RESEARCH CENTRE

Sophia Antipolis - Méditerranée

2020 ACTIVITY REPORT

Project-Team DIANA

Design, Implementation and Analysis of Networking Architectures

DOMAIN

Networks, Systems and Services, Distributed Computing

THEME

Networks and Telecommunications

Contents

Project-Team DIANA

Creation of the Team: 2013 January 01, updated into Project-Team: 2015 July 01

Keywords

Computer sciences and digital sciences

A1.1.13. – Virtualization

A1.2.1. – Dynamic reconfiguration

A1.2.2. – Supervision

A1.2.3. – Routing

A1.2.4. – QoS, performance evaluation

A1.2.5. – Internet of things

A1.2.9. – Social Networks

A1.3. – Distributed Systems

A1.3.4. – Peer to peer

A1.4. – Ubiquitous Systems

Other research topics and application domains

B6.2. – Network technologies

B6.2.1. – Wired technologies

B6.2.2. – Radio technology

B6.2.3. – Satellite technology

B6.3.2. – Network protocols

B6.3.3. – Network Management

B6.3.4. – Social Networks

B8.5.2. – Crowd sourcing

B9.1.1. – E-learning, MOOC

B9.5.1. – Computer science

B9.5.6. – Data science

B9.8. – Reproducibility

B9.10. – Privacy

1 Team members, visitors, external collaborators

Research Scientists

- Walid Dabbous [Team leader, Inria, Senior Researcher, HDR]
- Chadi Barakat [Inria, Senior Researcher, HDR]
- Arnaud Legout [Inria, Researcher, HDR]
- Damien Saucez [Inria, Researcher, from Oct 2020]
- Thierry Turletti [Inria, Senior Researcher, HDR]

Post-Doctoral Fellow

- Tingting Yuan [Inria, until Feb 2020]

PhD Students

- Othmane Belmoukadam [Univ Côte d'Azur]
- Yanis Boussad [Univ Côte d'Azur]
- Giuseppe Di Lena [Orange Labs, CIFRE]
- Mamoutou Diarra [Ekinops, CIFRE]
- Houssam Elbouanani [Inria]

Technical Staff

- David Migliacci [Inria, Engineer, until Apr 2020]
- Thierry Parmentelat [Inria, Engineer]

Interns and Apprentices

- Rabi Ahrara [Univ Côte d'Azur, until Jan 2020]
- Chanpiseth Chap [Univ Côte d'Azur, until Jan 2020]
- Chanpiseth Chap [Inria, from Mar 2020 until Aug 2020]
- Anass El Boujidi [Inria, from Oct 2020]
- Florinda Fragassi [Univ Côte d'Azur, until Jan 2020]
- Rossella Franco [Univ Côte d'Azur, until Jan 2020]
- Amal Krimi [Univ Côte d'Azur, from Oct 2020]
- Angelo Rodio [Univ Côte d'Azur, until Jan 2020]
- Angelo Rodio [Inria, from Mar 2020 until Aug 2020]
- Bernard Tamba Sandouno [Inria, from Mar 2020 until Aug 2020]
- Adeel Siddiqui [Univ Côte d'Azur, until Jan 2020]

Administrative Assistant

- Christine Foggia [Inria]

Visiting Scientist

- Wilson Borba Da Rocha Neto [Univ Campinas, from Feb 2020 until Apr 2020]

External Collaborators

- Mondi Ravi [Inria, until May 2020]
- Mondi Ravi [Self employed, from Jul 2020]

2 Overall objectives

2.1 Presentation of the team

The overall objective of the DIANA project-team is to provide network architectural support for improving citizen rights in the Internet. To do so, we work to provide service transparency and user data control in the context of hundreds of billions of both wired and mobile devices. Our methodology includes advanced measurement techniques, design and implementation of architectural solutions, and their validation in adequate experimental facilities.

The high complexity of the Internet architecture, protocols and services, and the economic interests of the big stakeholders result in a lack of transparency concerning information of high interest to the connected "citizen" such as possible privacy leaks, root cause of service degradation or lock-in behavior. It is therefore important to enhance the network to provide service transparency to citizens.

On the other hand, the ossification of the Internet architecture around the IP protocol makes introduction of new functionalities in the network quite difficult. Users currently have no control on their contents and depend on big companies (e.g., Google drive, iCloud, dropbox, Microsoft OneDrive) to easily access and share data at the expense of their privacy. However, the recent development of software-defined network and network functions virtualization concepts open the perspective of faster deployment of network functionalities, as it abstracts the whole network as a single piece of software, instead of a large number of heterogeneous and dedicated devices to be configured one-by-one.

