures, and could analyze the repartition of the number of faces per image in each database:

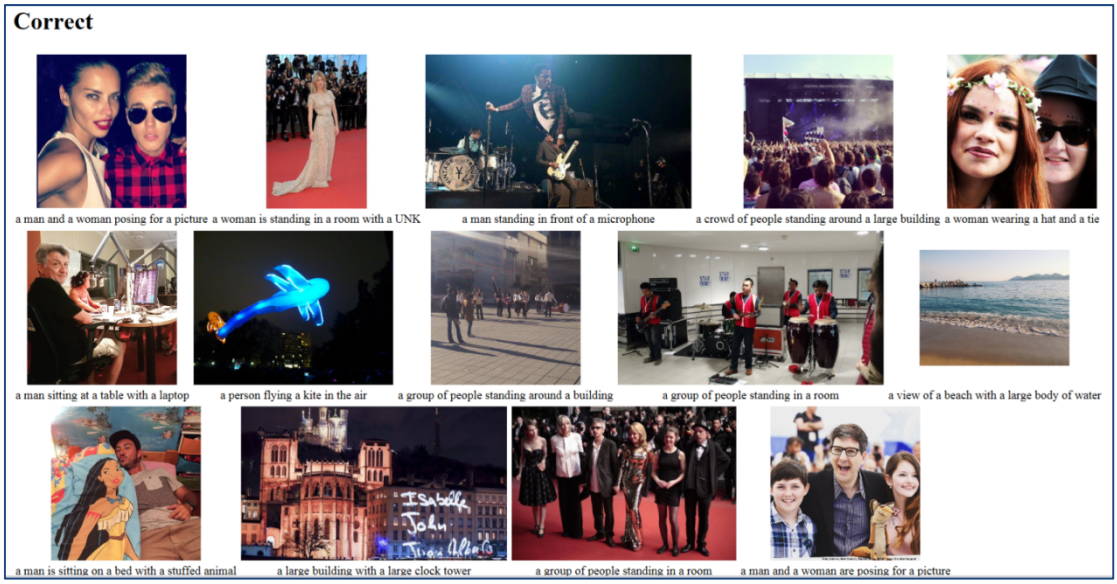


Repartition of number of faces in the database

We have also applied the NeuralTalk system developed at Stanford to generate captions for the images. We found that, despite the fact that NeuralTalk was built on a different database, the captions generated frequently contain valid information about the visual content of the picture:

Annotation correctly describes the visual content	Annotation partially describes the visual content	Annotation does not describe the visual content
25%	45%	30%

Some examples are given below:



Research Team

Bernard Merialdo (P.I.)

Benoit Huet

Antoine Ghorra (PhD student)

Danny Francis (PhD student)

2. MACHINE LEARNING

In this Section, we develop the various lines of work that deal with the design, formal analysis, and evaluation of scalable learning algorithms.

2.1 Scalable inference for probabilistic nonparametric statistical models

The study of complex phenomena through the analysis of data often requires us to make assumptions about the underlying dynamics. In modern applications of data science, we are facing the challenge of doing so when very little is known about the mechanistic description of many systems of interest; even when we do, the complexity of simulating such systems is so that we can't use it effectively. Nonparametric statistical models offer a principled way to analyze data so as to interpret the system under study as well as quantify the confidence in predictions. While nonparametric models are attractive from a theoretical point of view, they pose some serious computational challenges. Part of the research at EURECOM is devoted to solve or alleviate these problems with the ultimate goal of enabling practical and scalable nonparametric models to tackle important scientific questions in environmental and life sciences [RFGW16].

Gaussian processes based statistical models are a class of probabilistic nonparametric models. All key computations in Gaussian processes involve algebraic operations with the so-called kernel matrix, which is generally dense, with size growing with the square of the number of observations. The standard formulation of Gaussian processes involves the computation of the logarithm of the determinant of the kernel matrix and the solution of some linear systems. Alternatively, it is possible to reformulate the learning of these models by casting it as a stochastic optimization problem that can be solved in a way that requires solving linear systems only.

Here are the advancements that we made this year in the direction of making these operations faster/distributed, and in approximating models to achieve scalability:

Preconditioning

Linear systems can be solved using iterative methods such as the Conjugate Gradient algorithm. The advantage of using these algorithms for linear systems involving the kernel matrix is that the kernel matrix does not need to be stored, as these iterative solvers require iteratively carrying out matrix-vector products. In practice the conjugate gradient algorithm can converge very slowly depending on the conditioning of the system. While preconditioning is a well-known technique to accelerate convergence, its study in the context of Gaussian processes has been largely overlooked. We carried out an extensive comparison of preconditioners derived from approximations of kernel matrices and demonstrated that preconditioning helps in accelerating the learning of Gaussian processes compared to plain conjugate gradient and matrix factorization [COCF16].

Distributed Cholesky: An additional research line (partially funded by the French project eBOB) that the group recently considered consists in studying algorithmic primitives for algebraic operations on large-scale matrices. In particular, our current focus is on the design and analysis of parallel algorithms for the Cholesky factorization method, which is a decomposition of a (possibly large-scale) matrix into the product of a lower triangular matrix and its conjugate transpose, and that constitutes an essential building block for the line of work outlined in this section, e.g. for efficient numerical optimization and inference of non-parametric probabilistic statistical models.

Stochastic variational inference

The literature abounds of methods that introduce some degree of approximation in the formulation of Gaussian processes to gain scalability. We made a significant breakthrough in this area by proposing some advances in stochastic variational techniques for these models. We also leveraged the convenience of automatic differentiation when working with richly parameterized Gaussian process models and we explored various

alternative objective functions as opposed to the standard marginal likelihood. The combination of these elements allowed us to get Gaussian processes to work with datasets of the same scale of Deep Neural Networks (millions of observations) and to get unprecedented performance on some tasks where Deep Neural Networks are the preferred choice [KBCF16].

Deep probabilistic nonparametric model

A natural extension of a Gaussian process is the so-called Deep Gaussian process (DGP) where Gaussian processes are stacked on top one another in a way that the output of a Gaussian processes is the input of the next. DGPs allow for a probabilistic nonparametric modeling of composition of functions. This construction to model composition of functions naturally resembles that of Deep Neural Networks (DNNs). The connection between DGPs and DNNs has been extensively investigated in the literature, and DGPs with a variety of kernel functions can be interpreted as DNNs with an infinite number of neurons at each layer and specific activation functions. In contrast to DNNs, DGPs provide an elegant way to deal with the model-selection problem of determining a suitable number of neurons, as they are inherently nonparametric learning machines. Furthermore, they allow for a principled probabilistic framework to carry out learning of latent representations and hyper-parameters. Although attractive from a theoretical point of view, learning DGPs poses some significant computational challenges that arguably hinder their application to a wider variety of problems. In contrast, DNNs have been extremely successful in areas such as computer vision because of their amenability to GPU and distributed computations, automatic differentiation tools, and mature developments of regularization techniques, such as low-rank weight representations and dropout. We made a significant step in the direction of bridging the gap between DGPs and DNNs by showing how DGPs can be learned at scale by borrowing the key strengths of DNNs, while retaining a probabilistic formulation for accurate quantification of uncertainty. We showed how random Fourier feature approximations for DGPs with Radial Basis Function (RBF) kernels lead to trigonometric DNNs with low-rank weight matrices. We also extended these results to the case of arc-cosine kernels that can be approximated by DNNs with ReLU activation functions. We conveniently implemented stochastic variational inference in TensorFlow to infer the parameters of such approximate DGPs at scale, showing dramatic speed and accuracy improvements over previous approaches to infer DGPs [CBMF16a, CBMF16b].

Research Team

Maurizio Filippone (P.I.)
Pietro Michiardi
Kurt Cutajar (PhD student)
Duc-Trung Nguyen (PhD student)

2.2 Scalable, unsupervised clustering algorithms for arbitrary metric distances

Scalable clustering algorithms

Exploratory data analysis requires fundamental techniques to understand, describe and eventually extract value from large amounts of data. One of such techniques is data clustering. In this line of work [LDDMR16], we focus on a scalable approach for data clustering, that accommodates arbitrary similarity measures and that produces high quality clusters.

To overcome the quadratic nature of typical approaches to text clustering, our work focuses on the role of approximation in establishing a trade-off between high clustering quality and fast algorithmic runtime. We do so by designing a scalable approach for data clustering that works in two phases. In the first phase, we build an approximate k-NN graph through a parallel, iterative process reminiscent of the well-known NNDescent algorithm. The first phase concludes with a pruning stage, which strives at eliminating spurious links between items with low similarity. In the second phase, we use a parallel approach to identify density-connected components in the k-NN graph, which are a proxy for clusters of similar items.

Our focus is to understand the scalability/accuracy tradeoff that underlies our method: we do so through an extensive experimental evaluation, where we use real-life datasets including a 4 Million SPAM emails collected by Symantec Research Labs, a dump of 1.5 Million tweets collected using the Twitter API, and many other publicly available datasets. Our experimental results show that even rough approximations of k-NN graphs are sufficient to identify useful clusters, which we validate both using standard metrics and with the help of domain experts. Our method – which we implemented for the Apache Spark framework – is scalable and can be easily tuned to meet requirements stemming from different application domains.

One important consequence of our method is that it does not require accurate representations of pairwise similarity across data items to produce high quality, interpretable clusters. We support our claims, in addition to the experimental campaign, through manual inspection – by domain experts – of the quality of clusters output by our algorithm.

Online clustering algorithms

In this line of work [DPMM2016], we design an approximate algorithm that is able to modify a distributed k-NN graph by quickly adding are removing nodes. The distributed algorithm produces an approximate graph that is highly similar to the graph computed using a naive approach, although it requires the computation of far less similarities. Our algorithm starts by partitioning the k-NN graph using a balanced k-Medoids algorithm. This partitioning is used to improve nearest neighbor search based on the existing graph. Indeed, the algorithm has two main steps to add a new point to the graph: 1) use the distributed graph to search the k nearest neighbors of the new point and 2) update the graph: these neighbors are used as starting points to search existing nodes for which the new point is now a nearest neighbor. To search the nearest neighbors, inside each partition the algorithm uses a fast, sequential graph based nearest neighbor search procedure. We also perform a complete experimental evaluation of the algorithm, using both euclidean and non-euclidean datasets.

Applications of scalable clustering algorithms to Population Estimation

Statistical authorities promote and safeguard the production and publication of official statistics that serve the public good. One of their duties is to monitor the presence of individuals, region by region. Traditionally this activity has been conducted by means of censuses and surveys. Nowadays technologies open new possibilities such as a continuous sensing of the presences by leveraging the data associated to mobile devices, e.g., the behavior of users on doing calls. In this line of work [LGDDMNR2016] we propose a specifically conceived similarity function able to capture similarity between individuals call behaviors. Then we make use of a scalable clustering algorithm able to handle arbitrary metric leading to a good internal and external consistency of clusters. The approach provides better population estimation with respect to state of the art comparing with real census data. The scalability and flexibility that characterizes the proposed framework enables novel scenarios for the characterization of people by means of data derived from mobile users, ranging from the near real time estimation of presences to the definition of complex, uncommon user archetypes.

Research Team

Pietro Michiardi (P.I.)
Thibault Debatty (PhD student)
Fabio Pulvirenti (PhD student)
Matteo Dell'Amico (Symantec)
Laura Ricci (Prof. University of Pisa)
Alessandro Lulli (Visiting PhD student)

2.3 Machine Learning for Multimedia Analysis

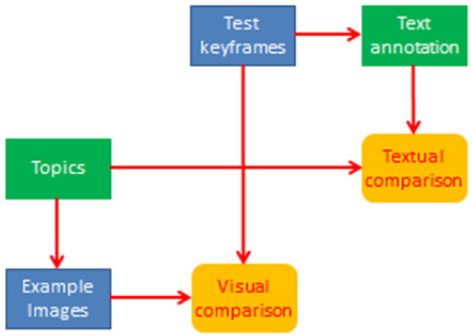
Multimedia Analysis requires complex models to extract and combine information from images, video, speech and text. We are focusing on the adaptation of Machine Learning techniques to build these models, apply them in applicative settings, and evaluate their performance. In particular, we are now studying the use of Deep Networks for creating and combining image and text representations as vectors.

EURECOM is a regular participant in the TRECVID international evaluation campaign. TRECVID is organized by NIST (US National Institute of Standards), and defines a number of tasks for which specific video data is provided to all registered participants. Participants run their systems on the same data and send the results back to NIST, where the evaluation is performed. A workshop is organized annually to summarize the results and detail the various scientific approaches that have been used.

In 2016, a new task was introduced, called "Ad-hoc Video Search" (AVS). Based on the textual description of a topic (a sentence such as "Find shots of one or more people, each walking into a building"), the goal is to retrieve all the video shots from a given database which correspond to that topic.

Each participating system has to provide a ranked list of at most 1,000 shots, which is later manually evaluated by the NIST annotators. The performance is measured as the Mean Average Precision (MAP) over a set of topics.
As this is a new problem, we have explored two orthogonal approaches:

- Starting from the text description of the topic, we use an image search engine to generate a training set of relevant images, which we use to perform a visual comparison with the search database.
- Starting from the images in the search database, we generate semantic descriptions using various types of concept detectors. We then perform a textual comparison with the content of the topics.



These approaches have shown reasonable results in the TV16 campaign. Based on more evaluation data, we plan to explore further the various parameters involved in the models, to better exploit the word and image embeddings.

Research Team
Bernard Merialdo (P.I.)
Benoit Huet
Danny Francis (PhD student)
Paul Pidou (Master student)

2.4 Design of learning algorithms for social computing systems and for networking applications

We work on the design of learning algorithms for two main applications: social computing systems and networking. Each application has its own constraints and specificities that brings challenging theoretical problems (in particular due to the large-scale nature of the system).

Matching and information disclosure in social computing systems

Individuals sharing data on today's social computing systems face privacy losses due to information disclosure that go much beyond the data they directly share. Indeed, it was shown that it is possible to infer additional information about a user from data shared by other users—this type of information disclosure is called attribute disclosure. Such studies, however, were limited to a single social computing system. In reality, users have identities across several social computing systems and reveal different aspects of their lives in each. This enlarges considerably the scope of information disclosure, but also complicates its analysis. Indeed, when considering multiple social computing systems, information disclosure can be of two types: attribute disclosure or identity disclosure—which relates to the risk of pinpointing, for a given identity in a social computing system, the identity of the same individual in another social computing system. This raises the key questions: how to measure and quantify these two privacy risks and how do they relate to each other?

In our work, we perform the first combined study of attribute and identity disclosure risks across multiple social computing systems. We first propose a framework to quantify these risks that combines our previous work on matching identities across multiple social computing systems with database notions such as k-anonymity, l-diversity and t-closeness. Then, our empirical evaluation on a real-world dataset from Facebook and Twitter then shows that, in some regime, there is a tradeoff between the two information disclosure risks, that is, users with a lower identity dis- closure risk suffer a higher attribute disclosure risk. We investigate in depth the different parameters that impact this tradeoff.

Causal methods for network performance

Communication networks are complex systems whose operation relies on a large number of components that work together to provide services to end users. As the quality of these services depends on different parameters, understanding how each of them impacts the final performance of a service is a challenging but important problem. However, intervening on individual factors to evaluate the impact of the different parameters is often impractical due to the high cost of intervention in a network. It is, therefore, desirable to adopt a formal approach to understand the role of the different parameters and to predict how a change in any of these parameters will impact performance. The approach of causality pioneered by J. Pearl provides a powerful framework to investigate these questions. Most of the existing theory is non-parametric and does not make any assumption on the nature of the system under study. However, most of the implementations of causal model inference algorithms and most of the examples of usage of a causal model to predict intervention rely on assumptions such linearity, normality, or discrete data.

In our work, we propose a methodology to overcome the challenges of working with real-world data and extend the application of causality to complex systems in the area of telecommunication networks, for which assumptions of normality, linearity and discrete data do no hold. Specifically, we study the performance of TCP, which is the prevalent protocol for reliable end-to-end transfer in the Internet. We first present and evaluate our methodology using TCP traffic obtained via network emulation, which allows us to experimentally validate the prediction of an intervention. We then apply the methodology to real-world TCP traffic sent over the Internet. Throughout the work, we compare the causal approach for studying TCP performance to other approaches such as analytical modeling or simulation and show how they can complement each other.

In a second time, we use our methodology to study the performance of DNS. For a user to access any resource on the Internet, it is necessary to first locate a server hosting the requested resource. The Domain Name System service (DNS) represents the first step in this process, translating a human readable name, the resource host

name, into an IP address. With the expansion of Content Distribution Networks (CDNs), the DNS service has seen its importance increase. In a CDN, objects are replicated on different servers to decrease the distance from the client to a server hosting the object that needs to be accessed. The DNS service should improve user experience by directing its demand to the optimal CDN server. While most of the Internet Service Providers (ISPs) offer a DNS service to their customers, it is now common to see clients using a public DNS service instead. This choice may have an impact on Web browsing performance. In this work we study the impact of choosing one DNS service instead of another and we compare the performance of a large European ISP DNS service with the one of a public DNS service, Google DNS, using our causal methodology. This allows us to explain and quantify the benefits obtained by clients using their ISP DNS service and to propose a solution to further improve their performance.

This research on learning for social computing systems is done in collaboration with Oana Goga (MPI-SWS) and Krishna Gummadi (MPI-SWS) and is partially funded by Institut Mines-Telecom through the Futur&Ruptures program. See our paper [AGLG16] for more details, the publication of which is still pending. The research on causal methods was done in collaboration with Hadrien Hours (former PhD student now at ENS Lyon), Alessandro Finamore (Telefonica) and Marco Mellia (Polito). See our publications [HBL16, HBLFM16] for details.

Research Team

Patrick Loiseau (P.I.)
Athanasios Andreou (PhD student)
Hadrien Hours (former PhD student now at ENS Lyon)
Oana Goga (MPI-SWS)
Krishna Gummadi (MPI-SWS)
Alessandro Finamore (Telefonica)
Marco Mellia (Polito)

3. GAME THEORY AND STATISTICAL LEARNING

Many modern online applications are based on statistical learning methods. On the other hand, those applications now often involve data generated by strategic self-interested agents, and game theory is the tool of choice to analyze strategic interactions between several agents. We work on theoretical developments in statistical learning, in game theory, and particularly at the frontier of both; and on their application in the area of network economics in general to build better performing systems and applications, and more specifically in the areas of social computing systems, security and privacy, cloud and smart grid.

3.1 Design of learning algorithms in adversarial environments

In many situations, in particular related to security the data from which one tries to learn is directly generated by a strategic agent seen as an adversary. That strongly impacts the performance of the learning algorithm. To design well-performing learning algorithms in such cases, we propose to model the system as a 2-players game. In 2016, we continued our work on designing robust classification algorithms in a one-shot setting; and on a dynamic setting with adversarial sequential learning with discount.

A game theoretic analysis of adversarial classification

Attack detection is usually approached as a classification problem. However, standard classification tools often perform poorly because an adaptive attacker can shape his attacks in response to the algorithm. This has led to the recent interest in developing methods for adversarial classification but there has not been a prior study that takes into account the attacker's tradeoff between adapting to the classifier being used against him with his desire to maintain the efficacy of his attack. Including this effect is key to derive solutions that perform well in practice.

In our work, we model the interaction as a game between a defender who chooses a classifier to distinguish between attacks and normal behavior based on a set of observed features and an attacker who chooses his attack features (class 1 data). Normal behavior (class 0 data) is random and exogenous. The attacker's objective balances the benefit from attacks and the cost of being detected while the defender's objective balances the benefit of a correct attack detection and the cost of false alarm. We provide an efficient algorithm to compute all Nash equilibria and a compact characterization of the possible forms of a Nash equilibrium that reveals intuitive messages on how to perform classification in the presence of an attacker. We also explore qualitatively and quantitatively the impact of the non-attacker and underlying parameters on the equilibrium strategies. Overall, our results allow designing better algorithms for adversarial scenarios.

A Set-valued Dynamic Programming Approach to Exact Regret Minimization

The regret-minimization paradigm has emerged as an effective technique for designing algorithms for online decision-making in adversarial environments. But so far, designing exact optimal policies that minimize the worst-case regret has proven to be a difficult task in general. In this work, we present a novel set-valued dynamic programming approach for characterizing such regret-optimal policies in repeated games with discounted losses and finite action sets.

Our approach first draws the connection between worst-case regret minimization and determining minimal achievable guarantees in repeated games with vector-valued losses. We then characterize the set of these minimal guarantees as the fixed point of a dynamic programming operator defined on the space of Pareto frontiers of convex and compact sets. This approach simultaneously results in the characterization of the optimal strategies that achieve these minimal guarantees, and hence of regret-optimal strategies in the original repeated game.

Based on this characterization, we propose a procedure based on approximate value iteration to compute epsilon-regret-optimal strategies for any epsilon > 0, for the case where the decision maker has only two available actions. As an illustration of this approach, we design a simple near-optimal strategy for the problem of prediction using expert advice for the case of two experts.

This work was done in collaboration with Vijay Kamble (Stanford), Jean Walrand (UC Berkeley), John Musacchio (UC Santa Cruz) and Lemonia Dritsoula (Google), and was partially funded by the France-Berkeley fund. See our papers [KLW16] and [DLM16], publications of which still pending.

Research Team
Patrick Loiseau (P.I.)
Vijay Kamble (Stanford)
Jean Walrand (UC Berkeley)
John Musacchio (UC Santa Cruz)
Lemonia Dritsoula (Google)

3.2 Economics of personal data

The main focus of our research on the economics of personal data is to understand how to quantify the value of personal information and how to incentivize users to release precise data; using game-theoretic models. As opposed to most other researches in the area, one main novelty of our work is that we take into account the externalities that are very important, that is the fact that data released by a user benefits other users. In much of our work, we assume that the data generated by users with a precision that they choose strategically is subsequently exploited through a statistical learning algorithm; so our work lies at the frontier of game theory and statistical learning. Contrarily to most other works, we do not study monetary payments but rather model learning accuracy as a public good and study how it is affected by the learning algorithm. In 2016, we continued as our work on finding ways to modify the learning algorithms so as to increasing the learning accuracy by taking into account the incentives of agents providing data, and we worked on a cooperative game-theoretic approach to quantify the value of personal data.

A cooperative game-theoretic approach to quantify the value of personal data in networks
The explosion of online services created a surge of interest on how to quantify the value of personal data. Recent studies have focused on its private cost, i.e., the amount at which users are willing to give it away. An appropriate quantification should, however, also take into account the benefit (or profit) users can get by releasing it. In our work, we argue that this makes the question much more complex because profit is not extracted directly from personal data but from the information derived from it, which has a very different economic nature: it is a public good and it has strong local externalities driven by the social network in which users are naturally embedded.

We propose a cooperative game-theoretic approach to quantify the value of personal data in networks. We model the system as a local public good game and we propose a natural extension to a cooperative game. We apply classical allocation solutions (the core and the Shapley value) to quantify the value of each user's personal data. We prove that the game and the proposed allocations have good mathematical properties. Finally, we analyze the impact of the social network structure.

This research was done in collaboration with Michela Chessa (former PostDoc now at University of Nice). See our paper [CL16] for more details, publication of which is still pending.

Research Team
Patrick Loiseau (P.I.)
Michela Chessa (former PostDoc now at University of Nice)

4. DISTRIBUTED SYSTEMS

In this section we develop the main lines of research in the domain of distributed systems, data processing frameworks, data management systems and cloud computing. In addition to fundamental research work, we also produce a number of open source software that have attracted vast interest from the industry.

4.1 Resource Allocation, scheduling and pricing in large-scale systems

Scheduling and pricing of cloud resources

Cloud computing has now become the main paradigm for computing, storage and many other applications. The design of appropriate pricing schemes for cloud resources is therefore of great importance for cloud providers to maximize their revenue. Given the capacity of a cloud, an important objective is to increase the resource utilization so as to allow more tenants to be served, increasing its revenue. In other words, the pricing problem cannot be separated from the question of tasks scheduling in the cloud. Our work addresses questions in both scheduling and pricing in cloud computing.

Due to the ubiquity of batch data processing in cloud computing, we consider a fundamental model in which a given set of batch tasks is to be scheduled on multiple identical machines and each task is specified by a value, a workload, a deadline and a parallelism bound. Within the parallelism bound, the number of machines allocated to a task can vary over time without affecting its workload. For this model, we obtain two core results: a quantitative characterization of a sufficient and necessary condition such that a set of malleable batch tasks with deadlines can be scheduled on the multiple machines, and a polynomial-time algorithm to produce such a feasible schedule. These core results provide a conceptual tool and an optimal scheduling algorithm that enable proposing new analyses and designs of algorithms and improving existing algorithms for extensive scheduling objectives. In particular, we obtain the best possible greedy algorithm, and the first dynamic programming exact algorithm for the objective of maximizing the sum of values of tasks completed before their deadline. We also provide new results for monotonic moldable tasks, for which the total workload increases when parallelizing a job; in the case of a single deadline.

On the other hand, cloud providers currently mostly use a fixed usage-based pricing, sometimes in association with a spot market to sell remaining resources and with reserved instances with long-term commitment. Given the current pricing models, tenants face difficulties to decide on which type of instance to purchase. Another contribution of our work is to extend the existing online learning algorithm of how to make decision to purchase on-demand and spot instances to the case with the instances with the long-term commitment (e.g., reserved instances or the instances an organization has already held). We also propose a framework to analyze the possibly optimal structure and scheduling policies of utilizing long-term, spot and on-demand instances. See our paper [WL16] for more details, publication of which is still pending.

Research Team

Patrick Loiseau (P.I.)
Xiaohu Wu (former PhD student now at Aalto University)

Scheduling algorithms for data-intensive scalable computing frameworks

This line of work is inscribed in the long-term quest to understand how to achieve efficiency, fairness and ultimately, high resource utilization in large-scale computing clusters.

Next, we describe a general approach to understand size-based scheduling [DCM16] and its properties, as a general policy for Big Data computing frameworks. Indeed, size-based schedulers have very desirable performance properties: optimal or near-optimal response time can be coupled with strong fairness. Despite this, however, such systems are rarely implemented in practical settings, because they require knowing a priori the amount of work needed to complete jobs: this assumption is difficult to satisfy in concrete systems. It is definitely more likely to inform the system with an estimate of the job sizes, but existing studies point to somewhat pessimistic results if size-based policies use imprecise job size estimations.
In particular, we take the goal of designing scheduling policies that explicitly deal with inexact job sizes. First, we prove that, in the absence of errors, it is always possible to improve any scheduling policy by designing a size-based one that dominates it: in the new policy, no job will complete later than in the original one. Unfortunately, size-based schedulers can perform badly with inexact job size information when job sizes are heavily skewed; we show that these issues, and the pessimistic results shown in the literature, are due to problematic behavior when large jobs are underestimated. Once the problem is identified, it is possible to amend size-based schedulers to solve the issue.

We do so by generalizing a well-known fair and efficient size-based scheduling policy; in addition, our solution deals with different job weights (that can be assigned to a job independently from its size). We provide an efficient implementation of the resulting protocol, which we call Practical Size-Based Scheduler (PSBS). Through simulations evaluated on synthetic and real workloads, we show that PSBS has near-optimal performance in a large variety of cases with inaccurate size information, that it performs fairly and that it handles job weights correctly. We believe that this work shows that PSBS is indeed practical, and we maintain that it could inspire the design of schedulers in a wide array of real-world use cases.

Additionally, in collaboration with Orange Labs Sophia Antipolis, we also focus on the problem of dynamic resource allocation. The problem comes from the realization that the de-facto standard approach to specify resource demands for scheduling applications in a cloud computing environment is to provide resource reservations: it is up to the user to specify the amount of CPU, RAM and IO resources to allocate to her Big Data processing tasks. In practice, however, our measurements and related work in the literature indicate that rarely resource reservations limits are met: this imply that resources might go wasted, as schedulers interpret resource reservations are actual resource occupation. Armed with this realization, we propose a novel mechanism (that we implement, for the sake of prototyping, for the YARN framework) to redeem untapped resources and allocate them to analytics jobs waiting in a queue to be served.

Our experiments, that were performed both in the EURECOM Cloud Computing Platform, and in the Orange Cloud production environment, indicate substantial improvements in global resource utilization, which also translate in decreased job runtimes.

Research Team
Pietro Michiardi (P.I.)
Matteo Dell'Amico (Symantec)
Damiano Carra (Prof. University of Verona)
Son-Hai Ha (PhD student)
Patrick Brown (Orange)

Container-based Analytics-as-a-Service: a cluster scheduler approach
This line of work represents a major research and development effort, toward the design, implementation and operation of a completely new cluster scheduler. Indeed, the last decade has witnessed the proliferation of numerous distributed frameworks to address a variety of large-scale data analytics and processing projects. First, MapReduce has been introduced to facilitate the processing of bulk data. Subsequently, more flexible tools, such

as Dryad, Spark, Flink and Naiad, to name a few, have been conceived to address the limitations and rigidity of the MapReduce programming model. Similarly, specialized libraries and systems like TensorFlow have seen the light to cope with large-scale machine learning problems. In addition to a fast-growing ecosystem, individual frameworks are driven by a fast-pace development model, with new releases every few months, introducing substantial performance improvements. Since each framework addresses specific needs, users are left with a wide choice of tools and combination thereof, to address the various stages of their projects.

The context depicted above has driven a lot of research in the area of resource allocation and scheduling, both from academia and the industry. These efforts materialize in cluster management systems that offer simple mechanisms for users to request the deployment of the framework they need. The general underlying idea is that of sharing cluster resources among a heterogeneous set of frameworks, as a response to static partitioning, which has been dismissed for it entails low resource utilization. Existing systems divide the resources at different levels. Some of them, e.g. Mesos and YARN, target low-level orchestration of distributed computing frameworks: to this aim, they require non-trivial modifications of such frameworks to operate correctly. Others, e.g. Kubernetes and Docker Swarm, focus on provisioning and deployment of containers, and are thus oblivious to the characteristics of the frameworks running in such containers. To the best of our knowledge, no existing tool currently addresses the problem of scheduling analytic applications as a whole, leveraging the intrinsic properties of the frameworks such applications use, but without requiring substantial modification of such frameworks.

The endeavor of this line of work is to fill the gap that exists in current approaches, and raise the level of abstraction at which scheduling works. We introduce a general and flexible definition of applications, how they are composed, and how to execute them. For example, a user application addressing the training of a statistical model involves: a user-defined program implementing a learning algorithm, a framework (e.g., Spark) to execute such a program together with information about its resource requirements, the location for input and output data and possibly hyper-parameters exposed as application arguments. Users should be able to express, in a simple way, how such an application must be packaged and executed, submit it, and expect results as soon as possible.

Our research shows that scheduling such applications represents a departure from what has been studied in the scheduling literature, and present the design of a new algorithm to address the problem. A key insight of our approach is to exploit the properties of the frameworks used by an application, and distinguish their components according to classes, core and elastic: the first being required for an application to produce work, the latter contributing to reduced execution times. Our heuristic focuses cluster resources to few applications, and uses the class of application components to pack them efficiently. Our scheduler aims at high cluster utilization and a responsive system. It can easily accommodate a variety of scheduling policies, beyond the traditional “first-come-first- served” or “processor sharing” strategies, that are currently used by most existing approaches. We study the performance of our scheduler using realistic, large-scale workload traces from Google and show it consistently outperforms the existing baseline approach which ignores component classes: application turnaround times are more than halved, and queuing times are drastically reduced. This induces fewer applications waiting to be served, and increases resource allocation up to 20% more than the baseline.

Finally, we conceived a full-fledged system, called Zoe, that schedules analytic applications according to our original algorithm and that can use sophisticated policies to determine application priorities. Our system exposes a simple and extensible configuration language that allows application definition. We validate our system with real-life experiments, and report conspicuous improvements when compared to a baseline scheduler, when using a representative workload: median turnaround times are reduced by up to 37% and median resource allocation is 20% higher.

Zoe is available as an open source project, and is currently in use in production by AirFrance / KLM, to power their infrastructure for Data Scientists teams in Orly and Amsterdam. Additionally, Zoe has been adopted as a de-

facto system fabric for the KPMG financial analytics suite. KPMG is heavily involved in the development of Zoe, with contributions including user interfaces, security and new back-ends. More information about Zoe is available through the following links:
http://zoe-analytics.eu/
https://github.com/DistributedSystemsGroup/zoe

Research Team
Pietro Michiardi (P.I.)
Francesco Pace (PhD student)
Daniele Venzano (Research Engineer)
Damiano Carra (University of Verona)
Andrea Abelli (KPMG Germnay)
Hoang-Xuan Quang-Nhat (KPMG Germany)

Optimizing demand-response mechanisms in smart grids

Demand-Response (DR) programs, whereby users of an electricity network are encouraged by economic incentives to re-arrange their consumption in order to reduce production costs, are envisioned to be a key feature of the smart grid paradigm. Several recent works proposed DR mechanisms and used analytical models to derive optimal incentives. Most of these works, however, rely on a macroscopic description of the population that does not model individual choices of users.
In our work, we conduct a detailed analysis of those models and we argue that the macroscopic descriptions hide important assumptions that can jeopardize the mechanisms' implementation (such as the ability to make personalized offers and to perfectly estimate the demand that is moved from a timeslot to another). Then, we start from a microscopic description that explicitly models each user's decision. We introduce four DR mechanisms with various assumptions on the provider's capabilities. Interestingly, we find that the optimization problems that result from our mechanisms are significantly more complex than the ones derived from macroscopic descriptions; and can be solved numerically only through a heuristic. We finally present numerical simulations that compare the different mechanisms and their sensitivity to population variability. At a high level, our results show that the performance of DR mechanisms under reasonable assumptions on the provider's capabilities are significantly lower than those suggested by previous studies, but that the gap reduces when the population's flexibility increases.

See our paper [BLN16] for more details.

Research Team
Patrick Loiseau (P.I.)
Alberto Benegiamo (former PhD student)
Giovanni Neglia (Inria)

4.2 In-memory data stores for data caching

Modern Web-scale applications attract very large numbers of end-users, which expect services to be responsive at all times: indeed, latency plays a crucial role in the perceived Quality of Experience (QoE), which determines to a large extent the popularity and success of competing services. Today's web resources have a complex structure, as they are composed by tens of objects, often served by a pool of back- end servers. In addition, objects usually do not have the same relevance: central panels, side panels or advertisements may have different values for end-users as well as content providers. To serve Web resources composed by such heterogeneous objects efficiently, modern web architectures make use of fast, in-memory storage systems that work as caches.

Traditionally, the hit rate (or probability) is a key performance metric to assess the properties of a caching scheme. The hit probability is the probability that a generic request for a given content will be served by the cache. Most of the existing literature implicitly assumes that a hit occurs if the content is stored in the cache at the moment of the request. In practice, however, in real caching systems the actual hit rate is often limited by the speed at which the cache can serve requests. In particular, Hard-Disk Drive (HDD) access times can be the key factor capping cache performance.

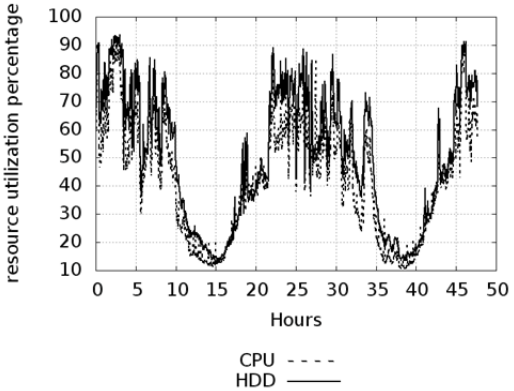


Figure 1. Graph showing the CPU and HDD utilization percentage of a production server.

As an illustrative example, Fig. 1 shows the percentage of CPU and HDD utilization, as reported by the operating system, over two days in the life of a production caching server. As the number of requests varies during the day, the resource utilization of the caching server varies as well: during peak hours, HDD utilization can exceed 95%. Such loads may cause the inability to serve a request even if the content is actually cached in the HDD. In case of a pool of cache servers, a solution based on dynamic load balancing may alleviate this problem by offloading the requests to another server. Nevertheless, this solution has its own drawbacks, because the rerouted queries are likely to generate misses at the new cache.
In this line of work [CM2016, NCFJMT2016], we study if and how the RAM can be used to alleviate the HDD load, so that the cache can serve a higher rate of requests before query-rerouting becomes necessary. Indeed, modern cache servers usually operate as a hierarchical cache, where the most recently requested contents are stored also in the RAM: upon arrival of a new request, content is first looked up in the RAM; if not found, the lookup mechanism targets the HDD. Hence, the RAM “shields” the HDD from most of the requests.

The question we ask in our work is: what is the optimal way to use the RAM? Which content should be duplicated in the RAM to minimize the load on the HDD? We show that, if content popularities are known, the problem can be formulated as a knapsack problem. More importantly, we design a new dynamic replacement policy that, without requiring popularity information to be known, can implicitly solve our minimization problem. Our policy is a variant of q-LRU. In q-LRU, after a cache miss, the content is stored in the cache with probability q and, if space is needed, the least recently used contents are evicted. We call our policy qi-LRU, because we use a different probability qi for each content i. The value qi depends on the content size and takes into account the time needed to retrieve contents from the HDD. Simulation results on real content request traces from the Akamai’s Content Delivery Network (CDN) show that our policy achieves more than 80% load reduction on the HDD with an improvement between 10% and 20% in comparison to the standard LRU.

A recent, and related line of work [NCM2017] in this domain stems from the realization that cache policies to minimize the content retrieval cost have been studied through competitive analysis when the miss costs are additive and the sequence of content requests is arbitrary. More recently, a cache utility maximization problem has been introduced, where contents have stationary popularities and utilities are strictly concave in the hit rates.

Our research bridges the two formulations, considering linear costs and content popularities. We show that minimizing the retrieval cost corresponds to solving an online knapsack problem, and we propose new dynamic policies inspired by simulated annealing, including DYNQLRU, a variant of QLRU. For such policies, we prove asymptotic convergence to the optimum under the Che's approximation. In a real scenario, popularities vary over time and their estimation is very difficult. DYNQLRU does not require popularity estimation, and our realistic, trace-driven evaluation shows that it significantly outperforms state-of-the-art policies, with up to 45% cost reduction. Our work has been accepted for publication and will appear in 2017 at the INFOCOM conference.

Finally, we notice that this work attracted the attention of AKAMAI Technologies: we are currently in contact with a team based in San Francisco to work on AKAMAI caching mechanisms and their evaluation in a production-grade environment.

Research Team
Pietro Michiardi (P.I.)
Damiano Carra (University of Verona)
Giovanni Neglia (INRIA)

4.3 Large-scale systems for RAW data cleaning and query processing

In-situ querying RAW data

In recent years, modern large-scale data analysis systems have flourished. Moreover, the batch-oriented nature of systems such as Apache Spark has been complemented by additional components (e.g., Storm and Spark streaming) that offer (near) real-time analytics on data streams. The communion of these approaches is now commonly known as the "Lambda Architecture" (LA). In fact, LA is split into three layers, i) the batch layer (based on e.g., Hadoop/Spark) for managing and pre- processing append-only raw data, ii) the speed layer (e.g., Storm/Spark streaming) tailored to analytics on recent data while achieving low latency using fast and incremental algorithms, and iii) the serving layer (e.g., Hive, SparkSQL, Impala) that exposes the batch views to support ad-hoc queries written in SQL, with low latency.

The problem with such existing large-scale analytics systems is twofold. First, combining components (layers) from different stacks, though desirable, raises performance issues and is sometimes not even possible in practice. For example, companies who have expertise in, e.g., Hadoop and traditional SQL-based (distributed) RDBMSs, would arguably like to leverage this expertise and use Hadoop as the batch processing layer and RDBMSs in the serving layer. However, this approach requires an expensive transform/load phase to, e.g., move data from Hadoop's HDFS and load it into a RDBMSs, which might be impossible to amortize, in particular in scenarios with a narrow processing window, i.e., when working on temporary data. Indeed, although many SQL-on-Hadoop systems emerged recently, they are not well designed for (short- lived) ad-hoc queries, especially when the data remains in its native, uncompressed, format such as text-based CSV files. To achieve high performance, these systems prefer to convert data into their specific column-based data format, e.g., ORC and Parquet. This works perfectly when both data and analytic queries (that is, the full workload) are in their final production stage. Namely, these self-describing, optimized data formats play an in- creasing role in modern data analytics, and this especially becomes true once data has been cleansed, queries have been well designed, and analytics algorithms have been tuned. However, when users perform data exploration tasks and algorithm tuning, that is when the data is temporary, the original data format typically remains unchanged – in this case, premature data format optimization is typically avoided, and simple text-based formats such as CSV and JSON files are preferred. In this scenario, current integrated data analytics systems can under-perform. Notably, they often fail to leverage decades old techniques for optimizing the performance of (distributed) RDBMSs, e.g., indexing, which is usually not supported.

In summary, contemporary data scientists face a wide variety of competing approaches targeting the batch and the serving layer. Nevertheless, we believe that these approaches often have overly strict focus, in many cases ignoring one another, thus failing to explore potential benefits from learning from each other.

In this line of work [TALAMV2016], we propose DiNoDB, an interactive-speed query engine that addresses the above issues. Our approach is based on a seamless integration of batch processing systems (e.g., Hadoop MapReduce and Apache Spark) with a distributed, fault-tolerant and scalable interactive query engine for in-situ analytics on temporary data. DiNoDB integrates the batch processing with the serving layer, by extending the ubiquitous Hadoop I/O API using DiNoDB I/O decorators. This mechanism is used to create, as an additional output of batch processing, a wide range of metadata, i.e., auxiliary data structures such as positional maps and vertical indexes, that DiNoDB uses to speed-up the interactive data analysis of temporary data files for data exploration and algorithm tuning. Our solution effectively brings together the batch processing and the serving layer for big data workflows, while avoiding any loading and data (re)formatting costs. While, clearly, no data analytics solution can fit all Big Data use cases, when it comes to ad- hoc interactive queries with a narrow processing window, DiNoDB outperforms state-of-the-art distributed query engines, such as Hive, Stado, SparkSQL and Impala.

Stream data cleaning

An additional line of work stems from the realization that a pervasive problem related to Big Data analytics is that of "dirty" data. Indeed, we live in a world where decisions are often based on analytics applications that process continuous streams of data. Typically, data streams are combined and summarized to obtain a succinct representation thereof: analytics applications rely on such representations to make predictions, and to create reports, dashboards and visualizations. All these applications expect the data, and their representation, to meet certain quality criteria. Data quality issues interfere with these representations and distort the data, leading to misleading analysis outcomes and potentially bad decisions.

As such, a range of data cleaning techniques were proposed recently. However, most of them focus on "batch" data cleaning, by processing static data stored in data warehouses, thus neglecting the important class of streaming data. In this paper, we address this gap and focus on stream data cleaning. The challenge in stream cleaning is that it requires both real-time guarantees as well as high accuracy, requirements that are often at odds. A naive approach to stream data cleaning could simply extend existing batch techniques, by buffering data records in a temporary data store and cleaning it periodically before feeding it into downstream components. Although likely to achieve high accuracy, such a method clearly violates real-time requirements of streaming applications. The problem is exacerbated by the volume of data cleaning systems need to process, which prohibits centralized solutions.

Therefore, our goal is to design a distributed stream data cleaning system, which achieves efficient and accurate cleaning in real-time. In our work, we focus on rule-based data cleaning, whereby a set of domain-specific rules define how data should be cleaned: in particular, we consider functional dependencies (FDs) and conditional functional dependencies (CFDs). Our system, called Bleach [TMV2016], proceeds in two phases: violation detection, to find rule violations, and violation repair, to repair data based on such violations. Bleach relies on efficient, compact and distributed data structures to maintain the necessary state (e.g., summaries of past data) to repair data, using an incremental equivalence class algorithm.

We further address the complications due to the long-term and dynamic nature of data streams: the definition of dirty data could change to follow such dynamics. Bleach supports dynamic rules, which can be added and deleted without requiring idle time. Additionally, Bleach implements a sliding window operation that trades modest additional storage requirements to temporarily store cumulative statistics, for increasing cleaning accuracy.

Our experimental performance evaluation of Bleach is twofold. First, we study the performance, in terms of throughput, latency and accuracy, of our prototype and focus on the impact of its parameters. Then we compare Bleach to an alternative baseline system, which we implement using a micro-batch streaming architecture. Our results indicate the benefits of a system like Bleach, which hold even with rule dynamics. Despite extensive work on rule-based data cleaning, we are not aware of any other stream data cleaning system

Research Team
Pietro Michiardi (P.I.)
Yongchao Tian (PhD student)
Marko Vukolic (IBM Zurich)

4.4 Multi-query optimization for large-scale, distributed query engines

Users that interact with "big data" constantly face the problem of extracting insight and obtain value from their data assets. Of course, humans cannot be expected to parse through terabytes of data. In fact, typical user interaction with big data happens through data summaries. A summary is obtained by grouping data along various dimensions (e.g., by location and/or time), and then showing aggregate functions of those data (e.g., count, sum, mean, etc.). Even graphical and interactive visualizations of data very often show aggregated results.

On-Line Analytical Processing (OLAP) tools and techniques exist to facilitate exploration of data, allowing to perform "slicing and dicing" by grouping data along multiple dimensions. In relational databases, this is facilitated by extending the traditional SQL GROUP BY clause with constructs such as ROLLUP, CUBE, and GROUPING SETS. These operations have seen extensive efforts to optimize their implementation in relational databases, which only apply to single servers or small clusters.

Despite the importance of data summarization, the field of data-intensive scalable computing (DISC) systems – where data can reach petabytes and be distributed on clusters of thousands of machines – has not seen much effort toward efficient implementations of the above concepts. In the Hadoop MapReduce ecosystem, high-level query languages such as Pig Latin and HiveQL offer simple implementations of the above constructs, which do not perform aggressive optimizations. In enterprise workloads, jobs coming from queries written in high-level languages are the majority; an optimized implementation of these operators is therefore truly desirable.

In this line of work [PM2016], we focus on optimizing a predominant operation in databases: data summarization. Users that interact with data (especially "big data") constantly feel the needs of computing aggregates to extract insights and obtain value from their data assets. Of course, humans cannot be expected to parse through Gigabytes or Terabytes of data. In fact, typically, users interact with data through data summaries. A summary is obtained by grouping data on various combinations of dimensions (e.g., by location and/or time), and by computing aggregates of those data (e.g., count, sum, mean, etc.) on such combinations. These summaries are then used as input data for all kinds of purposes such as joining with other data, visualization on dashboards, business intelligence decisions, data analysis, anomaly detection, etc. From this perspective, we consider data summarization as a crucial task that is performed extremely frequently. The workload and query templates of industrial benchmarks for databases justify this point. For instance, 20 of 22 queries in TPC-H and 80 out of 99 queries in TPC-DS are data summarization queries. In addition, a significant portion of queries in TPC-H and TPCDS access the same data, which means some data is "hotter" than others. A cross-industry study in actually confirms this observation. All together, these bestow a great chance for multiple Group By query optimization to achieve better performance.

The Group By operator is the building block of data summarization, as all other operators are its generalizations. For example, a Cube operator (introduced by Gray et al.) computes Group Bys corresponding to all possible

combinations of a list of attributes. Rollup and Cube operators allow users to compactly describe a large number of combinations of Group Bys. Numerous solutions for generating the whole space of data Cube and Rollup have been proposed. However, in the era of "big data", datasets with hundreds or thousands of attributes are very common (e.g. data in biomedical, physics, astronomy, etc.). Due to the large number of attributes, generating the whole space of data Cube and Rollup is inefficient. Also, very often users are not interested in the set of all possible Group Bys, but only a certain subset. The Grouping Sets operator facilitates this preference by allowing users to specify the exact set of desired Group Bys. In short, Cube, Rollup and Grouping Sets are convenient ways to declare multiple Group By queries.

In this line of work, we tackle the general problem of optimizing data summarization: how to efficiently compute a set of multiple Group By queries. This problem is known to be NP-complete, and all state of the art algorithms use heuristic approaches to approximate the optimal solution. However, none of prior works scales well with large number of attributes, and/or large number of queries. Therefore, we present a novel algorithm that: 1) Scales well with both large numbers of attributes and numbers of Group By queries. In our experiment, the latency introduced by our query optimization algorithm is several orders of magnitude smaller than that of prior works. As the optimization latency is an overhead that we should minimize, our approach is truly desirable; 2) Empirically performs better than state of the art algorithms: in many cases our algorithm finds a better execution plan, in many other cases it finds a comparable execution plan, and in only a few cases it slightly trails behind.

Research Team
Pietro Michiardi (P.I.)
Duy-Hung Phan (PhD student)

4.5 Software-defined Storage for Big Data Analytics

In this line of work, we study the I/O performance and the fairness properties that private cloud deployments expose to a very specific and demanding class of applications, namely "Big Data" applications. In particular, we have developed a software framework to proceed with the instrumentation, execution of a measurement campaign, and collection of raw log files to gain a better understanding of the overheads associated to virtualization, when compared to bare-metal performance. This work proceeds in two concurrent threads, which we outline below:

- Simple storage scenario: in this case, we consider a single host, and focus on low-level measurements to understand the intricate dependence between performance and system configuration. In particular, our active measurements emulate the workload imposed by a simple distributed file system, from the viewpoint of a single host. Our findings question current beliefs and the current best practices to configure Big Data applications running in virtualized environments. Indeed, our main results show that virtualization overheads are often negligible from the I/O perspective, but it is important to configure appropriately both operating systems and applications.

- Software-defined storage scenario: in this case, we consider the flexibility offered by the software-defined storage paradigm, and perform application-level performance measurements to understand the implications of a variety of storage layers, including: traditional distributed file systems, object storage and volume-based storage systems, such as the one commonly offered by public cloud vendors. Our findings indicate that data locality seems not to play a major role (contrary to the common belief) in determining the performance of individual Big Data applications. Instead, when a multi-tenant scenario is considered, concurrency hinders the task of allocating I/O resources, calling for alternative storage architectures, and informed placement of computing tasks, to achieve data locality.

Our work, which materializes in two main publications [IOSTACK2016, PMVCM2016], constitutes a natural conjunction with the research lines on scheduling.

Research Team
Pietro Michiardi (P.I.)
Francesco Pace (PhD student)
Daniele Venzano (Research Engineer)

5. RESEARCH OUTPUT

5.1 International Journals

[VV16] Viotti, Paolo; Vukolic, Marko, “Consistency in non-transactional distributed storage systems”, ACM Computing Surveys, Vol.49, N°1, June 2016.

[DCM16] Dell'Amico, Matteo; Carra, Damiano; Michiardi, Pietro, “PSBS: Practical size-based scheduling”, in IEEE Transactions on Computers, Volume 65, N°7, July 2016.

[CM2016] Carra, Damiano; Michiardi, Pietro, “Memory partitioning and management in Memcached”, in IEEE Transactions on Services Computing, September 2016, Vol PP, N°99.

[TALAMV2016] Tian, Yongchao; Alagiannis, Ioannis; Liarou, Erietta; Ailamaki, Anastasia; Michiardi, Pietro; Vukolic, Marko, “DiNoDB: an interactive-speed query engine for ad-hoc queries on temporary data”, in IEEE transactions on Big Data, December 2016, Vol.PP, N°99.

[IOSTACK2016] R. Gracia-Tinedo ; et. al., “IOStack: Software-defined object storage”, in IEEE Internet Computing, 18 Mars 2016, Vol. PP, N°99, ISSN: 1089-7801.

[RFGW16] J. M. Rondina, M. Filippone, M. Girolami, and N. S. Ward. Decoding post-stroke motor function from structural brain imaging. NeuroImage: Clinical, 12:372-380, 2016.

[HBLFM16] Hadrien Hours, Ernst Biersack, Patrick Loiseau, Alessandro Finamore, and Marco Mellia. A Study of the Impact of DNS Resolvers on CDN Performance Using a Causal Approach. Computer Networks, Special issue on “Traffic and Performance in the Big Data Era”, 2016.

[HBL16] Hadrien Hours, Ernst Biersack, and Patrick Loiseau. A causal approach to the study of TCP performance. ACM Transactions on Intelligent Systems and Technology, Special Issue on “Causal Discovery and Inference” (K. Zhang, J. Li, E. Bareinboim, B. Schölkopf, and J. Pearl, editors), 7(2):25:1–25:25, January 2016.

[KKL16] Raimo Kantola, Hammad Kabir, and Patrick Loiseau. Cooperation and End-to-End in the Internet. International Journal of Communication Systems, 2016.

[LZHT16] Li, Haojie; Zha, Zheng-Jun; Huet, Benoit; Tian, Qi, “Guest Editorial: Large-scale multimedia content analysis on social media”, Multimedia Tools and Applications, February 2016, Vol.75, N°3.

[QLHL15] Qi, Guo-Jun.; Larochelle, Hugo; Huet, Benoit; Luo, Jiebo; Yu, Kai, “Guest editorial: Deep learning for multimedia computing”, IEEE Transactions on Multimedia, Vol 17, N°11, November 2015.

[LM16] Li, Yingbo; Merialdo, Bernard, “Multimedia maximal marginal relevance for multi-video summarization”, Multimedia Tools and Applications, January 2016, Vol. 75, N°1, ISSN: 1380-7501.

[RMPFT16] Rizzo, Giuseppe; Meo, Rosa; Pensa, Ruggero G.; Falcone, Giacomo; Troncy, Raphaël, “Shaping city neighborhoods leveraging crowd sensors”, Information Systems, 5 July 2016.

[AST16] Assaf, Ahmad; Senart, Aline; Troncy, Raphaël, “Towards an objective assessment framework for linked data quality: Enriching dataset profiles with quality indicators”, International Journal on Semantic Web and

Information Systems (IJSWIS), Special Issue on Dataset Profiling and Federated Search for Linked Data, Vol. 12, N°3, 2016, ISSN: 1552-6283.

[KT16] Khrouf, Houda; Troncy, Raphaël, “EventMedia: a LOD dataset of events illustrated with media”, Semantic Web Journal (SWJ), Special Issue on Linked Dataset descriptions, 12 February 2016, Vol. 7, N°2, IOS Press, ISSN: 1570-0844.

5.2 International Conferences, Workshops, Preprints

[VMV16] Viotti, Paolo; Meiklejohn, Christopher; Vukolic, Marko, “Towards property-based consistency verification”, PAPOC 2016, Workshop on Principles and Practice of Consistency for Distributed Data, April 18th, 2016, London, UK / co-hosted with EUROSYS 2016.

[LVCQV16] Liu, Shengyun; Viotti, Paolo; Cachin, Christian; Quéma, Vivien; Vukolic, Marko, “XFT: Practical fault tolerance beyond crashes”, OSDI 2016, 12th USENIX Symposium on Operating Systems Design and Implementation, November 1-4, 2016, Savannah, GA, USA.

[LDDMR16] Lulli, Alessandro; Dell’Amico, Matteo; Michiardi, Pietro; Ricci, Laura, “NG-DBSCAN: Scalable density-based clustering for arbitrary data”, Proceedings of the VLDB Endowment, 2016, Vol.10, N°3, to appear in VLDB 2017.

[DPMM16] Debatty, Thibault; Pulvirenti, Fabio; Michiardi, Pietro; Mees, Wim, “Fast distributed k-nn graph update”, in Proc. of the IEEE BigData Workshop on Big Graphs, December 5th, 2016, Washington, USA.

[LGDDMNR2016] Lulli, Alessandro; Gabrielli, Lorenzo; Dazzi, Patrizio; Dell'Amico, Matteo; Michiardi, Pietro; Nanni, Mirco; Ricci, Laura, “Improving population estimation from mobile calls: A clustering approach”, in Proc. of IEEE Symposium on Computers and Communication, ISCC 2016, Messina, Italy, June 27-30, 2016.

[NCFJMT2016] Neglia, Giovanni; Carra, Damiano; Feng, Mingdong; Janardhan, Vaishnav; Michiardi, Pietro; Tsigkari, Dimitra, “Access-time aware cache algorithms”, in Proc. of ITC 2016, 28th International Teletraffic Congress, September 12-16, 2016, Würzburg, Germany.

[NCM2017] Neglia, Giovanni; Carra, Damiano; Michiardi, Pietro, “Cache Policies for Linear Utility Maximization”, to appear in IEEE INFOCOM 2017.

[TMV2016] Tian, Yongchao; Michiardi, Pietro; Vukolic, Marko, “Bleach: A Distributed Stream Data Cleaning System”. CoRR abs/1609.05113.

[PM2016] Phan, Duy-Hung; Michiardi, Pietro, “A novel, low-latency algorithm for multiple group-by query optimization”, in proc. of IEEE ICDE 2016, 32nd IEEE International Conference on Data Engineering, May 16-20, 2016, Helsinki, Finland.

[PMVCM2016] Pace, Francesco; Milanesio, Marco; Venzano, Daniele; Carra, Damiano; Michiardi, Pietro, “Experimental performance evaluation of cloud-based analytics-as-a-service”, in Proc. of IEEE CLOUD 2016, 9th IEEE International Conference on Cloud Computing, June 27-July 2, 2016, San francisco, USA.

[CBMF16a] K. Cutajar, E. V. Bonilla, P. Michiardi, and M. Filippone. Accelerating deep Gaussian processes inference with arc-cosine kernels. In Bayesian Deep Learning Workshop, NIPS, 2016.

[CBMF16b] K. Cutajar, E. V. Bonilla, P. Michiardi, and M. Filippone. Practical learning of deep Gaussian processes via random Fourier features, 2016. arXiv:1610.04386.

[COCF16] K. Cutajar, M. A. Osborne, J. P. Cunningham, and M. Filippone. Preconditioning kernel matrices. In Proceedings of the 32nd International Conference on Machine Learning, ICML 2016, New York City, USA, June 19-24, 2016, 2016.

[KBCF16] K. Krauth, E. V. Bonilla, K. Cutajar, and M. Filippone. AutoGP: Exploring the capabilities and limitations of Gaussian process models, 2016. arXiv:1610.05392.

[AGLG16] Athanasios Andreou, Oana Goga, Patrick Loiseau, and Krishna P. Gummadi. Identity vs. attribute disclosure risks for users with multiple social profiles. Preprint, 2016.

[KLW16] Vijay Kamble, Patrick Loiseau, and Jean Walrand. Regret-optimal strategies for playing repeated games with discounted losses. Preprint arXiv:1603.04981, 2016.

[DLM16] Lemonia Dritsoula, Patrick Loiseau, and John Musacchio. A game-theoretic analysis of adversarial classification. Preprint arXiv:1610.04972, 2016.

[CL16] Michela Chessa and Patrick Loiseau. A cooperative game-theoretic approach to quantify the value of personal data in networks. Preprint, 2016.

[WL16] Xiaohu Wu and Patrick Loiseau. Algorithms for scheduling malleable cloud tasks. Preprint arXiv:1501.04343, 2016.

[BLN16] Alberto Benegiamo, Patrick Loiseau, and Giovanni Neglia. Dissecting demand response mechanisms: the role of consumption forecasts and personalized offers. In Proceedings of the American Control Conference (ACC), July 2016.

[MPEH16] Merialdo, Bernard; Pidou, Paul; Eskevich, Maria; Huet, Benoit, “EURECOM at TRECVID 2016: The adhoc video search and video hyperlinking tasks”, TRECVID 2016, 20th International Workshop on Video Retrieval Evaluation, 14-16 Novembre 2016, Gaithersburg, Ma, USA.