In the DIANA project-team, we have two main research directions:

- designing and deploying a measurement plane providing network service transparency,
- defining and deploying an open network architecture for user control.

Our research program is presented briefly in the next section.

3 Research program

3.1 Service Transparency

Transparency is to provide network users and application developers with reliable information about the current or predicted quality of their communication services, and about potential leakages of personal information, or of other information related to societal interests of the user as a "connected citizen" (e.g. possible violation of network neutrality, opinion manipulation). Service transparency therefore means to provide information meaningful to users and application developers, such as quality of experience, privacy leakages, or opinion manipulation, etc. rather than network-level metrics such as available bandwidth, loss rate, delay or jitter.

The Internet is built around a best effort routing service that does not provide any guarantee to end users in terms of quality of service (QoS). The simplicity of the Internet routing service is at the root of its huge success. Unfortunately, a simple service means unpredicted quality at the access. Even though a considerable effort is done by operators and content providers to optimise the Internet content delivery chain, mainly by over-provisioning and sophisticated engineering techniques, service degradation is still part of the Internet. The proliferation of wireless and mobile access technologies, and the versatile nature of Internet traffic, make end users quality of experience (QoE) forecast even harder. As a matter of

fact, the Internet is missing a dedicated measurement plane that informs the end users on the quality they obtain and in case of substantial service degradation, on the origin of this degradation. Current state of the art activities are devoted to building a distributed measurement infrastructure to perform active, passive and hybrid measurements in the wired Internet. However, the problem is exacerbated with modern terminals such as smartphones or tablets that do not facilitate the task for end users (they even make it harder) as they focus on simplifying the interface and limiting the control on the network, whereas the Internet behind is still the same in terms of the quality it provides. Interestingly, this same observation explains the existing difficulty to detect and prevent privacy leaks. We argue that the lack of transparency for diagnosing QoE and for detecting privacy leaks have the same root causes and can be solved using common primitives. For instance, in both cases, it is important to be able to link data packets to an application. Indeed, as the network can only access data packets, there must be a way to bind these packets to an application (to understand users QoE for this application or to associate a privacy leak to an application). This is however a complex task as the traffic might be obfuscated or encrypted. Our objectives in the research direction are the following:

- Design and develop measurement tools providing transparency, in spite of current complexity
- Deploy those measurement tools at the Internet's edge and make them useful for end users
- Propose measurements plane as an overlay or by exploiting in-network functionalities
- Adapt measurements techniques to network architectural change
- Provide measurements as native functionality in future network architecture

3.2 Open network architecture

We are surrounded by personal content of all types: photos, videos, documents, etc. The volume of such content is increasing at a fast rate, and at the same time, the spread of such content among all our connected devices (mobiles, storage devices, set-top boxes, etc) is also increasing. All this complicates the control of personal content by the user both in terms of access and sharing with other users. The access of the personal content in a seamless way independently of its location is a key challenge for the future of networks. Proprietary solutions exist, but apart from fully depending on one of them, there is no standard plane in the Internet for a seamless access to personal content. Therefore, providing network architectural support to design and develop content access and sharing mechanisms is crucial to allow users control their own data over heterogeneous underlying network or cloud services.

On the other hand, privacy is a growing concern for states, administrations, and companies. Indeed, for instance the French CNIL (entity in charge of citizens privacy in computer systems) puts privacy at the core of its activities by defining rules on any stored and collected private data. Also, companies start to use privacy preserving solutions as a competitive advantage. Therefore, understanding privacy leaks and preventing them is a problem that can already find support. However, all end-users do not *currently* put privacy as their first concern. Indeed, in face of two services with one of higher quality, they usually prefer the highest quality one whatever the privacy implication. This was, for instance, the case concerning the Web search service of Google that is more accurate but less privacy preserving than Bing or Qwant. This is also the case for cloud services such as iCloud or Dropbox that are much more convenient than open source solutions, but very bad in terms of privacy. Therefore, to reach end-users, any privacy preserving solutions must offer a service equivalent to the best existing services.

We consider that it will be highly desirable for Internet users to be able to *easily* move their content from a provider to another and therefore not to depend on a content provider or a social network monopoly. This requires that the network provides built-in architectural support for content networking.

In this research direction, we will define a new *service abstraction layer* (SAL) that could become the new waist of the network architecture with network functionalities below (IP, SDN, cloud) and applications on top. SAL will define different services that are of use to all Internet users for accessing and sharing data (seamless content localisation and retrieval, privacy leakage protection, transparent vertical and horizontal handover, etc.). The biggest challenge here is to cope in the same time with large number of content applications requirements and high underlying networks heterogeneity while still providing

efficient applications performance. This requires careful definition of the services primitives and the parameters to be exchanged through the service abstraction layer.