[AFJMSKWOJHL16] Awad, George; Fiscus, Jonathan; Joy, David; Michel, Martial; Smeaton, Alan F; Kraaij, Wessel; Eskevich, Maria; Aly, Robin; Ordelman, Roeland; Jones, Gareth J. F; Huet, Benoit; Larson, Martha, “TRECVID 2016: Evaluating vdeo search, video event detection, localization, and hyperlinking”, TRECVID 2016, 20th International Workshop on Video Retrieval Evaluation, 14-16 Novembre 2016, Gaithersburg, Ma, USA.

[EJAOH16] Eskevich, Maria; Jones, Gareth J. F.; Aly, Robin; Ordelman, Roeland; Huet, Benoit, “Pursuing a moving target: Iterative use of benchmarking of a task to understand the task”, MEDIAEVAL 2016, MediaEval Benchmarking Initiative for Multimedia Evaluation Workshop, October 20-21, 2016, Hilversum, Netherlands.

[YPNH16] Yang, Shujun; Pang, Lei; Ngo, Cho-Wah; Huet, Benoit, “Serendipity-driven celebrity video hyperlinking”, ICMR 2016, ACM International Conference on Multimedia Retrieval, June 6-9, 2016, New York, USA.

[LWHW16] Liu, Xueliang; Wang, Feifei; Huet, Benoit; Wang, Feng, “E²SGM: Event enrichment and summarization by graph model”, MMM 2016, 22nd International Conference on Multimedia Modelling, January 4-6, 2016, Miami, USA / Also published in LNCS, Vol. 9517/2016.

[TSQHHL16] Tian, Qi; Sebe, Nicu; Qi, Guo-Jun; Huet, Benoit; Hong, Richang; Liu, Xiueliang, “Multimedia Modeling”, MMM 2016, 22nd International Conference on Multimedia Modeling, January 4-6, 2016, Miami, FL, USA / Also published as LNCS, Vol 9516+9517/2016.

[LGM16] Lorieul, Titouan; Ghorra, Antoine; Merialdo, Bernard, “Static and dynamic autopsy of deep networks”, CBMI 2016, 14th International Workshop on Content-based Multimedia Indexing, June 15-17, 2016, Bucharest, Romania.

[SRT16] Sygkounas, Efstratios; Rizzo, Giuseppe; Troncy, Raphaël, “A replication study of the top performing systems in SemEval Twitter sentiment analysis”, ISWC 2016, 15th International Semantic Web Conference, Resources Track, October 17-21, 2016, Kobe, Japan.

[DTTCVCL16] Destandau, Marie; Troncy, Raphaël; Todorov, Konstantin; Cecconi, Cécile; Voisin, Martine; Canno, Isabelle; Leresche, Françoise, “Using linked data for structuring and interlinking music catalogs: How three major french cultural institutions finally came to an agreement”, IFLA 2016, Workshop on Data In Libraries: The Big Picture, 10 August 2016, Chicago, USA.

[PRT16a] Palumbo, Enrico; Rizzo, Giuseppe; Troncy, Raphaël, “An ensemble approach to financial entity matching for the FEIII 2016 challenge”, DSMM 2016, ACM SIGMOD/PODS International Workshop on Data Science for Macro-Modeling with Financial and Economic Datasets, July 1, 2016, San Francisco, USA.

[PRT16b] Plu, Julien; Rizzo, Giuseppe; Troncy, Raphaël, “All models for one purpose and one purpose for all models: Enhancing entity linking by combining models”, ESWC 2016, 13th European Semantic Web Conference (ESWC'16), Open Extraction Challenge, May 29-June 2, 2016, Heraklion, Greece.

[MRTCG16] Mihindukulasooriya, Nandana; Rizzo, Giuseppe; Troncy, Raphaël; Corcho, Oscar; Garcia-Castro, Raul, “A two-fold quality assurance approach for dynamic knowledge bases: The 3cixty use case”, ESWC 2016, International Workshop on Completing and Debugging the Semantic Web (CoDeS'16), May 30, 2016, Heraklion, Greece.

[SRT16] Sygkounas, Efstratios; Rizzo, Giuseppe; Troncy, Raphaël, “Sentiment polarity detection from Amazon reviews: An experimental study”, ESWC 2016, 13th European Semantic Web Conference, Semantic Sentiment Analysis Challenge, First Ranked System, May 29-June 2, 2016, Heraklion, Greece.

[IREPT16] Ilievski, Filip; Rizzo, Giuseppe; van Erp, Marieke; Plu, Julien; Troncy, Raphaël, “Context-enhanced adaptive entity linking”, LREC 2016, 10th edition of the Language Resources and Evaluation Conference, 23-28 May 2016, Portoroz, Slovenia.

[REPT16] Rizzo, Giuseppe; van Erp, Marieke; Plu, Julien; Troncy, Raphaël, “NEEL 2016: Named Entity rEcognition & Linking challenge report”, WWW 2016, 6th International Workshop on Making Sense of Microposts (Microposts'16), NEEL Challenge, April 12, 2016, Montreal, Canada.

[LT16] Lisena, Pasquale; Troncy, Raphaël, “DOREMUS to Schema.org: Mapping a complex vocabulary to a simpler one”, EKAW 2016, 20th International Conference on Knowledge Engineering and Knowledge Management, Poster Track, November 19-23, 2016, Bologna, Italy.

[LAFTT16] Lisena, Pasquale; Achichi, Manel; Fernandez, Eva; Todorov, Konstantin; Troncy, Raphaël, “Exploring linked classical music catalogs with OVERTURE”, ISWC 2016, 15th International Semantic Web Conference, Poster Track, October 17-21, 2016, Kobe, Japan.

[EMPIPRW16] van Erp, Marieke; Mendes, Pablo; Paulheim, Heiko; Ilievski, Filip; Plu, Julien; Rizzo, Giuseppe; Waitelonis, Joerg. Evaluating entity linking: An analysis of current benchmark datasets and a roadmap for doing a better job. In 10th edition of the Language Resources and Evaluation Conference (LREC 2016), 23-28 May 2016, Portoroz, Slovenia.

[P16] Plu, Julien. Knowledge extraction in web media: At the frontier of NLP, machine learning and semantics. In 25th World Wide Web Conference (WWW 2016), PhD Symposium, April 11-15, 2016, Montreal, Canada.

[L16] Lisena, Pasquale. Modeling, exploring and recommending music in its complexity. In 20th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2016), Doctoral Consortium Track, November 19-23, 2016, Bologna, Italy.

6. DATA SCIENCE DEPARTMENT TEAM

Professors

FILIPPONE, Maurizio
HUET, Benoit
LOISEAU, Patrick
MERIALDO, Bernard
MICHIARDI, Pietro
TRONCY, Raphaël

Postdoc, Research Engineers and Visiting Scientists

BEN AHMED, Olfa
MILANESIO, Marco
PIDOU, Paul
VENZANO, Daniele
RIZZO, Giuseppe
VISINTAINER, Roberto

PhD Students

ANDREOU, Athanasios
CUTAJAR, Kurt
DEBATTY, Thibault
DOMINGUES, Remi
FRANCIS, Danny
GHORRA, Antoine
HA, Son-Hai
KAVOOSIFAR, Mohammad
KLOTZ, Benjamin
LISENA, Pasquale
NGUYEN, Duc-Trung
PACE, Francesco
PALAGI, Emilie
PALUMBO, Enrico
PHAN, Duy-Hung
PLU, Julien
PULVIRENTI, Fabio
REDONDO-GARCIA, Jose-Luis
TIAN, Yongchao
VIOTTI, Paolo
VU, Quan dong
WU, Xiaohu

COMMUNICATION SYSTEMS DEPARTMENT

INTRODUCTION

Mobile communications are facing unprecedented changes. The emergence of multimedia in the mobile world means a convergence of internet, broadcast and telecommunication networks.

In order to provide services to a large population of users, the redesign of communications strategies, radio spectrum sharing policies, QoS control modes and eventually the overall architecture of mobile networks must be carried out.

The primary goals of the Communication Systems department at EURECOM are twofold: 1) Making research advances at the fundamental level in domains ranging from radio to network architecture and communication theory, and 2) Provide the necessary technical background to engineers in the area.

As such the competencies of our department must cover the whole field, encompassing the physical aspects of propagating waves, the design of receivers, including the analog, digital, coding and signal processing aspects, the various networking protocol layers, all the way to the mobile services and their applications.

The research program of the Communication Systems Dept. is presented in five main parts:

- WIRELESS SYSTEMS and PROTOCOLS
- RESOURCE OPTIMIZATION AND CROSS LAYER DESIGN
- PROPAGATION AND SYSTEM ANALYSIS
- COMMUNICATION THEORY AND TECHNIQUES
- OpenAirInterface RADIO PLATFORMS
- PUBLICATIONS

The 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.

Note: An up-to-date (although more concise) description of the department's organization and general activities can be found under http://www.EURECOM.fr/cm/

1. WIRELESS SYSTEMS AND PROTOCOLS

The Research Group addresses “upper layers” of Mobile Communications Systems (above PHY).

It also supports the OpenAirInterface Protocols development (real time and emulation experimental platforms)

The objective of this research field is to address challenging problems proposed by novel wireless systems by developing efficient communication protocols. When it is possible and desirable, we evaluate these protocols on experimental platforms.

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

- M2M communications and Internet of Things
- 5G Networking (Software Define Networks, Network Virtualization Functions)
- Vehicular Communication & Network
- Heterogeneous Networks with heterogeneous RATs
- Human, Traffic and Transportation Mobility Modeling

In the last period the LTE standard has emerged and impacted our research activities. Heterogeneous access, all IP traffic, new types of traffic (Video, M2M, Sensor, etc.) using the LTE technology yield challenging research problems.

A new field of research involving collaboration between terminals is presented in the “Mobile Social and Opportunistic Network” section.

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)
- HETEROGENEOUS NETWORKS
- MOBILE SOCIAL AND OPPORTUNISTIC NETWORKS
- 5G NETWORKING (SDN, NFV, Cloud RAN, Orchestration, …)
- M2M and INTERNET of THINGS

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, SYSTUF, MCN, , CNES SMILE. Data Tweet, Netcom, SINETIC, 4GinVITRO, HIGHTS, direct contracts with Orange Labs, direct contracts with INTEL, Com4innov, direct contracts with the C2CCC. Projects started in fall 2015 Self Power IPCOM, RADIS, ELASTIC, COHERENT are under progress.

The group keeps on developing the protocol stacks software OpenAirInterface: LTE protocol stack on radio interface, 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.

1.1 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:

- Dependable Vehicular Communication & Networking
- Mobility Modeling of Vehicles and Vulnerable Road Users
- Cooperative Localization for Enhanced Positioning
- Control Algorithms for Autonomous Vehicles in Mixed Traffic Environments

The last point has been initiated this year (2016) and is expected to be playing a critical role in the future, as it is expected to also be integrated as input or requirements for the other three domains.

The department participated to a Panel discussion related to C-ITS Professional and Educational Activities in C-ITS during the ITS European World Congress [Här:16a], and a Keynote related to the future activities between ITS-G5 and LTE-V2X for C-ITS has been given in [Här:16b].

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

1.1.1 Dependable Vehicular Communications and Networking

Dependable vehicular wireless communications will provide mean to exchange between vehicles critical information required to increase the safety of passengers in traffic. Traffic safety applications are mainly devoted to traffic control and vehicle collision avoidance services and have been identified as a critical innovation factor for ITS. More recently, autonomous driving vehicles will also require dependable vehicular communications to exchange critical environmental information and distribute control decisions between vehicles.

One of the main challenge in vehicular safety-related communications is to provide reliable 1-hop broadcast of safety-related information. While IEEE 802.11p / ITS-G5 has been adopted as the lead technology for vehicular safety-related communications, 5G LTE-D2D might change the game. EURECOM addresses challenges for both technologies to provide dependable vehicular communication to C-ITS applications. Considering IEEE 802.11p / ITS-G5, the major challenge addressed in this year has been on the coexistence between ITS-G5 and WiFi in the ITS-G5 spectrum at 5.9Ghz. Although this spectrum has been initially reserved for safety-related ITS-G5 communications, the WiFi industry recently received the right to use these bands under the condition not to create any harmful interference. Accordingly, WiFi need to rely on a detect and avoid principle and leave the ITS-G5 channels in case any ITS-G5 transmission is detected. Two proposals (Detect & Vacate, Detect & Mitigate) are being proposed at the ETSI BRAN at the US FCC. We have been involved in the evaluations and improvements of the two proposals, and to contributing to an ETSI TS on that matter. Using the iTETRIS simulation platform, we implemented both approaches and evaluated the potential harmful interferences in a urban LOS/NLOS scenario. The outcome of this study showed that in the case of outdoor scenarios (WiFi is located outside in the streets), harmful interferences may be mitigated. However, if the WiFi in indoor, then further mitigation technics, such as reduced TX power should be further evaluated [Kha:16]. A contribution to ETSI TR

103 319 has been proposed. It is expected that in the future, the coexistence will be extended to also integrate LTE LAA and LTE-D2D in the ITS-G5 spectrum.

An activity has also been started on the initial design of the next generation of the ITS-G5 technology. The current ITS-G5 is based on the IEEE 802.11a, which is an outdated technology comparted to LTE-A and to the latest WiFi standards. Mandated by the C2CCC as an industrial contract, we proposed key enhancements fully backward compatible with current ITS-G5, which could lead to major performance gain of the next generation of the ITS-G5 technology. This activity will be further extended in 2017.

Considering 5G LTE-D2D, the challenge is to define a framework for UEs to exchange broadcast messages between each other with a minor involvement of eNBs. In 2016, the topic of LTE-D2D became very dynamic and we provided a survey of the 3GPP LTE-D2D mechanisms for V2X communications [Gal:16a]. Then, a complete evaluation of a random access strategy, as well as the potential impact on a LTE-D2D decentralized congestion control has been conducted and is currently under review as a journal publication [Gal:16b]. We further evaluated the impact of a TDMA mechanisms when integrated in the LTE-D2D frequency/time resource allocations. Results showed that TDMA is strongly impacted by half-duplex impairments. We provided an enhancement in the TDMA protocol to mitigate such impairment, and as such, we showed that the TDMA scheme to perform better than a random allocation as well as ITS-G5 (from a resource allocation perspective). This activity will be further extended in 2017 by the implementation of these mechanisms on OAI.

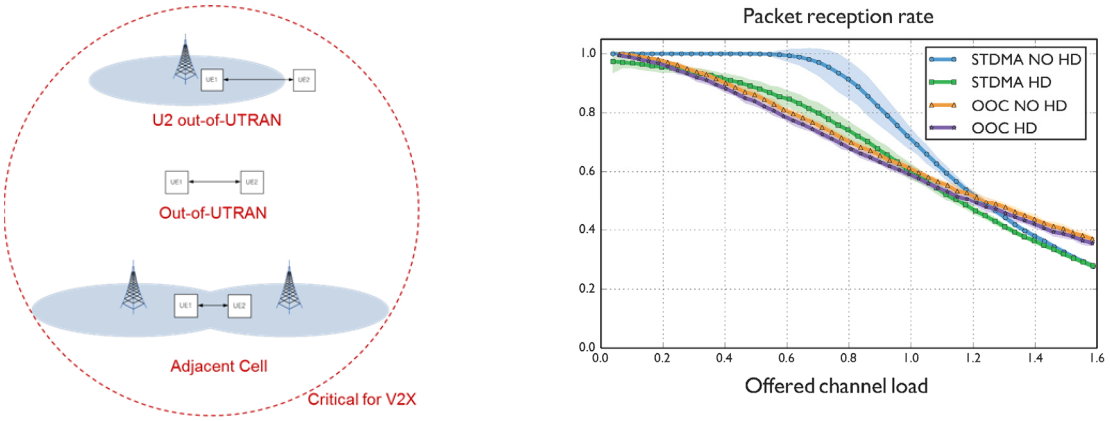


FIGURE 1: VEHICULAR LTE D2D BROADCAST COMMUNICATION (I) THREE USE CASES (II) PERFORMANCE COMPARISON (RANDOM VS. TDMA RESOURCE ALLOCATIONS)

Future autonomous driving vehicles will not only require dependable vehicular communications, but also dependable vehicular networks. It is expected that the state-of-art ICT strategies to be integrated to the C-ITS domain. In particular, we looked at two major aspects: Edge Caching as well as IoT Services. On the Edge Caching, future autonomous driving vehicles will require to obtain a large amount of data for its own navigation and environmental representation on the road. This will require the massive exchange of Local Dynamic Maps, which will need to be available as close as possible to the vehicles. In [Mah:16a], we proposed a mobility-aware caching mechanisms to prefetch the required data chunk in the caches of the appropriate Edge Nodes. We showed a 30% reducing in back-end traffic, as well as a drastic reduction of the required delay. This work has been further extended to compare with state-of-art caching strategies, and will also be further generalized in the context of Information-centric mechanisms. On the IoT Services, we proposed in [Dat:16g, Dat:16h, Dat:16l] an integrated architecture between the IoT/M2M and the ETSI ITS architectures in order to be able to integrate vehicles as connected IoT object.

We finally further worked on distributed floating car data (D-FCD) considering low penetration of ITS-G5 technologies, and including alternative Bluetooth connectivity between vehicles and with traffic lights. Results showed that considering 3% penetration of ITS-G5 technologies, the traffic state may be well evaluated using 30% penetration of Bluetooth [Jun:16].

1.1.2 Mobility Modeling of Vehicular and Vulnerable Road Users

Vehicular mobility modeling has been well studies over the past decade, which gave birth to models at various levels of granularity. But modeling and understanding mobility still remains a critical challenge for C-ITS. First, in the context of smart cities, strategic multi-modal population mobility remains particularly challenging to model and understand. Second, mobility is not restricted to vehicles and public transportation. A growing interest is centered around the so-called 'vulnerable road users', such as pedestrians, motorcycles and scooters, which are part of modern traffic, and which impact on C-ITS still remain misunderstood.

One major challenge addressed this year is to focus on the modeling of Motorcycles/scooters or more generally Powered-two wheelers (PTW). The global objective is to provide an analytic yet realistic representation of the movement of PTW among vehicular flows. One aspect investigated has been the distribution of the gap available between vehicles in multi-lanes highways, as it controls how PTW can move around them. We demonstrated that unlike previous assumptions, the distribution is not exponentially distributed but shows more a truncated Gaussian shape [Gas:16a]. We illustrated the benefit in terms of road capacity and travel time of a gradual penetration of PTW in urban traffic, and also proposed to adapt traffic lights to PTW.

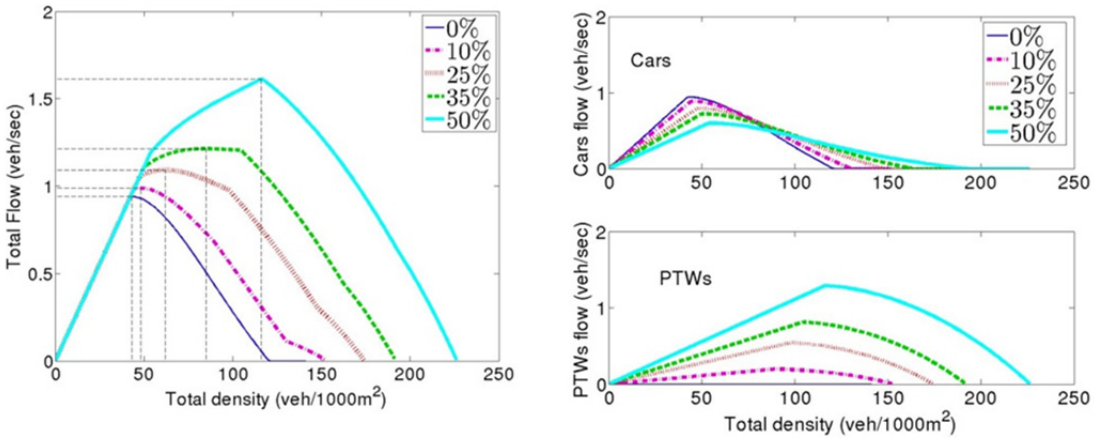


FIGURE 2: IMPACT OF SCOOTERS ON VEHICULAR TRAFFIC – SATURATION FLOW IS INCREASED (LEFT) AND THE TRAVEL TIMES ARE REDUCED FOR BOT H SCOOTERS AND CARS (RIGHT).

1.1.3 Cooperative Localization for Enhanced Positioning for C-ITS Applications

All C-ITS application, including cross-layer networking and service solutions assume the availability of GPS data, and marginally consider the lack of, or the related errors of GPS systems, in particular in obstructed areas, such as urban canyons and tunnels. Although network protocols and services may tolerate approximate GPS information, future safety-related C-ITS applications, such as Autonomous driving or detection of vulnerable road users (e.g. Scooters, Pedestrians) require extremely precise (<30cm) precision in all conditions.

Within the H2020 HIGHTS project (HIGHTS Project in the section dealing with on-going projects), we specifically addressed this issue. The general strategy is to consider neighboring vehicles as mobile anchors for cooperative localization, and as such benefit from potential higher quality positioning information from neighbors. In earlier investigations, we proposed cooperative fusion of positioning information based on Particle and Kalman filters. This year, we looked at the impact of exchanging information from the fusion filters between neighbors, and first proposed to adapt the communication strategy as fusion of the filtering error [Hoa:16c]. We then proposed an

abstraction mechanisms to compress the volume of data to be exchanged to a negligible part of the overall cooperative vehicular communication [Hoa:16b]. We finally observed that the ITS-G5 technology was not the most efficient one for range estimation and proposed to replace it with Ultra-Wide-Band Impulse Radio (IR-UWB) technology. We notably showed that if not accurately pre-processed, IR-UWB information cause more harm than good in cooperative fusion filters. But if pre-processed, IR-UWB manages to more than double the precision level [Hoa:16a]. We notably reached ~25cm precision in most of the cases.

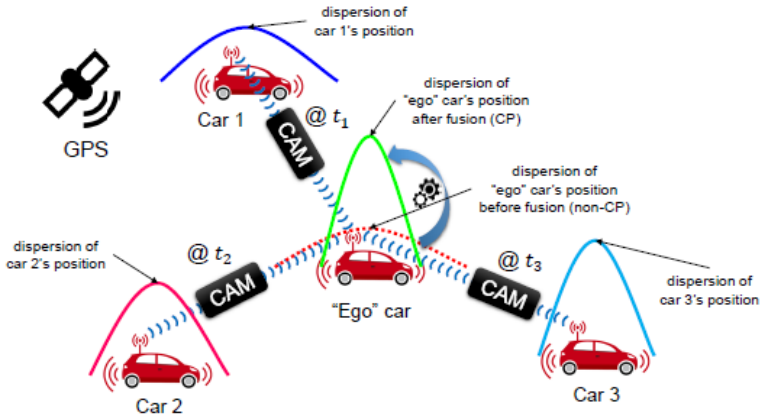


FIGURE 3: COOPERATIVE LOCALIZATION BASED ON VEHICLES CONSIDERED AS MOBILE LANDMARKS

1.1.4 Control Algorithms for Autonomous Vehicles in Mixed Traffic Environments

One major application domain of future C-ITS is related to autonomous driving vehicles. This domain is critical to the C-ITS research activities for many reasons. First, it represents a major application domain related to future Automations of Road Transportations. Second, control mechanisms for future autonomous vehicles will strongly influence the dimensioning and requirements of future vehicular communication and networking mechanisms. Finally, its particular mobility will need to be adapted and extended to account for lesser autonomous vehicles and vulnerable road users. In short, this domain is promising domain on its own but will also be strongly required to dimension the three other domains previously described.

Accordingly, we focused on control mechanisms for future autonomous vehicles in the particular case of mixed traffic scenario, where these advanced vehicles will need to share roads with other lesser autonomous or connected vehicles. Focusing on breaking mechanisms, we proposed in [Pat:16a, Pat:16b] an extension of ACC mechanisms to anticipate what a manually driving vehicles could do and accordingly adapt the control response. This is a new research domain in C-ITS, as well as for the department and we will further extend it in the next years.

FIGURE 4: AUTONOMOUS DRIVING VEHICLES IN MIXED TRAFFIC SCENARIOS

1.2 Standardization Efforts in C-ITS

The ComSys department is very active in C-ITS standardization activities. In 2016, it actively contributed to three different bodies: ETSI, CAR 2 CAR and IETF.

1.2.1 ETSI ITS and BRAN

In 2016, the CM dept. has been very active in in two ETSI ITS standardization groups: the ETSI ITS and the ETSI BRAN (Broadband Radio Access Networks) technical committees.

One major activity has been on the recent authorization for WLAN to access the 5.9GHz spectrum. Previously restricted to ITS-G5 technologies, WLAN devices have been authorized to access it under the strict condition of not creating any harmful interference. Accordingly, we have been asked by the ETSI BRAN, and in cooperation with Toyota USA, to evaluate different coexistence mechanisms and propose modifications to mitigate these interference.

In ETSI ITS, we further carried our activities in decentralized congestion control (DCC), this time from a Facilities-layer perspective. As an initial work, we implemented the current DCC_FAC specification and will evaluate its performance compared to pure access layer DCC. We will then provide feedback to the ETSI standards accordingly.

1.2.2 CAR 2 CAR

EURECOM is WG Communication co-chair and also Chair of the sub-WG on Congestion Control. Accordingly, we initiated an activity in defining the new requirements for DCC from the next Generation of C-ITS applications, such as Autonomous Driving Vehicles, Vulnerable Road Users or Rail Transportation. Yet the major activity in 2016 has been on an industrial contract between the CAR 2 CAR and EURECOM on two major domains.

First, we have been asked to integrate the ETSI BRAN co-existence mechanisms to the CAR 2 CAR profile, and propose recommendation on if and how alternative technologies could coexist with ITS-G5 (WLAN but also LTE-D2D). Second, we also initiated a highly innovative work on designing the next generation ITS-G5 technology, which could compete against the future LTE-V2X technology.

1.2.3 IETF

A New WG called IPWAVE has been created in 2016, which will focus on IPv6 mechanisms for vehicular communication and networking environments. EURECOM has been part of the early active contributors to the charter and contributed to two IETF drafts [Jeo:16, Pet:16]. The first one surveys the IPv6 challenges in vehicular communication and networking, whereas the second provides a specification on how to use IPv6 on ITS-G5 technology (without using ETSI ITS or IEEE 1609 architectures).

1.3 Heterogeneous Networks

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.

1.3.1 Mobile Data Offloading

Operators are having a difficult time keeping up with the exponential increase in data traffic demand. As a result, operators are investigating ways to offload as much data as possible through alternative means, such as WiFi networks and D2D communication. We have attempted to model both problems analytically, in order to better

understand the impact of different network parameters on performance, as well as the feasibility of each approach. We have also used these models to optimize various aspects of offloading. In [MehSpy16a], we analyze scenarios where users that opportunistically offload download requests to WiFi access points in range, if any, while they switch to the cellular network when there is no WiFi coverage. This gives rise to a queueing theoretic model with variable service, with the service rate being a function of a stochastic mobility model (for moving in and out of WiFi range). Furthermore, in [MehSpy16b] we consider the case of "delayed offloading", where users can postpone a download, up to a chosen deadline, until an open WiFi access point is found or the deadline expires (in which the content is fetched from the cellular connection). We propose a queueing analytic model to predict the performance then use it to optimize the average deadline as a tradeoff between the offloaded bytes (assuming bytes downloaded from the cellular infrastructure are more expensive money- or energy-wise) and user QoE (in terms of mean per flow delay). Finally, in [VigSpy16] we propose the use of vehicles acting as mobile relays and local caches. We then formulate the problem of optimal allocation of content to vehicular caches and derive efficient close-to-optimal algorithms as a function of mobility characteristics, content characteristics and storage capacity.

1.3.2 Optimal User Association in Small Cell Networks

User association in HetNets is a complex tradeoff, between maximizing the performance of the user in question, and improving the network-wide performance through load-balancing. In order to optimally make this decision in future networks, one needs to carefully consider the different types of traffic generated by UEs (best-effort and dedicated) as well as uplink and downlink traffic. In addition to traffic type, an important consideration in future small cell networks is that of limited backhaul capacity. In [SapSpy16], we formulate an optimization problem that considers traffic types, backhaul link capacity, and backhaul link topology. We show that this problem is convex, and use a penalty function approach to obtain optimal association rules that depend on the backhaul routing cost for each BS, in addition to the radio access properties and BS load. We then derive a fully distributed implementation and show that it converges to the optimal solution.

1.4 Mobile Social and Opportunistic Networks

Opportunistic or Delay Tolerant Networks (DTNs) are envisioned to support communication in case of failure or lack of infrastructure (disaster, censorship, rural areas). At the same time, the wide spread of smart phones and the access they provide to existing and novel suggests that direct mobile-to-mobile communication can be leveraged to harvest the large amounts of unused bandwidth between wireless devices in proximity. Epidemic algorithms are often the basis for communication in Opportunistic or Delay Tolerant Networks (DTNs). To ensure analytical tractability, existing analyses of epidemic spreading predominantly consider homogeneous contact rates between nodes. However, this assumption is generally not true in real scenarios. In [SerSpy16a], we consider classes of contact/mobility models with heterogeneous contact rates. Through an asymptotic analysis we prove that a first-order, mean value approximation for the basic epidemic spreading step becomes exact in the limiting case (large network size). We further derive simple closed form approximations, based on higher order statistics of the mobility heterogeneity, for the case of finite-size networks. We validate the analytic results through extensive simulations on synthetic scenarios, as well as on real traces to demonstrate that our expressions can be useful also in scenarios with significantly more complex structure. Our results are applicable not just to DTNs, but can also be applied to other contact networks (e.g. online social networks) to model information spread. Finally. In [SerSpy16b] we further study the joint impact of content request popularity and heterogeneous contact characteristics on DTN-based content spreading of edge cached content.

1.5 5G Networking

1.5.1 Network Slicing and Virtualization

A. NETWORK SLICING

The concept of Network Slicing through efficient resource and services sharing, is increasingly gaining momentum as a promising solution that is able to meet both challenges. It relies on cloud-based approaches for the extreme sharing of bandwidth, base stations, spectrum, processing power and storage, something that creates the illusion of infinite resources, while enabling novel business opportunities for Over the Top (OTT) service providers and vertical industries. Note that the concept of Network Slices is not new and has been recently refined by NGMN, adopted and adapted by the main Telecom manufacturers. A Network Slice can be defined as a composition of adequately configured network functions, network applications and underlying cloud infrastructures that are bundled together to meet the requirement of a specific use case or business model on a tenant basis.

Our activities around 5G Network Slicing were related to novel Network Slicing architectures that are able to manage the whole Network Slice life-cycle, i.e. from definition, instantiation, deployment and management, for integrated 5G communications involving the heterogeneous wireless domain. Although the Network Slicing concept seems as a natural evolution of virtualization, even for some simple use cases, its realization can be extremely complex. In order to simplify the steps towards integrated operational network slices, the relevant activities can be grouped in distinct phases. We consider for a simple modeling approach with the following three phases: planning, provisioning and operation.

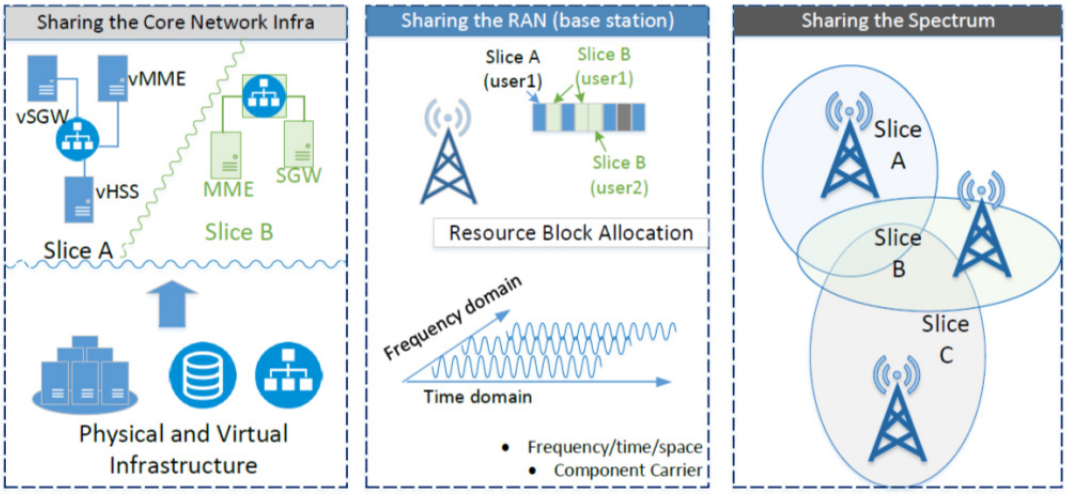


FIGURE 5: 3 TYPES OF SHARING: SHARING THE INFRASTRUCTURE, SHARING THE RAN AND SHARING THE SPECTRUM

By means of RAN resources sharing actually three dimensions of the problem exist [NN16]:

- **Sharing the Core Network Infrastructure:** The switch fabric and the core network/system (cpu, storage, memory) are virtualized and shared between tenants, for the deployment of physical or virtual network elements (e.g. SGW, MME, HSS).
- **Sharing the base station resources:** Different sharing schemes and scenarios exist where the focus stays on the way that base station (the eNodeB in LTE) resources are shared , like Resource Blocks (RB) in frequency/time/space domain.
- **Sharing the Spectrum resources**: This is related to techniques for dynamic spectrum sharing between different operators. Cognitive radio techniques fall under this category.

B. INFRASTRUCTURE SHARING

Given the resource constraints in the edge network and the highly-volatile service demand, providing differentiated performance Service Level Agreements (SLAs) to various tenants and service providers that compete for the edge network resources is not trivial. There is some previous work on SLA-driven VM placement in the generalized cloud, mostly based on demand forecasting for resource scheduling. However, in contrast to general cloud computing environments, in MEC, it is the limitation on the number of physical server resources that requires for extreme efficiency and proper modeling. Having this in mind, while also considering that a) multiple operators and stakeholders come into play concurrently to massively offer services, and b) there is lack of distinction between time-critical and non-time-critical services, the area of SLA-driven VM scheduling in the MEC context remains highly unexplored.

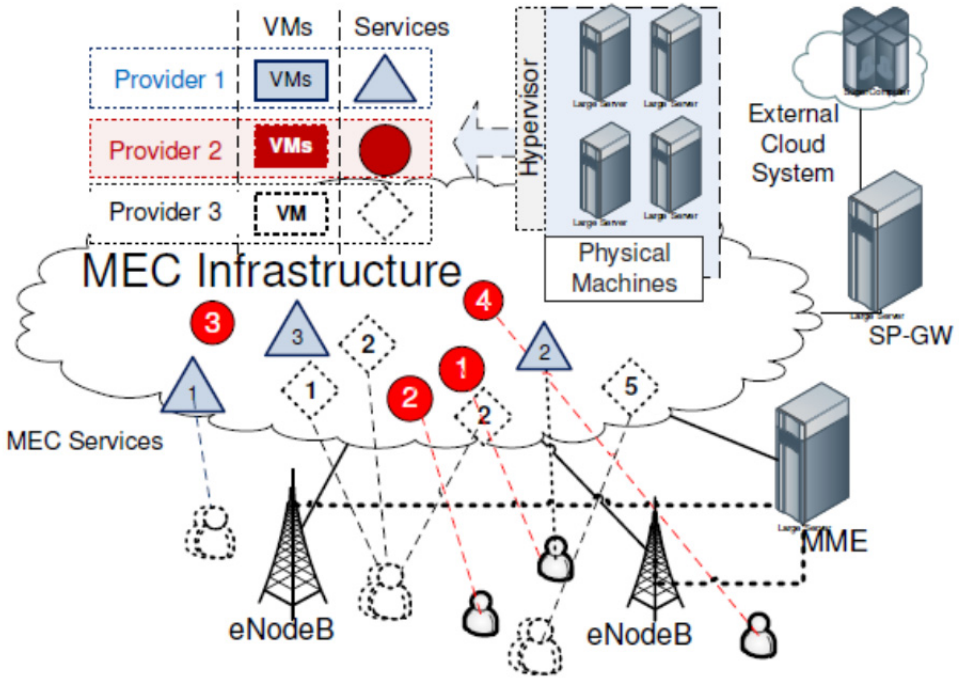


FIGURE 6: THE MEC CONCEPT IN LTE NETWORKS AND SERVICES

In [KK16b], we study a VM scheduling and placement problem, where the VMs are used to support services deployment at the edge network. In our approach, it is the SLA requirements set on a per provider basis, that drives the service deployment and the relevant VM instrumentation. Intuitively, a service with strict performance requirements must be preferred for deployment at the edge network, as compared to a service with loose performance requirements. In the case of multiple providers with a mixture of time-critical and non-time-critical services, a non-trivial scheduling problem arises for the MECIaaS provider, i.e., which service(s) from which provider(s) and where to be deployed at the edge. In our devised approach, the goal of the decision process is to maximize the MEC-IaaS provider revenue, while maximizing QoE for the customers of the time-critical services.

C. ALGORITHMS AND SOLUTIONS TO EASE NETWORK VIRTUALIZATION

Besides architectures and orchestration mechanisms, there is a need for solutions and algorithms to ease network virtualization for 5G. In this context, we studied the way of placing SDN controllers [KsBa:16] in the upcoming 5G architecture that relies on control plane and data plane function separation. The optimal placement of SDN is known to be NP-hard, hence we presented a solution, based on bargaining game, that finds a Pareto optimal solution to achieve a fair trade-off between reducing the SGW relocation and the load on each controller. Another enabler for network virtualization is the placement of VNF in the federated DCs, which is also known as NP-hard. In [YaFr:16], we first used a testbed platform to build a function (based on learning algorithms) to capture the relation between the QoE experienced by users and the resources dedicated to a VNF, in the context of HTTP video streaming. After that, the function where used in an optimization algorithm to find the optimal

placement of VNF in the DCs, which maximize user QoE. Since we are tackling an environment where users are mobile, we proposed in [TaKs:16b] the Follow Me Cloud (FMC) algorithm which ensures that VNFs are always placed in the optimal DC to serve mobile users. To reach this objective, we proposed service migration algorithms among DCs aiming at achieving zero downtime duration for end-user services.

1.6 5GDesign Patterns

In [KK16a], we focus on the control and user plane of future cloud based, SDN/NFV-enabled RAN designs. We exploit knowledge obtained from existing RAN designs, cloud computing methodologies, experience gained from SDN/NFV application, and state of the art mechanisms, while also from software engineering designs paradigms, like Service Oriented Architecture (SOA) and micro-services, to answer the following questions. How cloud computing technologies can be utilized in the mobile edge? How SDN and NFV design paradigms can be applied in the RAN? How Service Level Agreement driven SLA-Driven RAN methodologies can be applied, when diverge technologies are exploited? What is a Network Slice and how can it be implemented? What are the appropriate communication patterns in today's RAN implementations? Shall we focus on stateful or stateless communication paradigms? These are critical issues addressed and covered by this study. In this work we propose as essential and we analyze the following design patterns for cloud-based 5G communications: Distributed Shared Memory, Dedicated Data Plane Principle, Shared Control-Plane context, Agent-based VNF/VNA software, SLA-driven RAN Designs.

1.6.1 Wi-Fi Testbed Virtualization

Although the past few years have seen a massive expansion of public and home Wi-Fi installations around the world, the available wireless testbed infrastructures required to support large scale experimentally driven research are missing. The main reason is that in the best case, the number of available nodes on existing testbed installations is limited to up to hundreds. Taking into the account that many users access the testbed resources concurrently, to the best of our knowledge, there is no way and no available open wireless testbed that is able to support multiple concurrent experiments, with dense Wi-Fi installations per experiment.

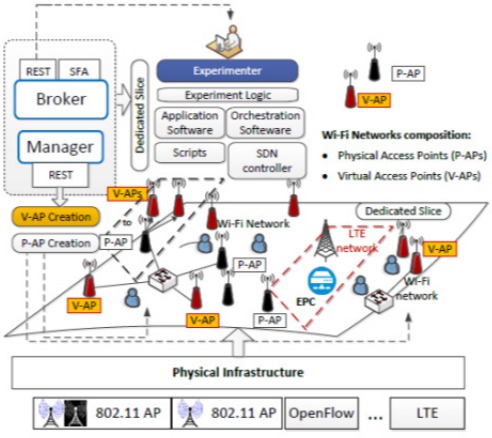


FIGURE 7: NETWORK PLANNING IN A TESTBED ENVIRONMENT WITH PHYSICAL (P-APS) AND VIRTUAL AP (V-APS)

In [KK16c] we present an approach for creating Virtual 802.11(a/b/g/e/n) Access Points (V-APs) in a wireless testbed. With the devised approach, by using a single testbed node, an experimenter can easily deploy up to a number of virtual access points. Alternatively, instead of reserving a physical node, he can just reserve a number of existing operational VAPs, which will be part of his slice reservation (see Fig.1). The time-consuming process of configuring the node as an Access Point (AP) is avoided and the time for network planning and configuration is

dropped dramatically. Thus, on one hand the experimenter is now able to create a cluster of Wi-Fi V-APs with a single REST-POST request, saving configuration effort and time, while letting the focus being on the experiment and the application. On the other hand better utilization of the testbed resources is achieved since clusters of hundreds of VAPs can be created over tens of testbed nodes.

1.7 Software-Defined Radio Access Network (SD-RAN)

Software Defined Networking (SDN) is among the key technologies considered in the context of evolving mobile networks. This is brought about by paradigm shifting ideas underlying SDN, which are the separation of the control from the data plane through a well-defined API (e.g., OpenFlow), the consolidation of the control plane and the flexibility introduced to the network through its programmability. These fundamental SDN ideas can contribute towards addressing various challenges faced by current and future mobile networks.

The radio access network (RAN) part of mobile networks offers even greater opportunities to benefit from SDN ideas [TCNN16]. One reason is that strategies and technologies being adopted to improve spectrum efficiency and scale system capacity — cell densification, use of multiple radio access technologies (e.g., LTE and WiFi), use of advanced PHY techniques like Coordinated Multipoint (CoMP), etc. — require a high level of coordination among base stations, which SDN can naturally enable. As another reason, softwarization of control in mobile networks, especially in the RAN, not only allows easier evolution to the future through programmability but also enables a wide range of use cases and novel services (as we will discuss later). At the same time, a software-defined RAN (SD-RAN) design is challenging given the unique nature of wireless resources to be managed and the stringent timing constraints associated with some keyRANcontrol operations (e.g.,MAC scheduling).

As a part of 5GPPP Coherent and FP7 FLEX projects, we designed, developed, and prototyped FlexRAN that to the best of our knowledge is the first open-source SD-RAN platform that incorporates an API to separate control and data planes that is tailored for the mobile RAN [XFNN16][XFNN16a]. FlexRAN controller design and implementation factors in the need to make real-time RAN control applications feasible. FlexRAN offers a great degree of flexibility to easily and dynamically realize different degrees of coordination among base stations reflecting centralized to fully distributed modes of operation. It offers programmability at two levels, one in the form of RAN control/management applications that can be built over the FlexRAN controller and the other within the controller to be able to update the implementation of any control function on the fly. FlexRAN is also transparent to the end-devices, aiding easier deployment and evolution.

Figure 8 provides a high-level schematic of the FlexRAN platform, which is made up of two main components: FlexRAN Control Plane and FlexRAN Agent API. The control plane follows a hierarchical design and is in turn composed of a Master Controller that is connected to a number of FlexRAN Agents, one for each eNodeB. The agents can either act as local controllers with a limited network viewand handling control delegated by the master, or in concert with other agents and the master controller. The control and data plane separation is provided by the FlexRAN Agent API which acts as the southbound API (a la OpenFlow) with FlexRAN control plane on one side and eNodeB data plane on the other side.

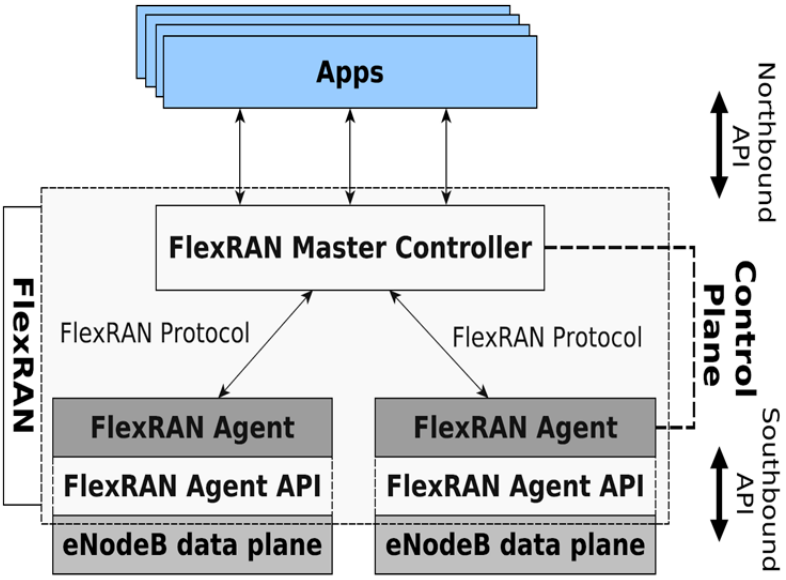


FIGURE 8: HIGH-LEVEL SCHEMATIC OF THE FLEXRAN PLATFORM

1.8 Critical Communications and Public Safety Networks

The EU FIRE+ Q4Health project aims to improve two existing FIRE platforms PerformNetworks and OpenAirInterface in order to provide better innovation services to third parties. This is done by a use case provided by the company Redzinc Services Ltd. that provides BlueEye, a wearable real-time video application for first responders, and VELOX, a virtual path slice solution to enable QoS. The project is based on a scenario in which first responders of a medical emergency (i.e. paramedics in an ambulance) have a wearable video equipment in the form of hands-free glasses with a dedicated LTE connection, and the objective is to guarantee the video transmission to a hospital where a doctor can monitor the condition of the patient in real time and suggest different treatments in its way to the emergency room. The main challenge is to achieve an interruption-free video broadcast while the ambulance pass through different LTE cells and available Wi-Fi hotspot, always within the accepted parameters defined for this type of traffic (under 150 ms for both audio and video transmissions).

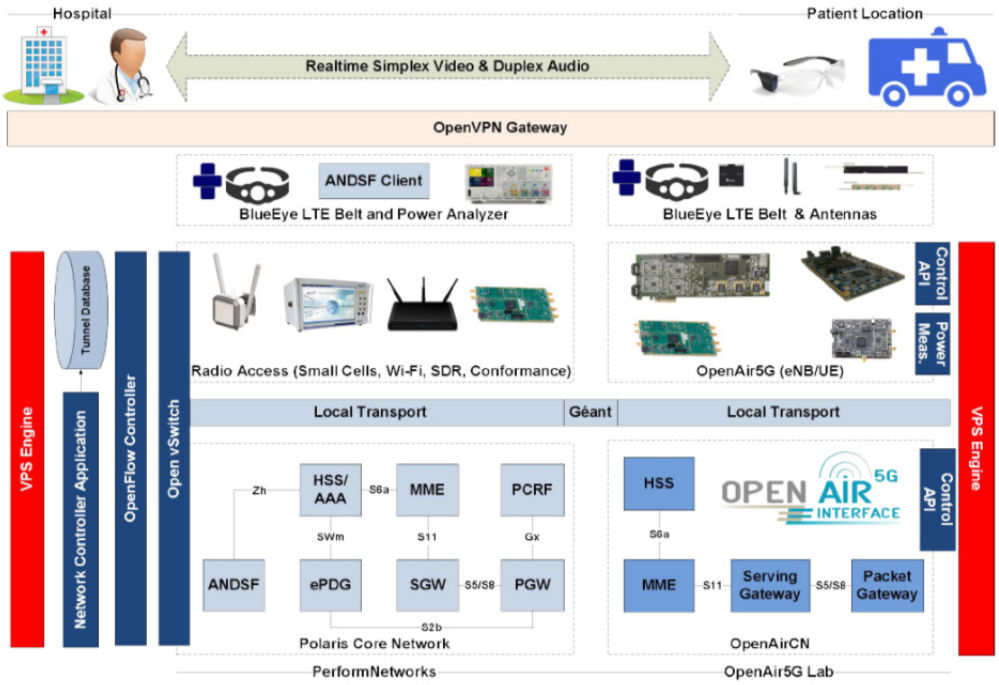


FIGURE 9: Q4HEALTH ARCHITECTURE

In [KK16e] and [KK16f] a set of experiments related to the combined experimental platform are described. The platform provides a very realistic end-to-end network, with access to the configuration of almost all of the levels of the stack. The functionality covered by the platform is focused in improving the innovation capabilities of the platform's users, but also with an eye in the latest communication trends trying to incorporate tools that enable the latest industry state of the art and best practices. The development of 5G prototypes both of EPC components and enhanced LTE radio access will boost the number of users the platforms can attract and will also result in scientific contributions.

Our focuses on the approach of the Q4Health project for optimising real-time video for emergency services over LTE, using the BlueEye platform, an eHealth video system, as the case study. The BlueEye platform is a wearable video system designed to assist first responders in public safety applications. Specifically, the project is being implemented as a set of experiments conducted over two FIRE platforms, namely PerformNetworks1[5] (formerly known as PerformLTE) and OAI. The motivation is to study video performance in scenarios with wearable live video, improving its response on LTE-A with a particular innovation focus on 3GPP release 12.

In addition the second volume of the book on the Public Safety Networks (PSNs) has been published [DCNN16]. This volume provides a system level view of some fundamental problems linked to the PSNs field. This book focus mainly on specific enabling technologies, the ones that can help the most on the deployment and usage of PSNs on real world scenarios.

1.9 Network Virtualization and Orchestration

It is generally agreed that cloud computing, Network Function Virtualization (NFV) and Software Defined Networking (SDN) will be key enabling technologies for 5G. Indeed, putting all these technologies together ensures several advantages in terms of network configuration flexibility, scalability, and elasticity, which are highly needed to fulfill the numerous requirements of 5G. Furthermore, 5G network management procedures should be as simple as possible; allowing network operators to orchestrate and manage the lifecycle of their Virtual Network Infrastructures (VNIs) and the corresponding Virtual Network Functions (VNFs), in a cognitive and programmable fashion.

Different challenges related to network virtualization towards 5G have been addressed by the communication system department during 2016. We classify these challenges into three topics:

- Network virtualization architectures
- Orchestration of virtual resources
- Algorithms and solutions to ease network virtualization

1.9.1 Network Virtualization Architectures

Different approaches have been proposed to virtualize network architecture, mainly using the ETSI NFV model as a reference. Depending on the targeted service, the virtualized network architecture maybe different. In [TaKs:16a], we proposed a generic model, namely Anything as a Service (ANYaaS), to manage important services such as mobile Content Delivery Network as a Service (CDNaaS), Traffic Offload as a Service (TOFaaS), and Machine Type Communications as a Service (MTCaaS). Ultimately, ANYaaS aims for enabling dynamic creation and management of mobile services through agile approaches that handle 5G network resources and services. Another work, we presented in [Dutta_GC], was on discussing and analysing the deployment of virtualized architecture dedicated to adapt the video rate to user bandwidth, using VNFs hosted at the Edge networks. The main idea is to instantiate a VNF at the edge, which transcodes on the fly the video to reduce or increase the video rate to maximize user Quality of Experience (QoE).

1.9.2 OAI and the Canonical (Ubuntu) JUJU Framework

Our recent collaboration with Canonical (the company behind Ubuntu), has led to the open-source exposure of the whole OAI LTE protocol stack for the EPC (MME and SP-GW and HSS), the eNodeB and Remote Radio Heads (RRH) elements using the JUJU framework and available on the charm store. The NFV concept in JUJU is built around software components called Charms. These are acting as Element Managers (EM) according to the ETSI NFV architecture and contain all the instructions necessary for deploying and configuring cloud-based services. JUJU Charms are actually the set of scripts that encapsulates the VNF. It contains all the logic you need to deploy, integrate, scale and expose the service to the out-side world. OAI is abstracted with JUJU charms while it is also a potential upstream for the ES OSM projects.

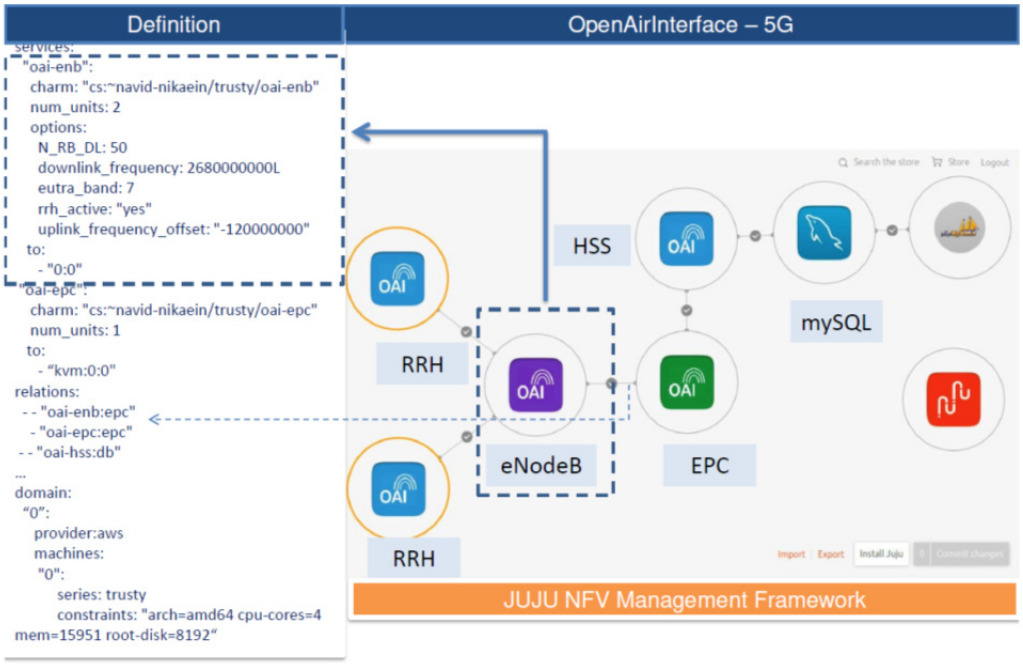


FIGURE 10: JUJU AND OAI: THE FIRST OPEN-SOURCE LTE VNF SOLUTION.

As shown in the Fig, the OAI solution can be chained together using JUJU charms relationships, while relationships can be used with other solutions like IMS and can be used to rapidly build applications over LTE such as voice and video. The implemented OAIeNB charm deploys a software managing both the radio front-end interface and the processing unit onto a single machine. With respect to the Cloud-RAN architecture paradigm, our implementation also supports different functional splits for the RRU and the BBU (see JUJU store website for more details on the OAI charms). This year we continued our activities around exposing OAI in the JUJU framework. A sample definition for a single VNF (where the OAI-eNodeB is exposed) is presented in the figure. A number of proof-of-concept demonstration have been shown during 2017 [NN16c][NN16d][NN16e]. https://jujucharms.com/q/oai

1.9.3 NFV and Multi-Domain NFV Orchestration

Instead of procedures simplification, new technologies introduced like cloud computing, SDN and NFV actually exacerbate the configuration effort required. What is now happening in practice is that the software, network and IT experts are further confused with all the additional software and hardware configuration fine tuning, even when simple functionality is needed. The goal of NFV orchestration is to meet a series of fundamental problems related to the deployment, operation and life-cycle management of network services exposed as Virtual Network Functions (VNFs). Without an orchestrator, the software tool-chain to provision resources is completely unrelated with the tools to deploy services. A specific API is used to provision resources but a completely different API is used to provision and expose services efficiently. This dramatically increases the time to deliver operational services.

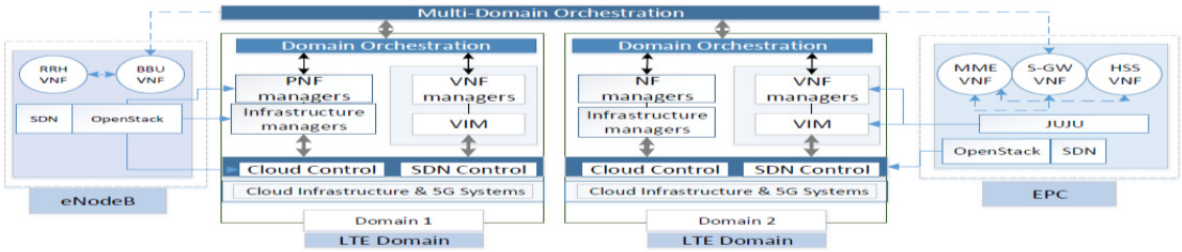


FIGURE 11: MULTI-DOMAIN NFV FOR LTE NETWORKS

In [KK16d] we focused on the problem of multi-domain orchestration for NFV over multi-technology environments. In order to facilitate service deployment in end-to-end setups (that span both vertically and horizontally the protocol stack), new orchestration designs are required that exploit and advance existing methodologies. We examined the challenges on multi-domain NFV orchestration for the general case, according to the current landscape and the existing technologies. We also proposed a reference architecture that jointly considers the challenges of NFV orchestration and supports the concept of Network Slicing. As a Network Slice, we define a composition of adequately configured network functions, network applications, and the underlying cloud infrastructure (physical, virtual or even emulated resources, Radio Access Network (RAN) resources etc.), that are bundled together to meet the requirements of a specific use case or business model. For the realization of the architecture, we provided an example use case for the LTE network and we provide an analysis for LTE-specific considerations.

1.10 Mobile Edge Computing (MEC)

1.10.1 MEC Proposed Architecture for LTE/LTE-A Networks

Towards 5G mobile networks, the low-latency and high bandwidth services are highly anticipated; however, legacy 3G and 4G networks now suffers from the mobile data surge. In this sense, pushing network services to the network edge has the potential to improve the traffic latency, user experience, and offload Internet traffic. Although the LTE/LTE-A network can highly benefit from the Mobile Edge Computing (MEC) principle, a detailed

MEC architecture is not currently in place. In this work, we propose a modular architecture for the Mobile Edge Host that is ETSI compliant and describe the functional mapping of the architecture to LTE systems. Proof-of-concept demonstrations based on the OpenAirInterface (OAI), a software implementation of LTE/LTE-A systems, present significant benefits of adopting the MEC concept in data caching use case.

An ETSI compliant modular Mobile Edge Computing (MEC) architecture with plug-in design is proposed in [CcAk:16]. The concept of ME host is discussed in the proposed MEC framework including its services and components' functionalities that aid the higher layer application development. Towards 5G network technology deployment, the mapping of the architectural components to the LTE/LTE-A system is also provided. Finally, a proof of concept is shown based on OAI under the distributed content caching use case considering different placements of ME host. As a result, the proposed MEC framework is studied under the RAN impact regime.

1.10.2 ETSI MEC and Low Latency-MEC

Moving forward 5G mobile networks, diverse usage scenarios are expected to support and enhancements are anticipated as well. Ultra-reliable-low latency is one of the most important dimensions; however, legacy 3G and 4G networks now suffers from the mobile data surge. In this sense, pushing network services to the edge has the potential to enhance user latency and experience, as well as to offload Internet traffic. Pushing network services to the network edge has the potential to improve the traffic latency, user experience, and offload Internet traffic.

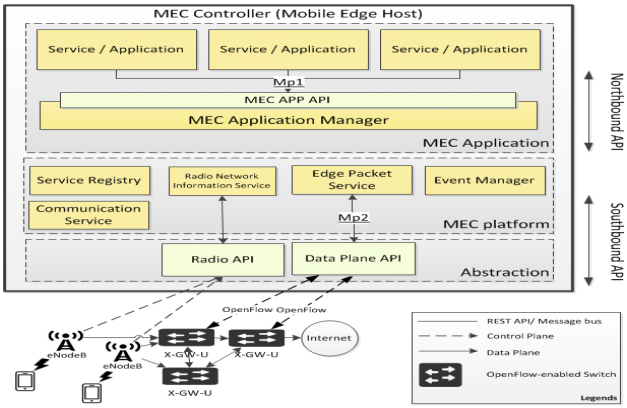


Figure 12: MEC Architecture Proposal

Introduced and specified by ETSI (ETSI MEC 001-005), Mobile Edge Computing (MEC), concurrently evolving to Multi-access Edge Computing towards 5G, is a platform enabling applications to run at the mobile edge and in close proximity to end-users. During 2016, the potential gain of building open Low-Latency-MEC platform (LL-MEC) has been observed and investigated. The research and design of LL-MEC exploiting OpenFlow has been conducted in order to build a flexible and scalable platform [CcAk:16]. The implementation of LL-MEC is initiated and progressing while several use-cases is being investigated. The main idea is to show the benefits of MEC towards 5G in different dimensions of important use-cases, e.g. enhanced mobile broadband, massive IoT, and low latency services.

1.11 Cloud RAN, Functional Split, and Ethernet-Based Fronthaul Transport Protocol

1.11.1 Toward a Fully Cloudified Mobile Network Infrastructure

Cloud computing enables the on-demand delivery of resources for a multitude of services and gives the opportunity for small agile companies to compete with large industries. Being able to leverage cloud computing technology effectively for the telco world is the focus of EU FP7 Mobile Cloud Networking (MCN) project. In this work a cloud-native radio access and core network architecture has been designed, prototyped and evaluated

[BS16][NN16]. Results show the efficiency and the simplicity that a MNO can deploy and manage the complete service lifecycle of fully cloudified, composed services that combine OTT/IT- and mobile-network-based services running on commodity hardware. The extensive performance evaluation of MCN using two key Proof-of-Concept scenarios that compose together many services to deliver novel converged elastic, on-demand mobile-based but innovative OTT services proves the feasibility of such fully virtualized deployments. Results show that it is beneficial to extend cloud computing to telco usage and run fully cloudified mobile-network-based systems with clear advantages and new service opportunities for MNOs and end-users.

1.11.2 Packetization and Scheduling of Packet-Based C-RAN Network

Being considered as a key enabler for 5G networks, Cloud-RAN (CRAN) offers advanced cooperation and coordinated processing capabilities and brings multiplexing gains. The high capacity and low latency fronthaul (FH) links requirement in the CRAN architecture can be reduced by a flexible functional split of baseband processing between remote radio units (RRUs) and Baseband units (BBUs). Under the wide adoption of Ethernet in data centers and the core network, we consider the Radio over Ethernet (RoE) as an off-the-shelf alternative for FH link in our work.

In [CcSr:16], we studied the impact of different functional splits on the FH capacity and proposed the use of a packet-based fronthaul network over RoE. Moreover, the joint impact of packetization methods and RRU-BBU functional splits on the FH rate and latency is surveyed and the potential multiplexing benefits in terms of numbers of RRHs one could support over a single Ethernet-based FH network is also provided in this work. Further in [CcNn:16], several packetization schemes and a unified packetization algorithm are proposed to minimize the FH delay incurred by the packetization over the packet-based CRAN network. Furthermore, different packet scheduling policies applied at the aggregated RRU gateway (see Fig. a) are surveyed in this work which aims to enhance the multiplexing gain in terms of the maximum number of RRUs that can be supported over the Ethernet-based FH network.

1.11.3 Cloud RAN Signal Processing Optimization with Single RF Antennas

Cloud radio access network (C-RAN) systems consisting of remote radio heads (RRHs) densely distributed in a coverage area and connected by optical fibers to a cloud infrastructure with large computational capabilities, have the potential to meet the ambitious objectives of next generation mobile networks. Actual implementations of C-RANs tackle fundamental technical and economic challenges. In the FP7 project HARP, we propose solutions to such challenges summarized in [ArtChe16]. In [ArtChe16], we present an end-to-end solution for practically implementable C-RANs by providing innovative solutions to key issues such as the design of cost-effective hardware and power-effective signals for RRHs, efficient design and distribution of data and control traffic for coordinated communications, and conception of a flexible and elastic architecture supporting dynamic allocation of both the densely distributed RRHs and the centralized processing resources in the cloud to create virtual base-stations (BSs). More specifically, we propose a novel antenna array architecture called load-controlled parasitic antenna array (LCPAA) where multiple antennas are fed by a single RF chain. Energy and spectral-efficient modulation and signaling schemes easy to implement are also provided. Additionally, the design presented for the mobile fronthaul (MFH) enables flexibility and elasticity in resource allocation to support BS virtualization. A layered-design of information control for the proposed end-to-end solution is presented. The feasibility and effectiveness of such LCPAA- and MFH-enabled C-RAN system setup has been validated through an over-the-air (OTA) demonstration.

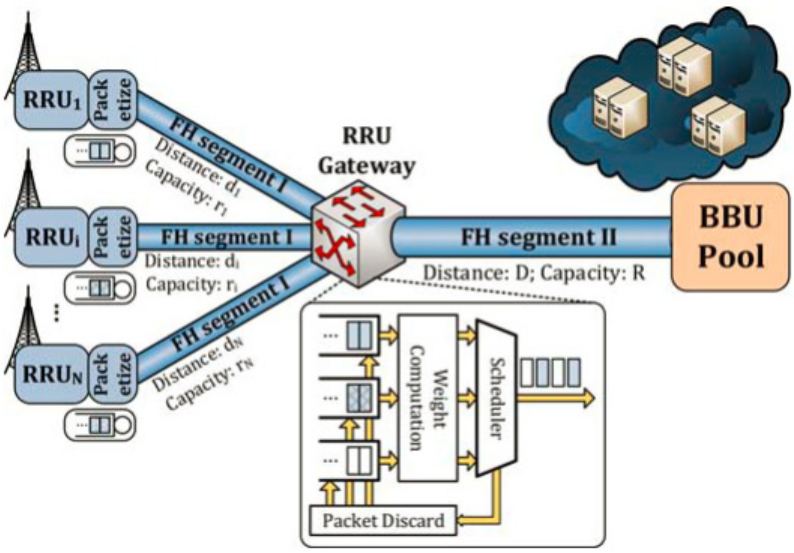


Figure 13: Considered C-RAN network topology

1.12 5G Small Cells

1.12.1 X2 HO and Multi Connectivity

A. LOAD-AWARE DECISION ALGORITHM IN NEXT-GENERATION HETNETS

Denser deployments in HetNets experience high spatio-temporal load variations, and thus require more advanced handover (HO) algorithms that consider both perspectives, jointly. In addition, power based algorithms (e.g., RSS-based) proposed in the literature cannot be applied, due to asymmetrical transmission powers among macrocells, and small cells (SCs). Especially, in such environments, a user equipment (UE) usually remains connected to a macrocell that offers high transmission power (and is potentially overloaded due to its high number of concurrent users), while is placed close to one or more underloaded SCs. We revisit the problem of HO, in the context of future HetNets deployments via an SDN approach.

A framework was constructed considering both user (i.e., RSS) and network (i.e., cell load) perspectives in [AkSn:16]. Specifically, the service delay of a random ow is analytically derived, as a key QoS criterion for the HO decision algorithm under the assumption of uneven transmission power regimes. This approach overcomes the shortcomings created by only considering RSS criteria in HO decision for HetNets. In addition, we propose a HO decision algorithm and sketch an implementable framework based on Software Defined Networking (SDN) architecture. Finally, we investigate the related tradeoffs involved, in different load variation scenarios.

B. X2 HANDOVER ANALYSIS IN LTE/LTE-A

Comparing the S1 and X2 handover in terms of the EPC signaling load and the results proofs that X2 handover can reduce EPC signaling load more than six times compared with S1 handover. X2 handover can be a sort of solution to decrease the load impact to the EPC and to increase the reliable inbound handover. The work models in related work provide LTE handover scheme on an open source platform operated on the ns-3 platform; however, it does not compare the impact of different parameters on the handover delay. Finally, this paper uses the ns-3 platform to compare the measured RSRP and RSRQ level under different parameters: vehicle velocity, eNB transmission power and distance between UE and eNB. However, there is no comparison on the handover latency on different parameters. To this end, we revisit the problem of X2 performance analysis developing the whole protocol in the OpenAirInterface (OAI) emulation platform.

We discuss and sum up the X2 handover protocol as well as its own characteristics, parameters and further extensions towards 5G network technologies in [AkNn:16]. Moreover, we investigate the impact of the handover parameters such as frequency offsets and hysteresis that are commonly used in the handover decision algorithms criteria. Finally, we analyze and characterize the performance of X2 in terms of delay using the OAI emulation platform focusing on the Evolved Universal Terrestrial Access Network (E-UTRAN). Results obtained from the OAI LTE/LTE-A emulation platform demonstrate that main delay bottleneck resides in the uplink synchronization of UE to the target eNB.

C. NETWORK GRAPH APPROACH AND MULTI-CONNECTIVITY IN 5G NETS

In principle, multi-connectivity (see Fig. b) is one of the most promising techniques that can be used to achieve high performance improvements towards 5G which is extended from the dual connectivity in LTE system since Release 12. Moreover, with the arrival of Internet-of-Things (IoT) and the network access commoditization through each kind of devices, the future networks needs not only to operate with multiple connections but also support heterogeneous QoS requirement.

To deal with this technique, we established a system-level multi-connectivity platform based on existing LTE/LTE-A system in order to examine the benefits of multi-connectivity in both network (e.g., network throughput enhancement) and user (e.g., QoS requirement fulfillment) perspective. This platform is based on network abstraction methodology to enable SDN-based control and coordination for the RAN using network graphs.

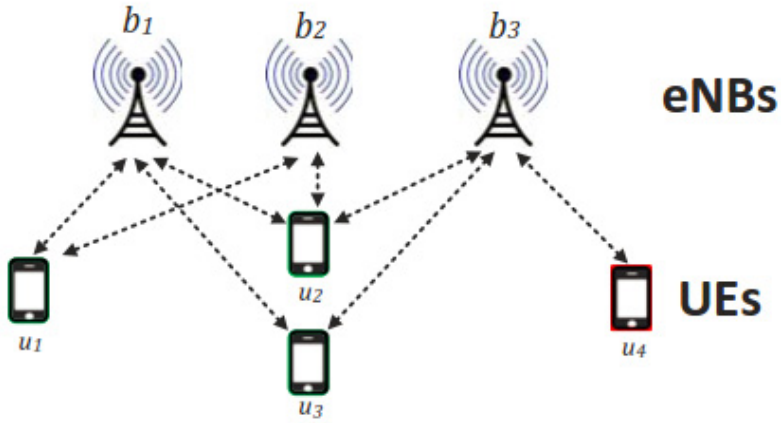


Figure 14: Multi-connectivity example

1.12.2 Self Backhauling

In 2015, we presented the concept of evolved UE (eUE) and evolved eNB (e2NB) to support the increasing mobility of cells for mobile scenarios requiring rapidly deployable networks. In 2016, we extended the studies by identifying scenarios for public safety use and the associated public safety network topologies. We summarized the current 3GPP work items to make LTE usable for public safety by providing the corresponding requirements like reliability, confidentiality, security, group and device-to-device communications. We underlined some challenges that future public safety networks will have to address as the expected traffic of said networks is increasing while volatile infrastructure operation is of utmost importance. We showed that the wireless interconnection of base stations through LTE (aka as self-backhauling) is not addressed by the 3GPP standards outside of the relay case while there is interest in providing such a possibility. We expressed that the two previously mentioned solutions (eUE, e2NB) could answer this matter by either evolving the UE capabilities to allow for multiple eNB communications and relaying (eUE) or by leveraging the LTE relay channel inside virtual UEs (e2NB) to allow self-backhauling in a mesh fashion, as partly shown on the following figure [FavApo16a][FavApo16b].

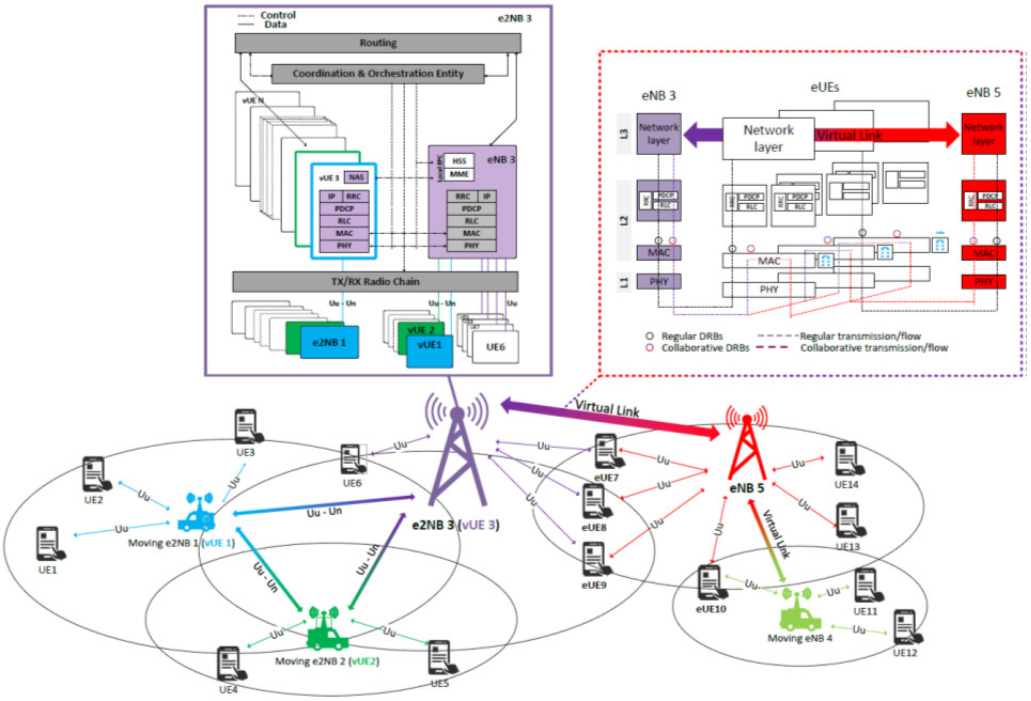


Figure 15: LTE Backhauling

We also worked on the performance evaluation of the e2NB by implementing a subset of the LTE relay channel (Un) on the OpenAirInterface platform to compare its performance against the classical UE channel (Uu) and by implementing an out-band e2NB relying on Uu to experiment multihop latency and data rates. This work is to be submitted.

1.13 M2M Communications and IoT

This research activity has started four years ago, with the vision that it may impact a lot of different application fields as Connected Cars and Smart Cities. Our activities mainly focus on applying the next-generation information and communication technology to all walks of life, embedding sensors and actuators to:

- Smart homes
- Health-care centres
- Smart power grids
- Roads & transportation systems
- Water systems
- Oil and gas pipelines

Internet of Things (IoT) and Machine-to-Machine (M2M) communication are the essences to achieve that. From the three fundamental operations of the IoT (sensing, processing, actuating), we have progressed to create an advanced, data-centric IoT architecture, securing its elements, data dissemination using Information Centric Networking (ICN) and Web of Things (WoT).

Our research activities include the following aspects:

- **Data Centric IoT Architecture**
  - Moving beyond infrastructure based IoT architecture to data centric architecture
  - Consumer centric IoT services

- IoT architecture compliant with ETSI M2M and oneM2M standards
- Design and development of building blocks like – generation subsystem (sensors), generic processing and storage subsystem (cloud or edge computing device), consumer subsystem (smartphones) and administration subsystem.
- Development of common service functions in the IoT like resource discovery, semantic based device description, device management, proxy (managing both smart and legacy things), notification and semantic based data processing.
- Creating a software framework to ease data centric and cross domain IoT application development.

- **Protocols**
  - Sensor Markup Language (SenML) and its extensions for actuators for uniform sensor and actuator metadata exchange.
  - Constrained Resource Environment (CoRE) Link Format for thing description and configuration.
  - Open Mobile Alliance Lightweight M2M (OMA LwM2M) technical specification for device management.
  - MQTT for pub-sub style M2M communications (current research item).
  - CoAP for communication among M2M components in any IoT architecture (current research item).
  - CCNx (an implementation of ICN) for dissemination of data to any consumers (current research item).
- **Security**
  - Study of security vulnerabilities in IoT architecture elements.
  - Creating a secure framework for the IoT architecture.
- **Usage scenarios**
  - Integrating connected vehicles into internet of things ecosystems by examining challenges and solutions.
  - Using our onem2m architecture based IoT framework for mobile crowd sensing in smart cities.
  - Exchanging local dynamic maps (LDM) among semi/highly autonomous vehicles.

The outcomes of these R/D activities have been demonstrated through prototypes at various ETSI events, W3C WoT meetings and IEEE conferences and we are able to quickly consider the advances as well as limitations of the proposed ideas.

1.13.1 IoT Architectural Components

The data centric approach we use relies on the formulating consumer IoT services that shift the focus on IoT data from the underlying IoT infrastructure. We aim at designing and developing a data dissemination mechanism based in Information Centric Networking. This is in line with the data centric principle of the IoT architecture. For this purpose, CCNx, an open source and optimized software stack of Named Data Networking (NDN) is utilized. We are also looking at creating a full stack software framework for easing IoT application development process. This will allow developers to focus on application logic rather than working on common service functions. The goal of this work is to provide developers with a set of open APIs and interfaces for rapid development, reduce time to market an IoT application and manage its lifecycle.

Semantic web technologies are increasingly becoming part of IoT architecture building blocks due to advantages on interoperability. Previously we have used CoRE Link Format to describe M2M devices (sensors, actuators) while in the advanced architecture we are moving towards semantic annotations in device descriptions. Such mechanism allows further granularity in an M2M world in terms of properties, events and actions. This approach is a part of the ANR DataTweet Project.

Security of IoT infrastructure and privacy of IoT data are essential for success and faster consumer adoption of IoT products and services. With an aim to achieve those, we have looked at vulnerabilities in our architecture and created a framework to secure its building blocks.

1.13.2 IOT Consumer Scenarios

Currently we are engaged with two consumer scenarios in the IoT ecosystem – (i) smart cities and (ii) connected cars. In smart cities, many problems like finding a parking spot, traffic status and important events could be solved using mobile crowd sensing (MCS). Although several research projects have considered MCS, two distinct challenges remain – (i) battery life for mobile application generating data for MCS server and (ii) lack of an interoperable solution for a global MCS service. Our oneM2M IoT architecture based solution is aimed at solving these two challenges.

With connected cars, we examined how they can be a part of the IoT ecosystem. Several challenges remain in terms of – (i) interoperability of vehicular resources, cloud platform and consumer devices, (ii) choice of appropriate connectivity, (iii) computation at an edge server and more. We provided a roadmap for solving the challenges, an IoT based framework and a prototype.

1.13.3 Engagement with Standard Development Organizations

The team is highly engaged in standardization efforts in W3C (Web of Things, SSN), ETSI M2M and oneM2M partnership projects:

- **Contribution to oneM2M WG5 (MAS: Management, Abstraction and Semantics**) – This is an ongoing effort from EURECOM and we adapt our research to latest release of oneM2M standards. Following have been disseminated so far.
  - Uniform vocabulary to describe sensors, measurements, units.
  - Designing interoperable domain knowledge
- **Contribution to IEEE Communication Society research group on IoT** – A Semantic based IoT architecture has been disseminated in the group. It will be part of a larger whitepaper schedule to release in 2017.
- **Contribution to W3C Web of Things (WoT) Interest Group on**
  - Completed leading a task force of discovery technologies leading to publication of a research paper.
  - Contribution to ongoing discussions on thing description, binding to APIs and Protocols and Security and Privacy in WoT
  - A joint whitepaper on semantic interoperability for the Web of Things has been published (combined effort from W3C, oneM2M and IEEE)
  - Another whitepaper titled “The WoT as a platform of platforms” has been submitted to the Internet Architecture Board.

1.13.4 Publication Topics in the Field of IoT

The research work of the Internet of things activity has been extensively published.

Advanced IoT Architecture

The data centric IoT architecture, its interoperation with oneM2M standard, securing its building blocks, adoption of advanced semantic reasoning and application development framework are highlighted in [DatBonHar16c], [DatBon16a], [DatBon16c], [Gyr16a], [Gyr16b].

Integration of ICN with IoT

NDN based data dissemination mechanism and adopting the same into the above architecture have been investigated in [DatBon16d], [DatBon16f].

Connected Car and IoT

Application of IoT in connected car domain along with challenges and solutions are presented in [DatBonHar16b].

Smart city and IoT

Mobile Crowd Sensing as a means of solving many smart city problems are presented in [DatBonHar16a]

Discovery, management and description of M2M devices

Semantic web technology based resource discovery, description and management of M2M devices connected to the Internet still pose a challenge. These issues are reported in [DatBon16b], [DatBon16e].

IoT and WoT Standardizations related works

Our contributions to IoT and WoT standard development organizations and adopting oneM2M architecture for research and development works are highlighted in [DatBonHar16c]. Several whitepapers and tutorials presentations disseminating our research works and prototypes are described in [GyrDat16], [Dat16a], [Dat16b], [DatBonHar16d], [Dat16c], [Dat16d], [Dat16e]

1.13.5 Mechanisms to Support Mobility in Distributed Mobility Management (DMM) with Software Define Networking

Software-Defined Networking (SDN) is a new and a very promising approach that refers to the ability to control, change and manage the behavior of the network and the network devices in a dynamic manner. It is achieved by decoupling the control plane from the data plane. The control plane is moved to a logical controller which determines how the packets should flow through the network, and configures the forwarding plan accordingly. In other words, the network devices act as a simple forwarding hardware instead of running a full stack of routing protocols as in the traditional network. Therefore, SDN allows the network to offer almost limitless capabilities for implementing new services deployed on top of the network abstraction layer and flexibly adapting to their requirements. Also, the network operator can significantly reduce the cost expended on network operation, maintenance and service development.

Since DMM and SDN have complementarity assets and share the same decoupled data and control plane paradigm, we propose in this paper a DMM solution based on SDN architecture (namely S-DMM). The mobility management intelligence is deployed as a software application on top of the controller, thus taking full advantages of SDN architecture. S-DMM has several salient features, such as: (i) distributed data plane - no need to deploy any mobility-related component at the access router, and mitigation of tunneling overhead and sub- optimal routing; (ii) centralized control plane - full view and control of network switches, and mitigation of tunneling and flow management (e.g., maintenance and keep-alive signaling), which are one of the major issues of the legacy DMM; (iii) per-flow mobility support - management of the mobility of traffic flows rather than nodes; and (iv) network independence of the underlying technologies. The experiment results from a real implementation show that S-DMM comes at no performance penalty (in terms of handover latency and end- to-end delay) compared to legacy DMM, yet at a slightly better management cost, which makes S-DMM a promising candidate for a mobility management solution in the context of 5G networks.

The results on S-DMM framework can be found in [NgTBon16a].

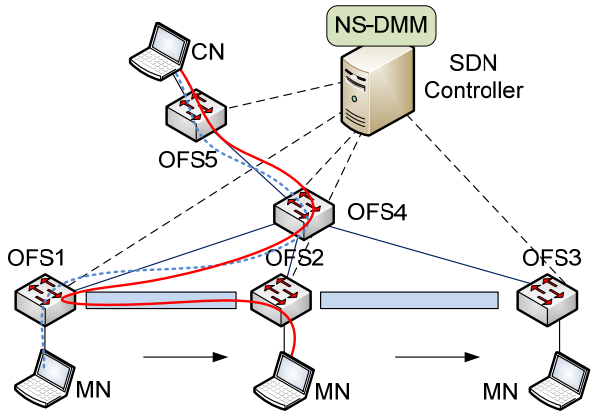


FIGURE 16: NETWORK-BASED SDN-DMM

2. RESOURCE OPTIMIZATION AND CROSS-LAYER DESIGN

2.1 Cooperative precoding in Multicell MIMO Channels with distributed CSIT

Cooperation of TXs in wireless networks allows to strongly improve the performance. To enable this cooperation, centralization of the processing in a so-called Cloud-RAN is a possible solution but it only comes with a high cost. Therefore, distributed cooperation of TXs is a practically relevant approach. Particularly challenging is the configuration where each transmitter receives its own CSIT version, i.e., corrupted with a transmitter dependent noise. In that scenario, conventional transmission schemes are known to perform badly. Although some improvements have been achieved in the particular case of two TXs, the case of more than two TXs is a challenging open problem. In the works [KerGes16,KerBazGes16], new robust transmission schemes are developed and it is shown how a strong improvement in the pre-log factor can be achieved. This is in particular obtained by developing a novel quantizer and a novel precoding scheme which enforce hierarchy between the TXs, thus strongly improving the coordination of the transmitted signals.

2.2 Optimal Use of Delayed and Imperfect CSIT at the Transmitter

How to optimally exploit the CSI at the TX is a critical question that has been open for many years. One of the most surprising advances in recent years has been to show that it was possible to exploit completely delayed CSIT to enhance the performance. This has led to a large amount of works with the goal of transforming this idea into practical performance gains. One of the most interesting model consists in assuming that both delayed and imperfect CSI are available at the TX. How to optimally exploit this CSI to maximize the pre-log factor of the sum rate was an open question that was solved in [KerGesZha16]. In this work, the optimal pre-log factor is derived as a function of the accuracy of the imperfect CSIT. The optimal pre-log factor is achieved via a novel scheme which deviates from existing efforts as it digitally combines interference, decodes symbols of any order in the MAT alignment scheme from [MadTse2012], and utilizes a hierarchical quantizer whose output is distributed across rounds in a way that minimizes unwanted interference. These jointly deliver, for the first time, the elusive DoF-optimal combining of MAT and ZF.

Rate Splitting precoding strategies for the MIMO channels with imperfect CSIT

In a multiuser MIMO broadcast channel, the rate performance is affected by multiuser interference when the channel state information at the transmitter (CSIT) is imperfect. To tackle the detrimental effects of the multiuser interference, a rate-splitting (RS) approach has been proposed recently, which splits one selected user's message into a common and a private part, and superimposes the common message on top of the private messages. The common message is drawn from a public codebook and decoded by all users. In [DaiCle16], we generalize the idea of RS into the large-scale array regime with imperfect CSIT. By further exploiting the channel second-order statistics, we propose a novel and general framework hierarchical-rate-splitting (HRS) that is particularly suited to massive MIMO systems. HRS simultaneously transmits private messages intended to each user and two kinds of common messages that are decoded by all users and by a subset of users, respectively. We analyze the asymptotic sum rate of RS and HRS and optimize the precoders of the common messages. A closed-form power allocation is derived which provides insights into the effects of various system parameters. Finally, numerical results validate the significant sum rate gain of RS and HRS over various baselines.

2.3 Cognitive and spectrum sharing systems

Decentralization Coordination of Transmitters in a License Shared Access (LSA) system

In [FilKerGes16], the operation of a Licensed Shared Access (LSA) system is investigated under minimal assumption in terms of information exchange between the transmitters so as to make the proposed scheme robust and practical. In contrast to most prior works on underlay cognitive radio, the coordination of the two TXs takes place under a realistic CSI scenario, where each TX has solely access to the instantaneous direct channel of its served terminal. Such a CSI knowledge setting brings about a formulation based on the theory of Team Decisions, whereby the TXs aim at optimizing a common objective given the same constraint set, on the basis of individual channel information. Consequently, a novel set of applicable precoding schemes consisting in letting the two TXs cooperate on the basis of the statistical information is proposed. We verify by simulations that this novel, practically relevant, coordinated precoding scheme outperforms the standard underlay CR approach.

Licensed Shared Access for the High Speed Railway Backhaul

The use of multiple antennas, location information, etc. for licensed shared access is being pursued in the FP7 project ADEL. The backhaul of a train and in particular a TGV requires very high data rates. On the other hand, the use of OFDM in 4G/5G communication systems together with fast mobility gives rise to significant Inter Carrier Interference (ICI). To mitigate this ICI, [GopSlo16a] considers a SIMO scenario and proposes to use receive beamforming to maximize the signal to (ICI) interference plus noise ratio (SINR) at any subcarrier. Independent Doppler values are considered on different multipath components. We show that it is important to first select an appropriate demodulation frequency before performing the FFT at the receiver. Location-agnostic and location-aware techniques are compared for this. With an appropriate carrier compensation, the channel variation can be approximated as linear across the OFDM symbol and receive beam forming can be done to compensate for the resulting structured ICI. For speeds up to 450km/hr, location-agnostic beam forming techniques achieve good performance with just 2 receive antennas. In [GopSlo16b], we explore a transmit beamformer design that takes into account ICI while maximizing the sum rate for a single user MIMO scenario. To perform the analysis, we first show that within the duration of an OFDM symbol, the channel variation can be assumed to be linear, even at 450kmph. We observe that, with linear channel variation, the interference due to ICI is similar to that in a multi-user interference broadcast channel (IBC). The overall problem can be viewed as that of water filling in the presence of Interference. Thus, we perform an iterative beamformer design (convergence proved) involving a Difference of Convex approach and alternating minimization. We also show that the Difference of Convex approach can be interpreted as an instance of majorization. This work was followed up in [GopSlo16c] to exploit further the excess CP available in an OFDM symbol. Along with the beamformer optimization, window coefficients were also jointly derived for the receiver that would further mitigate the ICI. The window coefficients were used to weight the OFDM symbol (including the excess CP) before the excess CP gets added to the useful part of the OFDM symbol. [GopSlo16b-c] assumed perfect channel state information at the transmitter (CSIT). [GopSlo17] addressed the fact that the CSIT could be imperfect. Gaussian CSIT with known mean and separable covariance is considered for partial CSIT. Expected WSR (EWSR) is chosen to be metric under partial CSIT and the beamformer is designed to optimize this metric. Large system approximation that also holds good when the antenna dimensions are small are employed to complete the analysis.

Important ADEL dissemination activities:

- EUR and INTEL organized a joint ADEL-SOLDER-SPEED5G workshop entitled “SAS5G: Spectrum Aggregation and Sharing for 5G Networks“ at the IEEE Int’l Symp. On Wireless Communication Systems (ISWCS), Poznan, Poland, Sept. 2016. The program can be found here: http://iswcs2016.org/welcome/menu/wp3 . The workshop was quite successful (e.g. chairs had to be added to the workshop room), in particular the panel discussion was well attended.

- EUR lead the organization of the ADEL Indian Summer School on Spectrum Aggregation and Sharing for 5G Networks (SSSAS5G) at EURECOM, Campus SophiaTech, Sophia Antipolis, France, Oct. 17-19, 2016. The program can be found here: http://www.euracon.org/index.php/2013-02-12-09-41-49/sssas5g . The event was quite successful. The presentations of the invited speakers can be found on this summer school website.
- EUR contributed to the ADEL booth at EuCNC2016 in Athens with an intermediate version of the T6.2 Proof of Concept demo. The ADEL booth won the "Best Booth Award", elected by the attendees, out of 32 booths!

Licensed Shared Access in Heterogeneous Cellular Scenarios Demo

To highlight the significance of the LSA paradigm a demo was setup using the EURECOM Open Air Interface. The demo used the Massive MIMO platform and it's Octave based interface for over the air transmission. The Massive MIMO antenna array was used to mimic two BSs of 23 antennas each. Additional 2 antennas were used for 2 UEs. Assuming a TDD framework, the UL channel estimates were transformed to the DL channel estimates with the calibration factors. Hence, as a first step, the antennas were calibrated. The estimation was done over several frames and was done in a non-coherent manner that minimized the requirement on the channel coherence time. As the pilots for channel estimation were spaced across the frequency axis, time domain based interpolation was done to estimate the channel across the entire frequency band of interest. Maximal ratio transmission (MRT), Zero Forcing (ZF) and Minimum Mean Square Error (MMSE) beamforming techniques were considered. For the MMSE beamformer, instead of estimating the interfering channel, the uplink covariance was estimated. This was then transformed to the DL covariance matrix by appropriately using the calibration parameters.

3. PROPAGATION AND SYSTEM ANALYSIS

3.1 Location Aided Wireless Communications

Indoor location estimation techniques are the subject of an ongoing Cifre PhD thesis with the company RivieraWaves, now CEVA. Location estimation is also studied in a PhD thesis on vehicular communications, in cooperation with the CEA-LETI lab. This work is also related to an ongoing H2020 ITS project called HIGHTS.

Improved Estimation of Location Related Parameters in Parametric Multi-Antenna Wireless Channels
Recently, indoor localization has attracted considerable. Indoor localization techniques can be divided into two broad categories: Radio-based (parametric) and Fingerprinting (databases). In the work discussed here, we focus on the former. In particular, we are interested in jointly estimating the Angles and Times of Arrival (AoAs and ToAs) of the multipath in a system that is comprised of a single user using a SIMO link using multi-carrier systems such as OFDM. In [BazSlo16a], we propose a Modified Minimum Description Length (MMDL) criterion that estimates the number of multipath components in the few snapshots regime, based on eigenvalue distribution results from large random matrix analysis. As for the estimation part, it is also well known that the multipaths are coherent and as a consequence all subspace methods (2D-MUSIC/2D-ESPRIT/2D-root-MUSIC/etc.) fail. We have proposed [BazSlo16b] a natural extension of the so-called “Spatial Smoothing” to the 2D case of “Spatio-Frequential Smoothing”. The advantage of the latter is that it naturally provides more subarrays to smooth over and therefore it attains a better performance. Furthermore, 2D-DIQML is asymptotically globally convergent and hence insensitive to the initialization. In [BazSlo16c], two novel single-snapshot algorithms to perform JADE are proposed using 2D Matrix Pencils whereas in [BazSlo16d] we propose a novel 2D-MUSIC algorithm that could perform JADE in the presence of local scattering, i.e. when each path is modelled as a “cluster” of incident rays showing similar ToAs and AoAs.

In [BazSlo16f], we investigate the performance of some sparse recovery and compressed sensing algorithms when applied to the Angle-of-Arrival (AoA) estimation problem.

In [BazSlo16i], a novel approach entitled ”Joint Angle and Delay Estimator and Detector”, or simply JADED, is presented. This approach allows simultaneous estimation of number of coherent/non-coherent sources and joint estimation of the angles and times of arrival of each source. The system is composed of a Uniform Linear Array (ULA) receiving known OFDM symbols from a user in an indoor environment, which is rich in multipath. Therefore, the objective is to apply the JADED approach to this scenario, which allows extraction of the Line-of-Sight component based on the first arriving path. Moreover, [BazSlo16j] presents an iterative Variational Bayes (VB) algorithm that allows sparse recovery of the desired transmitted vector. The VB algorithm is derived based on the latent variables introduced in the Bayesian model in hand. The proposed algorithm is applied to the Angle-of-Arrival (AoA) estimation problem and simulations demonstrate the potential of the proposed VB algorithm when compared to existing sparse recovery and compressed sensing algorithms, especially in the case of closely spaced sources. Furthermore, the proposed algorithm does not require prior knowledge of the number of sources and operates with only one snapshot.

Joint Antenna Array Parameter Estimation and Mutual Coupling Calibration
In [BazSlo16e], a novel algorithm for estimating the Angles of Arrival (AoA) of multiple sources in the presence of mutual coupling is derived. We first formulate an ”Equality Constrained Quadratic Optimization” problem, then derive a suitable MUSIC-like algorithm to solve the aforementioned problem, and thus obtain good estimates of the AoA parameters. Identifiability conditions of the proposed algorithm are also derived therein. Furthermore, the problem of Angle of Arrival estimation of multiple sources in the presence of mutual coupling is addressed again

in [BazSlo16g]. The presence of unknown mutual coupling between antenna array elements is known to degrade the performance of super resolution direction-finding algorithms. In this paper, we propose an iterative algorithm based on Alternating Projection in order to optimize the Maximum Likelihood cost function, which takes mutual coupling into account. Simulation results demonstrate the potential of the proposed algorithm, as it is compared to the Cramer-Rao bound of joint mutual coupling and Angle of Arrival estimation. In [BazSlo16h], the problem of Joint Angle and Delay Estimation (JADE) in the presence of mutual coupling is addressed. The system consists of a Single Input Multiple Output (SIMO) link in an OFDM communication setting. We have presented two resilient algorithms that could cope with mutual coupling. The first one is an extension of an existing algorithm in order to perform JADE, which allows resolvability of more signals. The second algorithm is an improvement of the first one, in the sense that it could handle more mutual coupling parameters. Simulation results show the difference between both algorithms. In addition, the algorithms are compared with the 2D-MUSIC method that performs JADE in the absence of mutual coupling, i.e. when a ”coupling-free” model is assumed.

In [BazSlo16m], We first present a result explaining why some traditional methods, that estimate Angles-of-Arrival (AoAs) of multiple sources in the presence of mutual coupling, suffer from an identifiability issue, when the number of coupling parameters exceeds a certain level. Then, we present a first method that estimates AoAs of sources when more coupling parameters are present, namely when the number of coupling parameters exceeds that certain level. Finally, we propose a refinement of the proposed algorithm, which could further enhance the AoA estimates. Simulation results have demonstrated the potential of the proposed method and its refined version, for different scenarios, as it enjoys better performance than existing methods. Then in [BazSlo16l], we derive a Mean-Squared-Error (MSE) expression of a recently proposed algorithm that could estimate the Angles-of-Arrival of multiple sources in the presence of more mutual coupling parameters, compared to traditional methods. The MSE expression is compared with the MSE of MUSIC with known mutual coupling parameters and to the Cramer-Rao bound (CRB) of any unbiased estimator that estimates the Angles-of-Arrival in the presence of mutual coupling. It is shown that the proposed method is asymptotically unbiased. In addition, it is also shown that the method attains CRB for large number of antennas with fixed coupling parameters and uncorrelated sources. For high SNR, the CRB is not necessarily attained, however, we study the gap between the derived MSE and the CRB. The proofs of the theorems presented in papers [BazSlo16m] and [BazSlo16l] are presented in [BazSlo16n].

4. COMMUNICATION THEORY AND TECHNIQUES

4.1 Single User Communication Techniques

mmWave MIMO Beam Alignment

Millimeter-wave (mmWave) communication is a promising technology for next-generation wireless systems. One challenging application lies in the vehicular domain, where mmWave should support ultra-fast and high-rate data exchanges among vehicles and between vehicles and infrastructure. To achieve ultra-fast initial access between nodes, we propose in [GarWym16] a location-aided beamforming strategy and analyze the resulting performance in terms of antenna gain and latency. We find that location information can significantly speed up initial access.

4.2 Massive Mimo Systems

MASSIVE MIMO SYSTEMS IN NON-RECIPROCAL CHANNELS WITH D2D

In Frequency Division Duplex (FDD) mode and, more generally, when uplink-downlink channel reciprocity does not hold, the success of massive MIMO as 5G technologies runs up against the high cost of channel state information (CSI) acquisition based on a classical closed loop feedback. In previous works, we proposed to leverage possible direct D2D links between neighboring user terminals to alleviate or annihilate the CSI acquisition costs. In [CheYin16], [CheYin16a], and [Cot16] we further develop this concept D2D concepts to make it more applicable and reliable in multiuser massive MIMO systems under limited feedback. In particular, we focus on two aspects: (i) under sufficient capacity of D2D communications, how to exploit D2D as data relaying to avoid instantaneous feedback in uplink [Cot16], and (ii) under insufficient capacity of D2D communications, how to adapt the feedback and precoding strategy according to the D2D communication quality [CheYin16], [CheYin16a].

In [Cot16], we consider fast fading massive MIMO systems in FDD mode and propose a communication scheme based on D2D cooperation and statistical CSI feedback. User terminals are conveniently clustered and cooperate within a cluster to create a virtual MIMO at a target user terminal via D2D communications. We propose and analyze the use of D2D CSI at a massive MIMO base station (BS) to enhance performance compared to statistical precoders oblivious of the underlying relaying links. In the case of uniform linear array (ULA) at the BS and slow fading channels for D2D links, the structure of an optimal precoder is characterized and an algorithm for optimal power allocation is proposed. The average sum rate of the proposed communication scheme is analyzed by numerical simulations and it is compared to a baseline protocol oblivious of the relaying CSI.

In [CheYin16], we study a cooperative feedback scheme, where the users first exchange their CSI with each other by exploiting D2D communications, then compute the precoder by themselves, and feedback the precoder to the base station (BS). Analytical results are derived to show that the cooperative precoder feedback is more efficient than the CSI feedback in terms of interference mitigation. Under the constraint of limited D2D communication capacity, we develop an adaptive CSI exchange strategy based on signal subspace projection and optimal bit partition. Numerical results demonstrate that the proposed cooperative precoder feedback scheme with adaptive CSI exchange significantly outperforms the CSI feedback scheme, even when the CSI is exchanged via limitedrate D2D communications. It is worth noting that the concepts proposed in [CheYin16] is protected by a filed patent [CheCot16]. In [CheYin16a], drawing inspiration from classical point-to-point MIMO, where efficient mechanisms are obtained by feeding back directly in the precoder (rather than channel) space,

this paper proposes a new approach to the feedback problem, where it bridges feedback in the channel space and that in the precoder space, by exploiting possible D2D coordination links. Using a so-called team decision approach for the feedback and precoding design, the users and the base station (BS) minimize a common mean squared error (MSE) metric based on their individual observations on the imperfect global channel state information (CSI). The solutions are found to take similar forms as the regularized zero-forcing (RZF) precoder, with additional regularizations to capture any level of uncertainty of the exchanged CSI, in case the D2D links are absent or unreliable. Numerical results demonstrate superior performance of the proposed scheme for an arbitrary D2D link quality setup.

Covariance CSIT in FDD Massive MIMO Systems

In the context of multi-user Massive MIMO frequency division duplex (FDD) systems, the acquisition of channel state information (CSI) cannot benefit from channel reciprocity. However, it is generally expected that covariance information about the downlink channel must be estimated and fed back by the user equipment (UE). As an alternative, it was also proposed to infer the downlink covariance based on the observed uplink covariance and a stored dictionary of uplink/downlink covariance matrices. This inference was performed through an interpolation in the Riemannian space of Hermitian positive definite matrices. In [DecGui16] we propose to rewrite the interpolation step as a Riemannian coding problem. In this framework, we estimate the decomposition of the observed uplink matrix in the dictionary of uplink matrices and recover the corresponding downlink matrix assuming that its decomposition in the dictionary of downlink matrices is the same. Moreover, since this space is of large dimension in the Massive MIMO setting, it is expected that these decompositions will be sparse. We then propose new criteria based on this further constraint.

Sum Utility Optimization in MIMO Multi-User Multi-Cell: Centralized and Distributed, Perfect and Partial CSIT, Fast and Slow CSIT

In [TabSlo16d] we consider a combined form of partial CSIT (Channel State Information at the Transmitter(s) (Tx)), comprising both channel estimates (mean CSIT) and covariance CSIT. In particular multipath induced structured low rank covariances are considered that arise in Massive MIMO and mmWave settings. For the beamforming optimization, we first revisit Weighted Sum Rate (WSR) maximization with perfect CSIT and introduce yet another equivalent approach: Weighted Sum Unbiased MSE (WSUMSE). We then turn to the partial CSIT case where we consider Expected WSR (EWSR) maximization for which EWSUMSE turns out to be a better approximation compared to the existing EWSMSE approach. These approaches also allow an uplink/downlink duality interpretation for partial CSI.

Sum Utility Optimization in MIMO Multi-User Multi-Cell: Large System Analysis

We study MISO linear precoders for downlink interference broadcast channels (IBC) systems issued from the coordinated beamforming (CoBF) techniques in which each BS sends data to its own users only but channel state information (CSI) is shared between all the BSs so that each BS can exploit its excess number of spatial dimensions to mitigate the interference generated in other cells. These optimal precoders are in general iterative algorithms, which need many iterations to converge. Moreover, we need to average over many channel realizations in order to simulate the system behavior. We propose a new algorithm based on large system approximation approach, which can be used to evaluate the performance of systems using optimal weighted sum rate (WSR) maximizing precoders.

We consider two popular approaches to design WSR maximizing precoders design: linking Weighted Sum rate (WSR) to Weighted Sum MSE (WSMSE) and the Difference of Convex functions (DC programming) approaches. The general concept, denoted as the interference alignment (IA), is to define a specific interference subspace at the interfered receiver. The interfering transmitter would have to design its transmit precoder to balance between projecting the resulting interference at the designated interference subspace at interfered receivers and projecting the desired signal away from the designated interference subspace at desired receiver.

The key advantage is that such algorithms based on large system approximation converge in a single iteration and without any need to average over realizations, which is computationally less complex; as pointed in [TabSlo16a] and [TabSlo16b] for the WSMSE and the DC approaches respectively.

The work in [TabSlo16a] has been extended to [TabSlo16c] and to [TabYua16] where we study the behavior of MIMO single stream linear precoders and where large systems analysis is used to design decentralized precoders respectively. Thus, we have used large systems not only to study precoders but to design them as well.

MASSIVE MIMO SIGNAL PROCESSING TECHNIQUES

Pilot decontamination for massive MIMO systems

In [YinCot16] and [YinCot16a], we address a very relevant problem of massive MIMO in uplink known as pilot contamination: channel estimation suffers from the residual error due to the unavoidable reuse of identical training sequences by user terminals in different cells. Interference, which originates from pilot reuse, can in principle be discriminated on the basis of the distributions of path angles and amplitudes. In [YinCot16] we propose novel robust channel estimation algorithms exploiting path diversity in both angle and power domains, relying on a suitable combination of the spatial filtering and amplitude based projection. The proposed approaches are able to cope with a wide range of system and topology scenarios, including those where, unlike in previous works, interference channel may overlap with desired channels in terms of multipath angles of arrival or exceed them in terms of received power. In particular we establish analytically the conditions under which the proposed channel estimator is fully decontaminated. Simulation results confirm the overall system gains when using the new methods.

4.3 Surveillance Networks

In [DanEliRos16] we provide an overview of semantic biometrics and discuss some of the techniques that have been proposed to extract them from image and video data. We also introduce a taxonomy for organizing and classifying soft biometric attributes, and enumerate the strengths and limitations of these attributes in the context of an operational biometric system. We explore the privacy aspects concerning the collection, use, retention and disclosure of soft biometric data. Finally, we discuss open research problems in this field. This survey is intended for researchers and practitioners in the field of biometrics.

4.4 Coded Caching In Wireless Networks

Building on the recent coded-caching breakthrough by Maddah-Ali and Niesen, the work in [ZhaEli16] considered the K-user cache-aided wireless multi-antenna (MISO) symmetric broadcast channel (BC) with random fading and imperfect feedback, and analyzed the throughput performance as a function of feedback statistics and cache size. In this setting, our work identifies the optimal cache-aided degrees-of-freedom (DoF) within a factor of 4, by identifying near-optimal schemes that exploit the new synergy between coded caching and delayed CSIT, as well as by exploiting the unexplored interplay between caching and feedback-quality. The derived limits interestingly reveal that --- the combination of imperfect quality current CSIT, delayed CSIT, and coded caching, guarantees

that --- the DoF gains have an initial offset defined by the quality of current CSIT, and then that the additional gains attributed to coded caching are exponential, in the sense that any linear decrease in the required DoF performance, allows for an exponential reduction in the required cache size.

4.5 Signal Processing Techniques Over Graphs

Large random graphs model systems consisting of a huge number of interacting entities and capture the essence of their interactions and/or relations. They are used to describe systems from the most disparate fields (social networks, connectivity of users in a wireless network, transportation and sensor networks, etc.) and are one of the most promising tools to extract information from very large data set or Big Data. For this reason, their study gathered an enormous momentum with the explosion of Big Data. Signal processing over graphs provides a theoretical framework to efficiently process large amount of data and simultaneously handle and utilize their intrinsic relations.

The presence of a denser subgraph in a large graph can model several interesting phenomena in real-world networks. The problem of detecting such a denser subgraph, called dense subgraph discovery finds applications in correlation mining, fraud detection in an auction network like ebay or in webgraphs, graph-mining and detecting communities in social networks, and other areas of signal processing, computer science, information theory. We tackle the problem of dense subgraph discovery in [KadCot16], [KadCot16a], and [KadAvr16] using global and local test statistics. In [KadCot16] and [KadCot16a] we describe a test statistic based on the L1-norm of the eigenvectors of a modularity matrix to detect the presence of an embedded Erdos-Rényi (ER) subgraph inside a larger ER random graph. We make use of the properties of the asymptotic distribution of eigenvectors of random graphs to derive the distribution of the test statistic under certain conditions on the subgraph size and edge probabilities. We show that the distributions differ sufficiently for well-defined ranges of subgraph sizes and edge probabilities of the background graph and the subgraph. This method can have applications where it is sufficient to know whether there is an anomaly in a given graph without the need to infer its location. The results we derive on the distribution of the components of the eigenvector may also be useful to detect the subgraph nodes. In [KadAvr16] and in an extended version submitted to IEEE Trans. on Signal Processing, we detect a hidden ER graph embedded in a larger sparser ER graph by applying Belief Propagation (BP). We derive two algorithms based on BP to perform subgraph detection in the presence of two kinds of side-information. The first variant of side-information consists of a set of known nodes from the subgraph called cues. The second variant consists of a set of candidate nodes that can be erroneous with a given probability. It has been shown that BP fails to correctly detect the subgraph when a parameter λ that can be interpreted as an effective signal-to-noise ratio falls below a definite threshold. In our work we consider the impact of side-information on this detection threshold. We analyze the two algorithms and show that both of them achieve asymptotically zero error for any value of the effective signal-to-noise ratio parameter λ. We validate our results through simulations on synthetic datasets as well as on a real dataset taken from the DBLP collaboration network.

4.6 On Board Blind Calibration for Satellite Systems

Wideband Digital Beamforming (WDBF) technology has become an essential part of satellite communication systems since it combines the benefits of the digital beamforming with the advantages of the wideband systems. The design of WDBFs is typically based on the assumption that the delay line is frequency independent. However, analog RF chains performing down-frequency conversions and Analog to Digital Converters (ADC) introduce delay contributions that are frequency dependent and severely impair the performance of the WDBF network. Therefore, in broadband multi-antenna satellites systems (SS) a crucial role is played by a correct

calibration to compensate for these delays. Usually, on-board calibration for bent-pipe SS equipped with WDBF is based on supervised calibration techniques using probe signals. In such a context it is appealing to develop blind calibration techniques that could outperform the supervised ones in terms of costs, spectral efficiency, power, mass of the equipment, and complexity. The impairments can be modelled as linear filters and equalized in the digital domain. In [BazCot16], we develop a blind calibration algorithm for bent pipe satellite systems (SS) equipped with multiple antennas and a transparent digital beamforming network which demultiplexes the signals in the space domain but it does not perform demodulation and decoding of the transmitted data. The system consists of multiple sources transmitting from different directions and in certain subcarriers. We propose a novel blind algorithm that (i) estimates the AoAs of all present sources in the bandwidth of interest and (ii) estimates the different gain/phases at each antenna per frequency up to an unknown impairment at a reference antenna. We provide identifiability conditions that ensure a successful parameter estimation. Finally, the potential of the proposed algorithm compared to the case of known calibration parameters is assessed by simulations.

4.7 UAV-aided Wireless Cellular Communications

In [CheYat17], we target the promising area of UAV-assisted wireless networking, by which communication-enabled robots operate as flying wireless relays to help fill coverage or capacity gaps in the networks. In order to feed the UAV's autonomous path planning and positioning algorithm, a radio map was exploited, which must be, in practice, reconstructed from UAV-based measurements from a limited subset of locations. Unlike existing methods that ignore the segmented propagation structure of the radio map, we proposed a machine learning approach to reconstruct a finely structured map by exploiting both segmentation and signal strength models. A data clustering and parameter estimation problem is formulated using a maximum likelihood approach, and solved by an iterative clustering and regression algorithm. Numerical results demonstrated significant performance advantage in radio map reconstruction as compared to the baseline.

We then considered the exploitation of UAVs in wireless networking, with which communication-enabled robots operate as flying wireless relays to help fill coverage or capacity gaps in the network. We focused on the particular problem of (automatic) UAV positioning, which is known to crucially affect the performance. Existing methods typically rely on statistical models of the air-to-ground channel, and thus, they fail to exploit the fine-grained information of line-of-sight (LOS) conditions at some locations. In [CheGes17], we develop an efficient algorithm to find the best position of the UAV based on the fine-grained LOS information. In spite of the complex terrain topology, the algorithm does not require a global exploration of a signal strength radio map. In addition, it has linear complexity to the geographical scale and is able to converge to the global optimal UAV position. Numerical results demonstrated that in a dense urban area, the UAV-aided wireless system with the optimal UAV position can almost double the end-to-end capacity from the BS to the user as compared to the traditional non-UAV-aided cellular system.

5. EXPERIMENTAL ACTIVITIES AND ERC PROJECT

5.1 OpenAirInterface-Related Activities

OpenAirInterface.org is EURECOM's experimental platform activity. The activity provides open-source (hardware and software) wireless technology platforms which target innovation in air-interface technologies through experimentation. Overall, in 2016 many projects continued from 2015 (Section 6.5.1) and several new projects began (Section 6.1.2).

In 2016 EURECOM formally welcomed three strategic members into the OAI software Alliance from the following companies to contribute to the OAI community in the format of strategic partnership:

- Orange
- TCL Communication Technology
- Ercom

This involves both financial and effort contributions to stimulate technical development around the OpenAirInterface software packages. TCL furthermore has a strong presence at EURECOM (6 engineers) to collaborate around 5G terminal development using OAI as a base.

At the time of writing (early 2017), Nokia has confirmed their willingness to join the Alliance as a strategic partner. Several strong research collaborations within Nokia in France, Germany and elsewhere have led to their decision along with the finalization of the OAI public license in accordance with Nokia and Orange.

In addition, EURECOM has established associate partnership agreements with the following institutions:

- University of Kent (United Kingdom)
- IIID (India)
- University of Washington (USA)
- Chinese Academy of Sciences (Beijing, China)
- iMinds (Belgium)
- University of Utah (USA)
- B-COM (France)
- Beijing University of Posts and Telecommunications (China)
- University of Pisa (Italy)
- Instituto de Telcomunicações (Portugal)
- University of Edinburgh (Scotland)
- Universität Bern
- Universidad Carlos III de Madrid (Spain)
- University of Essex (United Kingdom)
- TNO (Netherlands)
- UESTC (China)
- IIT Hydrabad (India)
- Universitat Politecnica de Catalunia (Spain)
- Com4Innov (France)
- Rutgers University (USA)
- Ruhr Universität Bochum (Germany)
- University of Thessaly (Greece)
- Universidad de Malaga (Spain)
- IMT (France)
- UPMC (France)

- Poznan University (Poland)
- Technical University of Munich (Germany)
- Amirkabir University of Technology (Iran)
- TU Dortmund (Germany)
- Sharif University of Technology (Iran)
- INRIA (France)
- Institute for Information Industry (Taiwan)
- Xidian University (China)
- Czech Technical University (Czech Republic)
- National Chiao Tung University (Taiwan)

ITU-T IMT-2020 Focus Group

The Focus Group on network aspects of IMT-2020 was established in May 2015 to analyse how emerging 5G technologies will interact in future networks as a preliminary study into the networking innovations required to support the development of 5G systems. In December 2015, the Focus Group received an extension to its lifetime. New Terms of Reference call for the group to engage open-source communities, influencing and taking advantage of their work by introducing them to the challenges that telecoms players must overcome in the development of the 5G ecosystem. OAI was explicitly mentioned and included in deliverables of the group. OAI was presented and demonstrated at the final ITU meeting of the group in Geneva. OAI was also used by members of the group in order to develop proof-of-concept demonstrations of network slicing methods for evolving 5G networks.

5.2 Workshops

The following workshops were organized in the OAI community:

- OAI workshop and Training sessions at EURECOM (May 2016). These were held over 3 days at EURECOM. The workshop was held on a single-day an attracted around 100 participants. The training sessions were held over two days and attracted around 80 paying participants.
- OAI 1-day workshop (Taiwan). This workshop was organized in Taiwan by the OAI Alliance associate partners III and NCTU. Two researchers from EURECOM attended and presented EURECOM and OAI activities to over 150 participants from Taiwan.

5.3 Demonstrations and Presentations

The following demonstration events were carried out:

- **FOSDEM 2016** : Presentation of OAI status and activities at the primary Open-Source software workshop in EUROPE on invitation by National Instruments
- **Mobile World Congress 2016**: Demonstration on Canonical booth. Presentation of OAI and demonstration for 3GPP RAN and CT chaimen along with ETSI director and ETSI 3GPP representative.
- **Deutsche Telekom ITG Diskussionssitzung "Funk für 5G - Lösungen für die vernetzte Welt von morgen**": Presentation and demonstration of OAI in a german industry 5G workshop on invitation by Deutsche Telekom Labs.
- **Celtic+ Event 2016** : Demonstration stand. Award for SPECTRA project (2015) results.
- **ETSI Research to Standardization Workshop:** Presentation and demonstration of CloudRAN OAI features.
- **5G Observatory:** Presentation of OAI.
- **5G Roundtable at NCC (India):** Joint presentation on OAI with IIT Hydrabad
- **EUCNC 2016:** Two Demonstration stands (and best demo award)

- **OPNFV Summit:** Two independent demonstrations of OAI were held at the summit, one by EURECOM and one by Swisscom.
- **Crowncomm 2016**: Presentation and Demonstration of OAI
- **ITU-FG on Fixed-Line Requirements for 5G**: Presentaiton and Demonstration of OAI at final meeting at the ITU headquarters in Geneva.
- **Workshop on Wireless Network Performance Evaluation:** Presentation of OAI to stimulate collaboration with users in USA

5.4 OAI Public License and Contributor License Agreement

Significant effort was dedicated to finalizing the OAI licensing strategy with industrial partners, in particular Orange and Nokia. The final form of the license has been achieved in 2016. The main contributions pertain to the use of open-source methodology for 3GPP-based systems which are governed by a strict intellectual property policy allowing for patent-pooling. This has traditionally been incompatible with open-source development because of the patent grant clause common to both Apache and GPL licenses. Although similar to the Apache V2.0 license, the OAI public license informs users that they must accept to pay royalties for intellectual property that is implemented in OAI under fair, reasonable and non-discriminatory (FRAND) terms, which is the method associated with 3GPP standards. Use o the patented procedures in OAI are granted royalty-free For non-commercial and testing use cases. In addition, a new contributor license agreement was also proposed which highlights the need for contributors to disclose their patented procedures when contributing software realizations of their intellectual property to the OAI community.

5.5 EU and National Projects on Experimental

Below we provide an overview of the department's experimental activities with the status on running and new projects involving the platform, followed by some highlights in our research

5.5.1 New projects in 2016

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 will further 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:

- The first task will consider the overall system aspects of the DUPLEX project. The refinement of the target scenario and the system requirements specification, taking into account real use cases and constraints, will be the starting point for the project implementation. Led by industrial partners, this task will ensure a consistent project development under common scope, requirements and working assumptions. The target system definition will indeed impose important constraints like for example the radio environment, the radio access technologies characteristics, the amount of available spectrum, the power available for transmissions, the required sensitivity and dynamics of the equipment, and any other relevant parameters
- Task 2 (digital cancellation) aims to go beyond the state of the in terms of digital techniques for self-interference cancellation.
- Task 3 (analog techniques) is devoted to the development of 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), which is an existing Software Defined Radio platform developed by EURECOM. 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.

Partners

- CEA-LETI
- EURECOM
- ETSA
- LEAT
- Thales Communications and Security

ANR grant: 875 936 euros

5.5.2 Ongoing Projects and Projects Finished in 2016

ELASTIC (FUI)

ELASTIC : Virtualisation des éléments physiques du réseau pour adresser le test en charge des réseaux mobiles nouvelle génération

ELASTIC extends the CloudRAN concept to include testing methodologies and dynamic deployment of HeNodeB (Home eNodeB). Its objectives are summarized as:

- Evolve testing platforms to take into account the “cloudification” of radio-access networks with a particular focus on network densification.
- Develop a CloudRAN platform based on open-source software in order to explore performance gains and energy efficiency afforded by collaborative and centralized processing.
- Show that CloudRAN concepts can be exploited in the case of HeNodeB

The ELASTIC consortium is comprised of

- ERCOM
- VIRTUOR
- EURECOM
- LIP6
- COM4INNOV

The consortium also includes a user-group comprising Orange and Monaco Telecom.

HIGHTS (H2020)

Partners: EURECOM, Jacobs University, DLR, Chalmers, CEA-LETI, BOSCH, BMW/Objective, IBEOS, ZigPos, BeSpoon, FBConsulting, Paulsconsultancy, TassInternational.

The HIGHTS project addresses high precision positioning services for future C-ITS applications, and aims at providing sub-meter (<0.5m) cooperative positioning for autonomous vehicles and vulnerable road users. HIGHTS consortium comprises strong industrial partners as well as a number of world-class universities, among others.

HIGHTS addresses this objective by combining traditional GNSS technologies with an innovative use of on-board sensing and infrastructure-based wireless communication technologies (e.g., Wi-Fi, ITS-G5, UWB tracking, Zigbee, Bluetooth, LTE...) to produce advanced, highly-accurate positioning technologies for C-ITS. The results will be integrated into the facilities layer of ETSI C-ITS architecture and in HIGHTS European-wide Positioning Service Platform.

One critical aspect is to incorporate the exchange of ranging/positioning maps and cached landmark data in cooperative communication standards, as well as develop a set of positioning & context semantics to let each C-ITS actor not only transmit its positioning and local contextual data, but also understand that of its neighbors in a global Internet-of-Thing (IoT).

EURECOM is Technical Coordinator and WP leader. EURECOM will be the main IoT innovator in the HIGHTS project. It will adapt the OneM2M IoT architecture to a vehicular context, propose IoT architecture, message and service extensions for Vehicle-to-IoT communications. EURECOM will also extend C-ITS protocols to support the critical requirements from cooperative positioning, in particular on the ITS-G5 decentralized congestion control (DCC) algorithms. Finally, EURECOM will also integrate its 5G Automotive innovations, in particular extending the LTE-D2D architecture from the SYSTUF project. EURECOM will evaluate these innovations on its V2IoT vehicular experimental platform on its campus, and contribute to IoT/M2M/ITS, ETSI ITS and 5G Automotive standards.

SINETIC (FUI)

Partners: Renault, Oktal, EURECOM, all4Tec, INRIA, ARMINE, IFSTTAR

SINETIC means “Systems that are INtegratEd and digital for a Transportation more Intelligent and Cooperative”. In a context where cooperative intelligent transport systems are increasingly present in the transportation industry, the SINETIC project aims to bring:

- Assistance in the design of cooperative systems
- Advanced multi- and co-simulation techniques to represent the technical and user impact of cooperative applications
- Establish and harmonize a multi-granularity simulation environment (system, sub-system and component level)
- Link the different simulation platforms provided by the consortium partners.

Within SINETIC, EURECOM brings its expertise in the iTETRIS C-ITS platform. EURECOM closely cooperate with Renault and Oktal in order to integrate a micro-micro ‘driver simulator’ into the iTETRIS platform, and as such to be able to closely model the impact of connected vehicles on future autonomous driving applications.

In 2015, we redesigned the ITETRIS application module in order to be able to connect and handle the driver simulator SCANER (Oktal) by the iTETRIS iCS module and to allow SCANER to take driving decision based on traffic knowledge exchanged by connected vehicles. An industrial proof-of-concept has been shown to Renault-Nissan Management in December 2015.

All innovation by EURECOM in SINETIC will be conducted on the iTETRIS platform.

RADIS - Réseau d'antennes distribuées

RADIS is a national project labelled by the Pole Region Systems & ICT Cluster (Paris) and Secure Communicating Systems (PACA) financed under the FUI GT Telecom. It is managed by Alcatel-Lucent Bell Labs France and includes Orange, E-Blink, SIMPULSE, and EURECOM. The objective of RADIS is to contribute to the evolution of radio access network architecture from 4G to 5G. In particular, the project aims at experimenting with distributed processing of radio signals between remote radio site locations (radio units) and centralized processing centers (server units) and architectures for splitting functionalities between different processing systems. EURECOM is a workpackage leader and responsible for supporting the use of OAI in the server units by the industrial members of the consortium and providing a software-based radio-unit for prototyping and validation.

SOLDER – Spectrum Overlay through Aggregation of heterogeneous DispERsed Bands (FP7 call 11)

The SOLDER project ended after 3 years in November 2016. The last year of the project was devoted to finalizing the proof-of-concepts (PoC) showing different aspects of aggregation of heterogeneous radio access technologies. EURECOM was leading the PoC 3a devoted to classical carrier aggregation in LTE-Advanced. The PoC consists of an OpenAirInterface eNB, a scheduler from IS-Wireless and a Sequans UE capable of LTE Rel10 carrier aggregation. During the course of the project, two important features were integrated into OpenAirInterface: (i) LTE Rel 10 carrier aggregation and (ii) the Functional Application Programming Interface 2.0 (FAPI 2.0) that allows to plug in 3rd party schedulers into OpenAirInterface. More details and results from trials can be found in [Kal16], [KalRou06].

Another aspect of the SOLDER project is the system level analysis of the networks aggregating LTE and WiFi networks. Our main concern was to develop analytical and semi-analytical frameworks in order to achieve system performance avoiding complex packet-level simulators. In this direction we finalized our effort to model the PHY and the MAC layer of both LTE and WiFi systems [ArvKal16a], [ArvKal16b]. Furthermore, we provided a compact mathematical framework in order to estimate the flow-level performance of the aforementioned systems (users' delay, BSs' congestion probability, etc.) [ArvSpy16a]. Additional, we explored the tradeoffs of different association schemes in Heterogeneous Networks environment [ArvSpy16b]. Additionally, we investigated analytically the tradeoff between energy efficiency and users' QoE with respect to network's density [ArvKal16c].
Last but not least the paper [HolAij06] describes a management architecture for the aggregation of multi-RAT networks.

Advanced Signal Processing in Massive MIMO (funded by Huawei French Research Center)

Please see the activities and results under the section on Massive MIMO research.

Exploiting Channel Reciprocity in Massive MIMO (funded by Huawei French Research Center)

Please see the activities and results under the section on Massive MIMO research.

SHARING (Celtic+)

The project finished in 2016. It's main goal of SHARING was to propose cost/power efficient and high capacity broadband solutions by: 1.) Enabling a flexible interference management concept in order to trigger spectral efficiency increase in future heterogeneous networks, 2.) Introducing smart and innovative offloading strategies, as well as joint RRM solutions across radio, core and backhaul networks 3.) Proposing a novel integrated architecture incorporating: seamless Inter-Radio Access Technology (RAT) service continuity, fixed/wireless convergence, device-to-device transmissions, last mile broadband backhaul, and efficient licensed/license-exempt/unlicensed spectrum usage.

NETCOM (DGE FUI)

Partners: 3ROAM, EURECOM, Insight Sip, Télécom Paris Tech, ERCOM, LEAT UNS/CNRS

To enable a unique user experience with 4G/4G+ phones, the project NETCOM (FUI - SCS) aims to develop a device that fits easily in the radio of the mobile operator network. With Netcom, 4G mobile operator may, at lower cost, increase radio coverage in the area and improve the availability of bandwidth in its operating range providing a much improved user comfort to subscribers.

This project is about research on innovative features and algorithms to achieve a new product including :

- a specific electronics,
- a reconfigurable baseband modem
- an embedded software including advanced protocols base station and mobile to connect to the 4G network / LTE
- spatial diversity and polarization antennas
- a multi-band radio adapted to the operator's frequency.

The product will be tested and validated in emulation and on a life-size experimental platform, in partnership with academic and industrial reputed in their area. Some of the building blocks from the project are the subject of byproducts.

The project requires an investment in Research and Development, thirty men years as well as specialized knowledge in different fields. The project partners bring this unique and complementary expertise in the various components needed for LTE. The NETCOM partners have the habit of collaborative work, already tested on various projects (SYMPA, 4G Invitro, Spectra ...).

The achievement of the project objectives will allow the release of commercial products for industry players: Relay Node for LTE 3Roam, the Analog Front End Modules for Insight SIP and advanced versions of Mobipass tools for Ercom.

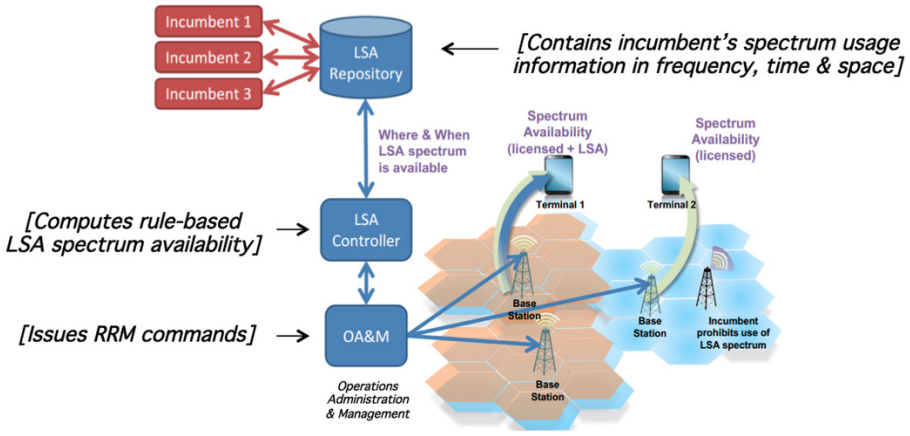
In NETCOM, EURECOM is developing algorithms for Radio Resource Management and interference management. This activity takes into account the optimization of the overall level of interference at the macro cell level, including several relays.

ADEL (FP7 CALL 11)

ADEL stands for "Advanced Dynamic Spectrum 5G mobile networks Employing Licensed shared access". To deliver the next order of magnitude gains in terms of overall spectral and radio efficiency envisioned for 2020, wireless access will have to be revisited both from a policy as well as from a technology innovation perspective. On the policy side, with the emergence of heterogeneous and small cell networks, the original “licensed vs. unlicensed” spectrum usage model has recently given way to the “licensed shared access (LSA)” paradigm wherein incumbent operators may allow other ones to share their spectrum at specific times and places, according to an agreed set of rules. Our plan is to address the following key challenges within the LSA wireless access paradigm:

- The dynamic and optimized allocation of spectral and power resources at a short time scale (on the order of seconds to even milliseconds).
- The guarantee of Quality of Service to the users of all participating spectrum-sharing networks
- The minimization of the overall energy expenditure of LSA networks.
- As key technology enablers towards these goals, we propose the utilization of:
- Decentralized spectrum sharing techniques that allow both faster decision making and less control overhead.
- Advanced collaborative sensing between the cooperating wireless networks and individual nodes for better network coordination.

- Advanced frequency agile transceiver techniques.
- Self-optimization techniques at the LSA networks to further minimise the EMF radiation and the interference caused to the incumbent network parts.



More information can be found at http://www.fp7-adel.eu/.

DataTweet (ANR-INFRA)

DataTweet proposes to develop a public service for opportunistic communications for crowd sourcing in Smart Cities. The DataTweet project proposes to explore the idea of a ubiquitous public data service for transmitting short messages in a similar way to Twitter. Any user of the service or a data source will be able to send a short message at a very low rate to some destination address over various access networks: open 802.11 hotspots, base stations of LTE, car-to-infrastructure stations of 802.11p. Partners of the DataTweet Project are: INP Grenoble, UMPC, Technicolor, EURECOM and the city of Grenoble.

The related IoT architecture considering 'Cars as Connected Objects' has been demonstrated at the ETSI OneM2M workshop [Dat:15b]. The next step will be to extend the OpenAirInterfaceITS ETSI ITS-G5 PHY/MAC with an ETSI ITS stack, and implement the oneM2M compatible M2M architecture and semantic technologies on OpenAirInterface.

All contributions of EURECOM to DataTweet will be done on the OpenAirInterface platform or on EURECOM's C-ITS experimental testbed.

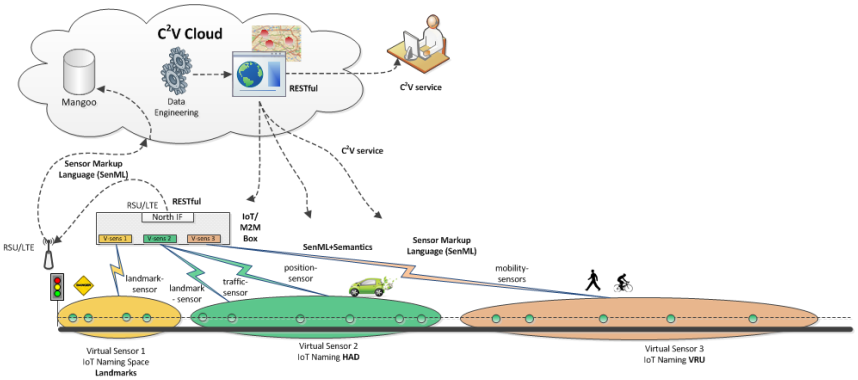


FIGURE 0.17 EURECOM'S V2IOT ARCHITECTURE AND THE EURECOM'S IOT/M2M DATATWEETBOX.

COHERENT (H2020)

The exponential growth of mobile traffic, drastically increasing of network complexity, and the strong need for inter-network coordination of wireless network resources call for breakthroughs in control, coordination and flexible spectrum management in 5G heterogeneous radio access networks. The COHERENT project aims to address these problems by researching, developing and validating a novel control framework for future mobile networks. The key innovation of COHERENT is to develop a unified programmable control framework to coordinate the underlying heterogeneous mobile networks as a whole.

The COHERENT control framework has two unique features to deal with the insufficiency of current control solutions for inter-network coordination. First, theories and methods to abstract the low layer network states and behaviors of different underlying mobile networks are developed, which provides a simplified but sufficient abstracted network view for network-wide control and resource coordination. Network abstraction will significantly reduce the signaling overhead, making scalable network-wide control solutions feasible, and enable more flexible spectrum management, which are key for the success of 5G networks. Second, based on the abstracted network view, common interfaces and software-development kits will be developed to enable programmability in controlling and coordinating heterogeneous mobile networks. The programmable control will provide operators a flexible and cost efficient way to implement new control functions and thus to support new services. The innovative impact of the COHERENT project is in enabling a unified control and coordination framework for heterogeneous mobile networks by combining innovative approaches on abstraction of low layers in underlying mobile networks, software defined networking, and flexible spectrum management. COHERENT will build a true proof-of-concept prototype to demonstrate the applicability and benefits of its approach.

Q4HEALTH (H2020)

This project is an innovation action focused on the optimization of real time video for emergency services over LTE. The project is implemented as a set of experiments conducted over the FIRE platforms PerformLTE and OpenAirInterface. The motivation is to study video performance in scenarios with wearable live video for first responders, improving its response on LTE-A with a particular innovation focus on 3GPP release 12. To achieve this goal six different experiments will be performed focused on resolving a set of six challenges identified as well as addressing a range of KPIs

Q4HEALTH faces different challenges, the inability of applications to negotiate a QoS agreement with the network, the delays introduced on live video, the appropriate scheduling algorithms on the access nodes, the service availability on indoor scenarios and the communication between geographically correlated entities. These challenges will be approached from different perspectives, the applications that will be extended to provide information regarding the type of traffic as well as their traffic requirements to the EPC and the scheduler in the RAN; the radio access where different scheduling strategies will be explored for emergency video; and the core network where mechanism to perform QoS reservation, techniques for seamless mobility between heterogeneous access technologies and SDN techniques to improve communication will be studied. Q4HEALTH has defined 20 KPIs and these will be formulated at the start as a baseline and at the end of the project we will measure the KPIs on an integrated optimized experiment in order to validate the project success.Q4HEALTH will participate in the EIT KIC as well as FIRE and 5GPPP events in order to disseminate our results in to the wider innovation ecosystem. All the project results and extensions will be showcased to maximize the exposures to other companies that might also exploit the outcomes of the project.

5GPAGODA (H2020)

5GPAGODA aims at federating Japanese and European 5G Testbeds to Explore Relevant Standards and Align Views on 5G Mobile Network Infrastructure Supporting Dynamic Creation and Management of Network Slices for Different Mobile Services. 5G!Pagoda represents the next evolution step in softwarized networks as supported by NFV, SDN and aimed at by the 5G network evolution. The top objectives of 5G!Pagoda are Transport and Risk Management for critical issues such as Climate Action and Secure Societies.

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 is in the context of wireless communications, and it aims 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.

Big Data and Biometrics: How unique is a face? (Postdoctoral grant from Labex UCN)

The project is motivated by the challenge of efficiently processing big data from facial images and facial biometrics. Given the massiveness of this data (currently running in the Exabytes), existing processing efforts fail in terms of reliability and speed. We believe that any progress in this area must come by understanding the true information content of such images, in essence understanding how unique and distinctive faces are. Towards this we propose a never-before-attempted exposition that combines our expertise in computer vision and information theory, towards understanding the entropy of faces, using custom-made `descriptive-complexity' measures that we will design, that will be based on how long of a code is needed to describe different faces. The longer the code, the more unique these faces are. An important outcome of this approach will come from our ability to understand the true limits of processing big facial data which, in layman's terms, could allow us to differentiate if suboptimal processing performance is due to the use of suboptimal algorithms, or if it is due to the fact that the class of facial images inherently does not have enough uniqueness /distinctiveness.

SPOTLIGHT (H2020-MSCA-ITN-2016)

Spotlight is a Innovative Training Network project started the 1st of January 2017 and has a duration of 4 years. With the aim of meeting and surpassing the requirements set for the 5G-and-Beyond mobile data network, SPOTLIGHT creates a fully-integrated and multi-disciplinary network of Early Stage Researchers (ESRs) that analyzes, designs, and optimizes the performance of a disruptive new mobile network architecture: the SPOTLIGHT architecture. This architecture promises to break performance limitations present to the currently loosely inter-connected, resource-fragmented and isolated in operation mobile network ecosystem, by transforming the currently loosely-coupled multitude of heterogeneous and multi-layered RATs to a flat coalition of massively distributed antenna sub-systems that are optimally orchestrated by a cloud-empowered network core. SPOTLIGHT primary aim will be to support for the first time self-including yet ultra-reliable radio communications at the edge network. To further reduce response time and enhance network resilience, all functions necessary for mobile communications will be subject of i) massive parallelization in cloud platforms at the network core and ii) big data analysis running on-top of a virtual pool of shared energy, radio, computing and storage resources at the network.

Research Contract with Thales Alenia Space – On Board Blind Calibration

The objective of this industrial project that started in the last months of 2015, is to analyze and discuss the problem of blind calibration in satellite systems for bent-pipe satellite equipped with a uniform linear antenna array and a digital transparent processor. The calibration consists in equalizing the distortion introduced by the receive antenna chains (e.g. down-frequency conversion) without demodulation and decoding of the transmitted signals. Classical blind calibration techniques are not suitable for this application since they require signal demodulation. The outcome of this work has been published in [BazCot17]. This industrial project ended in August 2016.

5.6 Equipment

The department acquired two CMW500 LTE testers (Rohde Schwarz) from the former Nvidia team in Sophia Antipolis. In collaboration with Telmat Industrie, the department fabricated the following platforms in 2016:

- 30 ExpressMIMO2 boards
- RF front-end solutions (passive and active) for LTE bands 7,38,42/43

Experimental research results

Exploiting Channel Reciprocity in Massive MIMO

Massive MIMO is one of the key enabling technologies for next generation of wireless communication systems. Our work in this topic has been focused on building the OpenAirInterface Massive MIMO testbed, which is one of the world's first LTE compliant base station equipped with large antenna array [JiaKalOpen16] [JiaKalOpen16bis]. A reduced scale version was demonstrated in June in EUCNC and has won the best booth award [JiaKalExp16].

On the theoretical part, we have mainly focused our research on the TDD channel reciprocity calibration. Some theoretical analysis on the impact of reciprocity accuracy on the final beamforming performance was carried out and have been published in SNOW workshop [JiaKalHow16] and ICC [JiaKalHow16bis] workshop.

Furthermore, we introduced a general framework which uniforms all existing calibration methods and proposed a fast calibration algorithm. This work will be presented in a journal paper which is still in preparation.

We have also proposed a calibration parameter tracking and phase jump fast detection method and a calibration method for hybrid structure massive MIMO systems. These work are submitted to our industrial partners for two patents application.

MIMO receiver architectures

We developed Reduced Complexity Maximum Likelihood Interference Aware Receiver with Parallel Detection (R-ML PIA) in OpenAIrInterface Platform (OAI). The Look-Up Table based Physical Layer Abstraction of DLSCH channel following Mutual Information Effective SNR mapping methodology was presented in the contribution mentioned at the end of this paragraph. The results showed the low complexity of the proposed method and sufficient level of accuracy to be used for LTE Link Adaptation [LukKal16].

We developed Reduced Complexity Maximum Likelihood Interference Aware Receiver with Successive Interference Cancelling (R-ML SIC IA) Receiver compared its performances in terms of throughput and performed an analysis of computational effort with the R-ML PIA receiver. It was shown that in the moderate and high SNR regime our SIC-based receiver significantly outperforms PIA receiver and the computational effort is reduced greatly [LukKal16a].

Inspired by the drive-test campaign, we developed the HARQ support for SIC and PIA receivers and showed that the SIC receiver benefits from the multi-round interference-free LLR combining. We also investigated the optimal retransmission scheme to be applied at eNodeB, when one of the codewords is successfully decoded [LukKal17].

5.7 ERC Advanced Project (PERFUME)

PERFUME is 5 year ERC Advanced funded project which started Fall 2015.

Electronic, communicating, devices have become critical to our daily lives and will increasingly impact the development of our economies and societies. Faced with the heterogeneity in the nature of such connected objects (from smartphones, tablets, wireless relays, sensors, to autonomous robot-like devices) and the growing

demands for better quality-of-service and greater data traffic volumes, future communication networks will require more intelligent organization solutions that what is currently available for traditional cellular communications.

On the bright side, our devices are becoming locally smarter and powerful (more computing power to run data intensive applications, more memory, more sensing capability). Unfortunately, these amazing capabilities are geared at local, individual performance, rather than maximizing performance collectively, as a network.

In response, PERFUME envisions a radical shift in network design and aims at enabling collective intelligence among devices (such as for instance the growing number of objects which are connected to the internet) by tapping on mutual (although partial, limited) communication capabilities between them.

PERFUME combines several fields of theoretical expertise:

- Information and communication theory
- Optimization, game, and coordination theory
- Team decision theory
- Autonomous control theory

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

6. REFERENCES

6.1 List of PhD theses completed in 2016 in the Communication Systems Department

Matzakos, Panagiotis “Scheduling and congestion management policies for QoS provision in disruption tolerant networks “. Thesis supervisor C. Bonnet. Doctoral school: EDITE Telecom Paris.

Robin Thomas, “Higher-order MIMO Detection and Half-Duplex Relay Strategies for 4G+ and 5G Networks”. Thesis supervisor: R. Knopp. Doctoral school: EDITE Telecom Paris.

Nikolaos Sapountzis, “Network Layer Optimization for Next Generation Heterogeneous Networks”. Thesis supervisor: T. Spyropoulos. Doctoral school: EDITE Telecom Paris.

6.2 Conferences and Journal Publications

[ArvKal16a] G. Arvanitakis; F. Kaltenberger; "Stochastic Analysis of Two-Tier HetNets Employing LTE and WiFi", EUCNC 2016, 25th European Conference on Networks and Communications, June 27-30, 2016, Athens, Greece.

[ArvKal16b]Technical report: G. Arvanitakis; F. Kaltenberger; "PHY and MAC layer modeling of LTE and WiFi RATs”, Research Report RR-16-317, March 24th, 2016.

[ArvKal16c] G. Arvanitakis; F. Kaltenberger, "How to Save Energy in Mobile Networks without Impacting the User Experience" submitted to ICC 2017, IEEE International Conference on Communications.

[ArvSpy16a] G. Arvanitakis; T. Spyropoulos; F. Kaltenberger; “An analytical model for flow-level performance of large, randomly placed small cell networks", GLOBECOM 2016, IEEE Global Communication Conference, 4-8 December 2016, Washington DC, USA.

[ArvSpy16b] G. Arvanitakis; T. Spyropoulos; F. Kaltenberger, "An Analytical Model for Flow-level Performance in Heterogeneous Networks" submitted to IEEE/ACM Transactions on Networking.

[AkSn:16] Load-aware handover decision algorithm in next-generation HetNets Alexandris, Konstantinos; Sapountzis, Nikolaos; Nikaein, Navid; Spyropoulos, Thrasyvoulos WCNC 2016, IEEE Wireless Communications and Networking Conference, 3-6 April 2016, Doha, Qatar.

[AkNn:16] Analyzing X2 handover in LTE/LTE-A, Alexandris, Konstantinos; Nikaein, Navid; Knopp, Raymond; Bonnet, Christian WINMEE 2016, Wireless Networks: Measurements and Experimentation, May 9, 2016, Arizona State University, Tempe, Arizona, USA.

[ArtBov16] Artuso, Matteo; Boviz, Dora; Checko, Aleksandra; Christiansen, Henrik L; Clerckx, Bruno; Cottatellucci, Laura; Gesbert; David; Gizas, Bobby; et al. Enhancing LTE with cloud-RAN and load-controlled parasitic antenna arraysIEEE Communications Magazine, Vol. PP, N°99, November 2016.

[BazCot16] Ahmed Bazzi, Laura Cottatellucci and Dirk Slock, Blind on board wideband RF calibration for multi-antenna satellites, to appear in Proc. of the 41st IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), March 2017.

[BazSlo16a] A. Bazzi, D. T.M. Slock, L. Meilhac, “Detection of the number of superimposed signals using modified MDL Criterion: a random matrix approach,” Int’l Conf. on Acoustics, Speech, and Signal Processing (ICASSP), March 2016.

[BazSlo16b] A. Bazzi, D. T.M. Slock, L. Meilhac, “On spatio-frequential smoothing for joint angles and times of arrival estimation of multipaths,” Int’l Conf. on Acoustics, Speech, and Signal Processing (ICASSP), March 2016.

[BazSlo16c] A. Bazzi, D. T.M. Slock, L. Meilhac, “Single Snapshot Joint Estimation of Angles and Times of Arrival: A 2D Matrix Pencil Approach,” Int’l Conf. on Communications (ICC), May, 2016.

[BazSlo16d] A. Bazzi, D. T.M. Slock, L. Meilhac, “On Joint Angle and Delay Estimation in the Presence of Local Scattering ,” Int’l Conf. on Communications (ICC), May, 2016.

[BazSlo16e] A. Bazzi, D. T.M. Slock, L. Meilhac, “Online Angle of Arrival Estimation in the Presence of Mutual Coupling,” IEEE Workshop on Statistical Signal Processing (SSP), June, 2016.

[BazSlo16f] A. Bazzi, D. T.M. Slock, L. Meilhac, “A Comparative Study of Sparse Recovery and Compressed Sensing Algorithms with Application to AoA Estimation,” IEEE Workshop on Signal Processing Advances in Wireless Communications, July, 2016, Invited Paper.

[BazSlo16g] A. Bazzi, D. T.M. Slock, L. Meilhac, “On a Mutual Coupling Agnostic Maximum Likelihood Angle of Arrival Estimator by Alternating Projection,” IEEE Global Conference on Signal and Information Processing (GlobalSIP), December 7-9, 2016.

[BazSlo16h] A. Bazzi, D. T.M. Slock, L. Meilhac, “A mutual coupling resilient algorithm for joint angle and delay estimation,” IEEE Global Conference on Signal and Information Processing (GlobalSIP), December 7-9, 2016.

[BazSlo16i] A. Bazzi, D. T.M. Slock, L. Meilhac, “JADED-RIP: Joint Angle and Delay Estimator and Detector via Rotational Invariance Properties,” IEEE Int’l Symposium on Signal Processing and Information Technology (ISSPIT), 2016.

[BazSlo16j] A. Bazzi, D. T.M. Slock, L. Meilhac, “Sparse Recovery using an Iterative Variational Bayes Algorithm and Application to AoA Estimation” IEEE Int’l Symposium on Signal Processing and Information Technology (ISSPIT), 2016.

[BazCot16k] A. Bazzi, L. Cottatellucci, D. T.M. Slock, “Blind on board wideband antenna RF calibration for multi-antenna satellites,” Int’l Conf. on Acoustics, Speech, and Signal Processing (ICASSP), March, 2017, Invited Paper.

[BazSlo16l] A. Bazzi, D. T.M. Slock, L. Meilhac, “Performance analysis of an AOA estimator in the presence of more mutual coupling parameters,” Int’l Conf. on Acoustics, Speech, and Signal Processing (ICASSP), March, 2017.

[BazSlo16m] A. Bazzi, D. T.M. Slock, L. Meilhac, “On mutual coupling for ULAS: estimating AOAS in the presence of more coupling parameters,” Int’l Conf. on Acoustics, Speech, and Signal Processing (ICASSP), March, 2017.

[BazSlo16n] A. Bazzi, D. T.M. Slock, L. Meilhac, “On AoA Estimation in the Presence of Mutual Coupling: Algorithms and Performance Analysis,” IEEE Transactions on Signal Processing, 2017, Submitted.

[BS16] Toward a fully cloudified mobile network infrastructure, Sousa, Bruno et al. IEEE Transactions on Network and Service Management, Vol. 13, N°3, September 2016, ISSN: 1932-4537.

[CcAk:16] MEC architectural implications for LTE/LTE-A networks Chang, Chia-Yu; Alexandris, Konstantinos; Nikaein, Navid; Katsalis, Kostas; Spyropoulos, Thrasyvoulos MOBIARCH 2016, 11th ACM Workshop on Mobility in the Evolving Internet Architecture, October 3, 2016, New York, USA.

[CcSr:16] Impact of packetization and functional split on C-RAN fronthaul performance Chang, Chia-Yu; Schiavi, Ruggero; Nikaein, Navid; Spyropoulos, Thrasyvoulos; Bonnet, Christian ICC 2016, IEEE International Conference on Communications, May 23-27, 2016, Kuala Lumpur, Malaysia.

[CcNn:16] Impact of packetization and scheduling on C-RAN fronthaul performance Chang, Chia-Yu; Nikaein, Navid; Spyropoulos, Thrasyvoulos GLOBECOM 2016, IEEE Global Communications Conference, Exhibition and Industry Forum, 4-8 December 2016, Washington, USA.

[CheCot16] Junting Chen, Laura Cottatellucci, Haifan Yin, and David Gesbert, Adaptive CSI sharing mechanism for D2D enabled precoding in massive MIMO, PTC Application Number PCT/EP2016/069606, Aug. 18,2016.

[CheYin16] Chen, Junting; Yin, Haifan; Cottatellucci, Laura; Gesbert, David, Efficient feedback mechanisms for FDD massive MIMO under user-level cooperation2nd round revision for submission to IEEE Transactions on Wireless Communications, Also submitted on ArXiv on November 22nd, 2016.

[CheYin16a] Chen, Junting; Yin, Haifan; Cottatellucci, Laura; Gesbert, David Dual-regularized precoding: A robust approach for D2D-enabled massive MIMO 50th Asilomar Conference on Signals, Systems and Computers, November 6-9, 2016, Asilomar, USA

[Cot16] Cottatellucci, Laura, D2D CSI feedback for D2D aided massive MIMO communications, SAM 2016, Sensor Array and Multichannel Signal Processing Workshop, 10-13 July 2016, Rio de Janeiro, Brazil

[CheGes16a] J. Chen and D. Gesbert, “Joint User Grouping and Beamforming for Low Complexity Massive MIMO Systems”, in Proc. IEEE Int. Workshop Signal Process. Adv. in Wireless Commun., Edinburgh, UK, Jul 3 – 6, 2016.

[CheGes17] J. Chen and D. Gesbert, “Optimal Positioning of Flying Relays for Wireless Networks: A LOS Map Approach”, to appear in IEEE Int. Conf. Commun., 2017.

[CheYat17] J. Chen, U. Yatnalli, and D. Gesbert, “Learning Radio Maps for UAV-aided Wireless Networks: A Segmented Regression Approach”, to appear in IEEE Int. Conf. Commun., 2017.

[DaiCle16] M. Dai, B. Clerckx, D. Gesbert, G. Caire "A Rate Splitting Strategy for Massive MIMO with Imperfect CSIT", IEEE Trans. Wireless Communications, to appear 2016.

[DCNN16]Wireless public safety networks 2 : A systematic approach, Câmara, Daniel; Nikaein, Navid, Elsevier, June 2016, 1st edition, ISBN:9781785480522.

[DanEliRos16] Antitza Dantcheva, Petros Elia and Arun Ross, “What else does your biometric data reveal? A survey on soft biometrics,” IEEE Transactions on Information Forensics and Security 11 (3), 441-467.

[DatBon16a] Easing IoT application development through DataTweet frameworkDatta, Soumya Kanti; Bonnet, ChristianWF-IoT 2016, IEEE World Forum on Internet of Things, 12-14 December 2016, Reston, VA, USA

[DatBon16b] A categorization of discovery technologies for the Internet of ThingsBröring, Arne; Datta, Soumya Kanti; Bonnet, ChristianIOT 2016, 6th International Conference on the Internet of Things, November 7-9, 2016, Stuttgart, Germany

[DatBon16c] Securing DataTweet IoT architecture elementsDatta, Soumya Kanti; Bonnet, ChristianICCE-ASIA 2016, International Conference on Consumer Electronics, October 26-28, 2016, Seoul, Korea

[DatBon16d] Interworking of NDN with IoT architecture elements: challenges and solutionsDatta, Soumya Kanti; Bonnet, ChristianGCCE 2016, IEEE 5th Global Conference on Consumer Electronics, October 11-14, 2016, Kyoto, Japan

[Dat16a] Connected car as an IoT serviceDatta, Soumya KantiTutorial at BMW Summer School 2016, July 18-23, 2016, Lake Tegernsee, Bavaria, Germany

[DatBonHar16a] OneM2M architecture based IoT framework for mobile crowd sensing in smart citiesDatta, Soumya Kanti; Ferreira da Costa, Rui Pedro; Bonnet, Christian; Härri, Jérôme EUCNC 2016, 25th European Conference on Networks and Communications, June 27-30, 2016, Athens, Greece

[DatBonHar16b] Integrating connected vehicles in Internet of Things ecosystems: Challenges and solutions Datta, Soumya Kanti; Ferreira Da Costa, Rui Pedro; Härri, Jérôme; Bonnet, Christian SMARTVEHICLES 2016, 3rd International Workshop on Smart Vehicles: Connectivity Technologies and ITS Applications, co-located with the WoWMoM 2016, June 21, 2016, Coimbra, Portugal

[DatBon16e] Describing things in the Internet of Things : From CoRE link format to semantic based descriptions Datta, Soumya Kanti; Bonnet, Christian ICCE-TW 2016, IEEE International Conference on Consumers Electronics-Taiwan, May 27-29, 2016, Nantou, Taiwan

[DatBon16f] Integrating named data networking in Internet of Things architectureDatta, Soumya Kanti; Bonnet, Christian ICCE-TW 2016, IEEE International Conference on Consumers Electronics-Taiwan, May 27-29, 2016, Nantou, Taiwan

[Dat16b] Fog computing for consumer centric Internet of ThingsDatta, Soumya KantiTutorial at TENSYMP 2016, IEEE Region 10 Symposium, May 9-11, 2016, Sanur, Bali Island, Indonesia

[DatBonHar16c] DataTweet: an architecture enabling data-centric IoT servicesDatta, Soumya Kanti; Bonnet, Christian; Ferreira Da Costa, Rui Pedro; Härri, JérômeTENSYMP 2016, IEEE Region 10 Symposium, May 9-11, 2016, Sanur, Bali Island, Indonesia

[DatBonHar16d] Web of things for connected vehiclesDatta, Soumya Kanti; Ferreira Da Costa, Rui Pedro; Bonnet, Christian; Härri, Jérôme WWW 2016, 25th International World Wide Web Conference, W3C Track, April 11-15, 2016, Montreal, Canada

[Dat16c] Cross-domain IoT application development for smart homeDatta, Soumya Kanti; Bonnet, ChristianPresentation at ETSI Workshop "Internet of Things in the Smart Home", March 21-22, 2016, Sophia-Antipolis, France

Dat16d] Horizontal IoT application development using semanticsDatta, Soumya KantiTutorial Presentation at T2TRG meeting, in colocation with the IoT Semantic Interoperability Workshop, March 17-18, 2016, San Jose, USA

[Dat16e] The web of things as a platform of platformsRaggett, Dave; Datta, Soumya Kanti IOTSI, IoT Semantic Interoperability Workshop 2016, 17-18 March 2016, San José, USA

[Dat:16a] Easing IoT application development through DataTweet frameworkDatta, Soumya Kanti; Bonnet, ChristianWF-IoT 2016, IEEE World Forum on Internet of Things, 12-14 December 2016, Reston, VA, USA

[Dat:16b] A categorization of discovery technologies for the Internet of ThingsBröring, Arne; Datta, Soumya Kanti; Bonnet, ChristianIOT 2016, 6th International Conference on the Internet of Things, November 7-9, 2016, Stuttgart, Germany

[Dat:16c] Securing DataTweet IoT architecture elementsDatta, Soumya Kanti; Bonnet, ChristianICCE-ASIA 2016, International Conference on Consumer Electronics, October 26-28, 2016, Seoul, Korea

[Dat:16d] Semantic web meets internet of things (IoT) and web of things (WoT)Gyrard, Amélie; Patel, Pankesh; Datta, Soumya Kanti; Ali, Muhammad IntizarISWC 2016, 15th International Semantic Web Conference, Tutorial, October 17-21 2016, Kobe, Japan

[Dat:16d] Interworking of NDN with IoT architecture elements: challenges and solutionsDatta, Soumya Kanti; Bonnet, ChristianGCCE 2016, IEEE 5th Global Conference on Consumer Electronics, October 11-14, 2016, Kyoto, Japan

[Dat:16e] Semantic interoperability for the Web of ThingsMurdock, Paul; Bassbouss, Louay; Datta, Soumya Kanti; et al.Research report, 2016

[DecGui16] Decurninge, Alexis; Guillaud, Maxime; Slock, Dirk TM, “Riemannian coding for covariance interpolation in massive MIMO frequency division duplex systems”, 9th IEEE Sensor Array and Multichannel Signal Processing (SAM) Workshop, 10-13 July 2016, Rio de Janeiro, Brazil.

[FavApo16a] Towards moving public safety networks Favraud, Romain; Apostolaras, Apostolos; Nikaein, Navid; Korakis, Thanasis IEEE Communications Magazine, Vol. 54, N°3, March 2016, ISSN: 0163-6804

[FavApo16b] Public safety networks: Enabling mobility for critical communications Favraud, Romain; Apostolaras, Apostolos; Nikaein, Navid; Korakis, Thanasis Book Chapter N°4, in "Wireless Public Safety Networks 2 : A systemic approach", Wiley-ISTE, ISBN: 9781785480522

[FilKerGes16] Miltiades Filippou, Paul de Kerret, David Gesbert, Tharmalingam Ratnarajah, Adriano Pastore, and Georges A. Ropokis, ``Spatial CSIT allocation policies for network MIMO channels", IEEE Transactions on Wireless Communications, vol.15, no.8, August 2016

[FilKer16] M. Filipou, P. de Kerret, D. Gesbert, T. Ratnarajah, A. Pastore, G. ropokis, "Coordinated Shared Spectrum Precoding with Distributed CSIT", IEEE Trans. on Wireless Communications, Aug. 2016.

[FilRop16] M. Filipou, G. Ropokis, D. Gesbert, T. Radnarajah "Joint Sensing and Reception Design of SIMO Hybrid Cognitive Radio Systems", IEEE Transactions on Wireless Communications, Sept. 2016. Available under arxiv at http://arxiv.org/abs/1509.04043

[Gal:16a] Analytic performance comparison of unsupervised LTE D2D and DSRC in a V2X safety context Gallo, Laurent; Härri, Jérôme
Research report RR-14-298, latest updated on September 2016.

[Gal:16b] Vehicular safety critical communications: a case study for unsupervised LTE D2D
Gallo, Laurent; Härri, Jérôme
Research Report RR-16-327

[GarWym16] Garcia, Nil; Wymeersch, Henk; Ström, Erik G; Slock, Dirk TM, “Location-aided mm-wave channel estimation for vehicular communication”, 17th Int’l IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), July 3-6, 2016, Edinburgh, UK.

[GyrDat16] Semantic web meets internet of things (IoT) and web of things (WoT)Gyrard, Amélie; Patel, Pankesh; Datta, Soumya Kanti; Ali, Muhammad IntizarISWC 2016, 15th International Semantic Web Conference, Tutorial, October 17-21 2016, Kobe, Japan

[Gyr16a] Reusing and unifying background knowledge for Internet of Things with LOV4IoTGyrard, Amélie; Atemezing, Ghislain; Bonnet, Christian; Boudaoud, Karima; Serrano, MartinFICLOUD 2016, IEEE 4th International Conference on Future Internet of Things and Cloud, 22-24 August 2016, Vienna, Austria

[Gyr16b] LOV4IoT: A second life for ontology-based domain knowledge to build Semantic Web of Things applicationsGyrard, Amélie; Bonnet, Christian; Boudaoud, Karima; Serrano, MartinFICLOUD 2016, IEEE 4th International Conference on Future Internet of Things and Cloud, 22-24 August 2016, Vienna, Austria

[GyrAte16] Reusing and unifying background knowledge for Internet of Things with LOV4IoTGyrard, Amélie; Atemezing, Ghislain; Bonnet, Christian; Boudaoud, Karima; Serrano, MartinFICLOUD 2016, IEEE 4th International Conference on Future Internet of Things and Cloud, 22-24 August 2016, Vienna, Austria

[GyrBon16] LOV4IoT: A second life for ontology-based domain knowledge to build Semantic Web of Things applicationsGyrard, Amélie; Bonnet, Christian; Boudaoud, Karima; Serrano, MartinFICLOUD 2016, IEEE 4th International Conference on Future Internet of Things and Cloud, 22-24 August 2016, Vienna, Austria

[Gas:16a] Mixed traffic flow with 2-wheelers: Macroscopic modeling and analysisGashaw, Sosina Mengistu; Goatin, Paola; Härri, JérômeBMW-EURECOM-TUM Summer School 2016, July 18-23, 2016, Tegernsee, Bavaria, Germany

[GopSlo16a] K. Gopala, D. Slock, "Doppler compensation and Receive Beamforming forHigh Mobility OFDM transmissions in multipath", EAI Int'l Conf. on Cognitive Radio Oriented Wireless Networks (CROWNCOM), Grenoble, France, 2016.

[GopSlo16b] K. Gopala, D. Slock, "MIMO OFDM capacity maximizing beamforming for large Doppler scenarios", IEEE Int'l Workshop on Signal Processing Advances in Wireless Communications(SPAWC), Edinburgh, UK, 2016.

[GopSlo16c] K. Gopala, D. Slock, "High doppler MIMO OFDM capacity maximizing spatial transceivers exploiting excess cyclic prefix", Spectrum Aggregation and Sharing for 5G Networks (SAS5G) Workshop at IEEE ISWCS, Poznan, Poland, 2016.

[GopSlo17] K. Gopala, D. Slock, "Robust MIMO OFDM transmit beamformer design for large doppler scenarios under partial CSIT", Proc. International Conf. on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, USA, 2017.

[Hoa:16c] Breaking the gridlock of spatial correlations in GPS-aided IEEE 802.11p-based cooperative positioning Hoang, Gia-Minh; Denis, Benoit; Härri, Jérôme; Slock, Dirk TM
IEEE Transaction on Vehicular Technology, December 2016, Vol.65, N°12

[Hoa:16c] Cooperative localization in GNSS-aided VANETs with accurate IR-UWB range measurements Hoang, Gia Minh; Denis, Benoit ; Härri, Jérôme; Slock, Dirk TMWPNC 2016, 13th IEEE Workshop on Positioning, Navigation and Communications, 19-20 October 2016, Bremen, Germany

[Hoa:16a] On communication aspects of particle-based cooperative positioning in GPS-aided VANETs Hoang, Gia-Minh; Denis, Benoit; Härri, Jérôme; Slock, Dirk TMIV 2016, IEEE Workshop on Cooperative Communication and Positioning, in conjunction with 2016 IEEE Intelligent Vehicles Symposium, June 19-22, 2016, Gothenburg, Sweden

[Här:16b] Cooperative connected vehicles (C2V): where IoT meets C-ITSHärri, Jérôme Invited Talk at Leibniz Universität Hannover, February 3rd, 2016, Hannover, Germany

[HolAij06] Holland, Oliver; Aijaz, Adnan; Kaltenberger, Florian; Foukalas, Fotis; Vivier, Guillaume; Buczkowski, Mateusz; Pietrzyk, Slawomir, "Management architecture for aggregation of heterogeneous systems and spectrum bands," IEEE Communications Magazine, Vol.54, N°9, September 2016

[Här:16a] Educational and professional development in ITSHärri, JérômeITS 2016, 11th ITS European Congress, Panel Talk, June 6-9, 2016, Glasgow, Scotland

[Jeo:16] Survey on IP-based Vehicular Networking for Intelligent Transportation Systems
Jeong, Pau; Cespedes, Sandra; Benamar, Nabil; Härri, Jérôme IETF IPWAVE draft, draft-jeong-ipwave-vehicular-networking-survey-00, 2016.

[Jun:16] Bayesian network based speed estimation in case of C2X data at very low equipment rates Junghans, Marek; Leich, Andreas; Härri, JérômeTRISTAN 2016, 9th Triennial Symposium on Transportation Analysis, 13-17 June 2016, Oranjestad, Aruba.

[JiaKalExp16] X. Jiang, F. Kaltenberger, R. Knopp, D. Slock, L. Deneire, "Exploiting channel reciprocity in massive MIMO", EUCNC 2016, European Conference on Networks and Communications, June 27-30

[JiaKalHow16] X. Jiang, F. Kaltenberger, “How accurately should we calibrate a massive MIMO TDD system?” SNOW 2016, 7th Nordic Workshop on System and Network Optimization for Wireless, 17-19 January, 2016, Cortina Dampezzo, Italy

[JiaKalHow16bis] X. Jiang, F. Kaltenberger, L. Deneire, “How accurately should we calibrate a massive MIMO TDD system?” ICC 2016, International Conference on Communications, Workshop, May 23-27, 2016, Kuala Lumpur, Malaysia

[JiaKalOpen16] X. Jiang, F. Kaltenberger, R. Knopp, H. Maatallahc, “OpenAirInterface Massive MIMO Testbed: A 5G Innovation Platform”, white paper, & Video, www.openairinterface.org?page_id=1760 , Aug 2016.

[KadAvr16] Kadavankandy, Arun; Avrachenkov, Konstantin; Cottatellucci, Laura; Rajesh, Sundaresan, Subgraph detection with cues using belief propagation,
Research Report, October 2016 submitted also to ICC 2017.

[KadCot16] Kadavankandy, Arun; Cottatellucci, Laura; Avrachenkov, Konstantin Characterization of L1-norm statistic for anomaly detection in Erdos Renyi graphsResearch Report RR-16-315, March 16, 2016

[KadCot16a] Kadavankandy, Arun; Cottatellucci, Laura; Avrachenkov, Konstantin, Characterization of L1-norm statistic for anomaly detection in Erdos Renyi graphsCDC 2016, 55th IEEE Conference on Decision and Control, December 12-14, 2016, Las Vegas, USA

[Kal06] Kaltenberger, Florian, “Carrier aggregation implementation on Open Air Interface platform,” WinnComm Europe 2016, October 11th, 2016, Paris, France

[KalRou06] Kaltenberger, Florian; Roux, Cédric; Buczkowski, Mateusz; Wewior, Macie, “The OpenAirInterface application programming interface for schedulers using carrier aggregation,” ISWCS 2016, 13th International Symposium on Wireless Communication Systems: SAS5G: Workshop on Spectrum Aggregation and Sharing for 5G Networks (ISWCS'2016-SAS5G Workshop), September 20-23, 2016, Poznan, Poland

[KerBazGes16] Paul de Kerret, Antonio Bazco, and David Gesbert, ``Enforcing coordination in network MIMO with unequal CSIT", Proc. Asilomar Conference 2016.

[KerGesZha16] Paul de Kerret, David Gesbert, Jingjing Zhang, and Petros Elia, "`Optimally bridging the gap from delayed to perfect CSIT in the K-user MISO BC"', Proc. ITW 2016.

[KerGes16] Paul de Kerret, and David Gesbert, ``Network MIMO: Transmitters with no CSI can still be very useful", Proc. ISIT 2016.

[Kha:16] Can ITS-G5 & Wi-Fi Coexist?Mohammed Khan, Irfan; Jérôme HärriBMW-EURECOM-TUM Summer School 2016, July 18-23, 2016, Tegernsee, Bavaria,

[KK16a], 5G Architectural Design Patterns, K. Katsalis, N Nikaein, E Schiller, R Favraud, TI Braun ,Communications Workshops (ICC), 2016 IEEE International Conference, 32-37,2016

[KK16b] SLA-driven VM Scheduling in Mobile Edge Computing, K. Katsalis, TG Papaioannou, N Nikaein, L Tassiulas, IEEE CLOUD, 2016

[KK16c] Building Virtual 802.11 Testbeds Towards Open 5G Experimentation, K. Kousias, K Katsalis, D Stavropoulos, T Korakis, L Tassiulas, IEEE, WCNC, 2016

[KK16d] Multi-Domain Orchestration for NFV: Challenges and Research Directions, K. Katsalis, N. Nikaein, A. Edmonds, IUCC, Granada Spain, 2016

[KK16e] Q4Health: Mission critical communications over LTE and future 5G technologies, Garcia-Perez, Cesar A; Rios, A.; Merino, P.; Katsalis, K.; Nikaein, N.; Figueiredo, R.; Morris, Donal; O? Callaghan, T.; Rodriguez, P., ICT Fire Book, November 2016

[KK16f] Q4HEALTH: Quality of Service and prioritisation for emergency services in the LTE RAN stack, C. A. García-Pérez et al 2016 European Conference on Networks and Communications (EuCNC), Athens, 2016, pp. 64-68.

[LukKal16] E. Lukashova, F. Kaltenberger, R. Knopp, C. Bonnet "PHY Layer Abstraction for SU-MIMO LTE System Employing Parallel Interference-Aware Receiver", IWSLS 2016, 1st International Workshop on Link- and System Level Simulations, Vienna, Austria, 2016.

[LukKal16a] E. Lukashova, F. Kaltenberger, R. Knopp, "Implementation and Analysis of Interference—Aware ML SIC receiver for LTE SU-MIMO", submitted to IEEE ICC 2017 Wireless Communications Symposium, Paris, France, May 2017.

[LukKal17] E. Lukashova, F. Kaltenberger, R. Knopp, "Single-User MIMO ML Successive Interference Canceling Receiver withHARQ-IR Protocol", accepted at WSA 2017 Smart Antenna Workshop, Berlin, Germany, March 2017.

[Mah:16] Mobility-aware edge caching for connected carsMahmood, A; Casetti, C; Chiasserini, C.F; Giaccone, P; Härri, JérômeWONS 2016, 12th IEEE Conference on Wireless On-Demand Networks systems and Services, 20-22 January 2016, Cortina d'Ampezzo, Italy

[MehSpy16a] F. Mehmeti, T. Spyropoulos, “Performance Analysis of Mobile Data Offloading in Heterogeneous Networks,” to appear in IEEE Trans. Mobile Computing

[MehSpy16b] F. Mehmeti, T. Spyropoulos, “Performance modeling, analysis and optimization of delayed mobile data offloading under different service disciplines,” to appear in ACM/IEEE Transactions on Networking.

[NN16] Towards RAN slicing in 5G, Nikaein, Navid, FG IMT-2020 Workshop and Demo Day: Wireline Technology Enablers for 5G, 7 December 2016, Geneva, Switzerland

[NN16b] Towards a cloud-native radio access network, Nikaein, Navid; Schiller, Eryk; Favraud, Romain; Knopp, Raymond; Alyafawi, Islam; Braun, Torsten, Book Chapter of "Advances in Mobile Cloud Computing and Big Data in the 5G Era", Springer, 2016, ISBN: 978-3-319-45143-5

[NN16c] Virtual network functions orchestration in heterogeneous 5G networks, Nikaein, Navid, EUCNC 2016, European Conference on Networks and Communications, Demo, June 27-30, Athens, Greece.

[NN16d] OAI in the cloud: NFV modelling, composition and chaining in cloudified 5G systems, Nikaein, Nikaein; Gupta, Rohit; Knopp, Raymond; Bonnet, Christian; Gauthier, Lionel; Kaltenberger, Florian; Roux, Cédric OPNFV Summit 2016, Open Platform for NFV Summit, Demo, June 20-23, 2016, Berlin, Germany.

[NN16e] NFV modelling, composition and chaining in cloudified 5G systems, Nikaein, NavidETSI Workshop 2016, From Research to Standardization, Demo, May 10-11, 2016, Sophia-Antipolis, France.

[Ngu:16a] SDN-based distributed mobility management for 5G networksNguyen, Tien-Thinh; Bonnet, Christian; Härri, JérômeWCNC 2016, IEEE Wireless Communications and Networking Conference, Track 3 - Mobile and Wireless Networks, 3-6 April 2016, Doha, Qatar

[Pauletal16] Semantic interoperability for the Web of ThingsMurdock, Paul; Bassbouss, Louay; Datta, Soumya Kanti; et al.Research report, 2016

[Pat:16a] Coordinated maneuvering supporting mixed conventional and autonomous vehiclesPatel, Raj Haresh; Jérôme HärriBMW-EURECOM-TUM Summer School 2016, July 18-23, 2016, Tegernsee, Bavaria,

[Pat:16b] Coordinated braking strategies supporting mixed autonomous and conventional vehicles Patel, Raj Haresh; Jérôme HärriResearch Report RR-16-321, 22 August 2016

[Pet:16] Transmission of IPv6 Packets over IEEE 802.11 Networks in mode Outside the Context of a Basic Service Set, Petrescu, Alex; Benamar, Nabil; Härri, Jérôme; Huitema, Christian; Lee, Jong-Hyouk; Ernst, Thierry; Li, Tony; IETF IPWAVE draft, draft-ietf-ipwave-ipv6-over-80211ocb-00, 2016.

[SerSpy16a] P. Sermpezis, T. Spyropoulos, "Delay analysis of epidemic schemes in sparse and dense heterogeneous contact environments", to appear in IEEE Transactions on Mobile Computing.

[SerSpy16b] P. Sermpezis, T. Spyropoulos, “Effects of Content Popularity on the Performance of Content-Centric Opportunistic Networking: An Analytical Approach and Applications. IEEE/ACM Trans. Netw. 24(6): 3354-3368 (2016)”

[SapSpy16] N. Sapountzis, T. Spyropoulos, N. Nikaein, U. Salim, ”Optimal downlink and uplink user association in backhaul-limited HetNets,” in Proc of IEEE INFOCOM 2016.

[TabSlo16a] Tabikh, Wassim; Slock, Dirk TM; Yuan-Wu, Yi; “Weighted sum rate maximization of correlated MISO interference broadcast channels under linear precoding: A large system analysis”, VTC 2016-Spring, IEEE 83rd Vehicular Technology Conference, 15-18 May 2016, Nanjing, China

[TabSlo16b] Tabikh, Wassim; Slock, Dirk TM; Yuan-Wu, Yi; “Weighted sum rate maximization of MISO interference broadcast channels via difference of convex functions programming: A large system analysis”, SSP 2016, IEEE Workshop on Statistical Signal Processing, 26-29 June 2016, Palma de Mallorca, Spain

[TabYua16] Tabikh, Wassim; Yuan-Wu, Yi; Slock, Dirk TM; ” Decentralizing multi-cell maximum weighted sum rate precoding via large system analysis”, EUSIPCO 2016, 24th European Signal Processing Conference, 28 August-2 September 2016, Budapest, Hungary

[TabSlo16c] Tabikh, Wassim; Yuan-Wu, Yi; Slock, Dirk TM;” A large system analysis of weighted sum rate maximization of single stream MIMO interference broadcast channels under linear precoding”, ISWCS 2016, International Symposium on Wireless Communication Systems, 20-23 September 2016, Poznan, Poland.

[TabSlo16d] Tabikh, Wassim; Slock , Dirk TM; Yuan-Wu, Yi, “Optimal beamforming with combined channel and path CSIT for multi-cell multi-user MIMO”, Information Theory and Applications (ITA) Workshop, Jan. 31-Feb. 5, 2016, San Diego, USA.

[TolKal16] A. Tolli, J. Kaleva, G. Venkatraman, D. Gesbert, "Joint UL/DL Mode Selection and Transceiver Design for Dynamic TDD Systems", in proc. of IEEE Symposium on Transceivers and Signal Processing for 5G Wireless and mm-Wave Systems, GLOBALSIP16, 2016.

[TCNN16] Towards software defined 5G radio access networks, Chen, Tao; Nikaein, Navid
Newsletter, Software Defined Networks, March 2016.

[VigSpy16] L. Vigneri, T. Spyropoulos, C. Barakat, “Storage on wheels: Offloading popular contents through a vehicular cloud,” in Proc. of IEEE WoWMoM 2016

[WanFan16] Wang, Lusheng; Fang, Fei; Nikaein, Navid; Cottatellucci, Laura, An analytical framework for multilayer partial frequency reuse scheme design in mobile communication systems, IEEE Transactions on Vehicular Technology, Vol. 65, N°9, September 2016, ISSN: 0018-9545

[XFNN16] FlexRAN: A flexible and programmable platform for software-defined radio access networks, Foukas, Xenofon; Nikaein, Navid; Kassem, Mohamed M; Kontovasilis, Kimon CONEXT 2016, 12th International on Conference on emerging Networking EXperiments and Technologies, Irvine, California, USA.

[XFNN16a] FlexRAN: A flexible and programmable platform for software-defined radio access networks, Foukas, Xenofon; Nikaein, Navid; Kassem, Mohamed M.; Marina, Mahesh K.; Kontovasilis, Kimon, ATC 2016, All Things Cellular Workshop, in conjunction with ACM MobiCom 2016, 7 October 2016, New York, USA.

[YinCot16] Yin, Haifan; Cottatellucci, Laura; Gesbert, David; Müller, Ralf R.; He, Gaoning Robust pilot decontamination: a joint angle and power domain approachICASSP 2016, 41st IEEE International Conference on Acoustics, Speech and Signal Processing, March 20-25, 2016, Shanghai, China

[YinCot16a] Yin, Haifan; Cottatellucci, Laura; Gesbert, David, Müller, Ralf, R; He, Gaoning, Robust pilot decontamination based on joint angle and power domain discriminationIEEE Transactions on Signal Processing, 2016, Vol.64, N°11, ISSN: 1053-587X

[ZhaEli16] Jingjing Zhang and Petros Elia, “Fundamental limits of cache-aided wireless BC: Interplay of coded-caching and CSIT feedback,” ALLERTON 2016, 54th Annual Allerton Conference on Communication, Control, and Computing, September 27-30, 2016, Monticello, IL.

Research activity report 2016

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