Two concurring factors make the concept behind SAL feasible and relevant today. First, the notion of scalable network virtualization that is a required feature to deploy SAL in real networks today has been discussed recently only. Second, the need for new services abstraction is recent. Indeed, more than fifteen years ago the Internet for the end-users was mostly the Web. Only ten years ago smartphones came into the picture of the Internet boosting the number of applications with new functionalities and risks. Since a few years, many discussions in the network communities took place around the actual complexity of the Internet and the difficulty to develop applications. Many different approaches have been discussed (such as CCN, SDN) that intend to solve only part of the complexity. SAL takes a broader architectural look at the problem and considers solutions such as CCN as mere use cases. Our objectives in this research direction include the following:

- Identify common key networking services required for content access and sharing
- Detect and prevent privacy leaks for content communication
- Enhance software defined networks for large scale heterogeneous environments
- Design and develop open Content Networking architecture
- Define a service abstraction layer as the thin waist for the future content network architecture
- Test and deploy different applications using SAL primitives on heterogeneous network technologies

3.3 Methodology

We follow an experimental approach that can be described in the following techniques:

- Measurements: the aim is to get a better view of a problem in quantifiable terms. Depending on the field of interest, this may involve large scale distributed systems crawling tools; active probing techniques to infer the status and properties of a complex and non controllable system as the Internet; or even crowdsourcing-based deployments for gathering data on real-users environments or behaviours.
- Experimental evaluation: once a new idea has been designed and implemented, it is of course very desirable to assess and quantify how effective it can be, before being able to deploy it on any realistic scale. This is why a wide range of techniques can be considered for getting early, yet as significant as possible, feedback on a given paradigm or implementation. The spectrum for such techniques span from simulations to real deployments in protected and/or controlled environments.

4 Application domains

The DIANA project-team conducts research activities to provide network architectural support for improving citizen rights in the Internet. The main application domains of the teams are:

- Network and quality of experience measurement
- Detection of private information leaks
- Industrial deterministic networks
- Data center networks
- Deployment of future open radio networks
- Realistic simulations and reproducible experiments

5 Social and environmental responsibility

Public health politics and scientist evaluating the impact of EMF radiations on human beings all face the same challenge: How to assess the real exposure of human beings in order to correlate it with observed symptoms and illness. This problem is even harder considering that the period of observation must be long (years), and the number of observed persons must be large.

ElectroSmart is a technological breakthrough that will have scientific, technical, and societal applications, notably on public health politics, by providing the scientific community and potential users with a unique measuring instrument, methods, and models to exploit the invaluable data gathered by the instrument.

6 Highlights of the year

6.1 Awards

The paper "From Encrypted Video Traces to Viewport Classification", by Othmane Belmokaddam and Chadi Barakat had the best paper award at the 16th International Conference on Network and Service Management (CNSM'2020), Virtual Conference, November 2020 [16].

6.2 Large deployment

The ElectroSmart has witnessed a very large deployment in 2020. As of December 16, 2020, the application has been downloaded 760,000 times and has 200,000 active users. The ElectroSmart application made it possible to collect 8 billion measurements in 150 countries. This represents, to our knowledge, the largest data base of human exposure to electromagnetic waves produced by wireless technologies.

7 New software and platforms

7.1 New software

7.1.1 ACQUAmobile

Name: Application for prediCting Quality of User Experience at Internet Access

Keywords: Android, Internet access, Performance measure, Quality of Experience

Scientific Description: ACQUA is an Application for prediCting QUality of Experience (QoE) at Internet Access. It is developed by the Diana team at Inria Sophia Antipolis – Méditerranée and was supported by Inria under the ADT ACQUA grant. The scientific project around ACQUA is supported by Inria Project Lab BetterNet and the French National Project ANR BottleNet. The project also got the approval of Inria COERLE and French CNIL for the part on experimentation with real users. ACQUA presents a new way for the evaluation of the performance of Internet access. Starting from network-level measurements as the ones we often do today (bandwidth, delay, loss rates, jitter, etc), ACQUA targets the estimated Quality of Experience (QoE) related to the different applications of interest to the user without the need to run them (e.g., estimated Skype quality, estimated video streaming quality).

The ACQUA Android application is supposed to be on one hand the reference application for QoE forecasting and troubleshooting for end users at their Internet access, and on the other hand, the feedback channel that allows end users to report to us (if they are willing) on their experience together with the corresponding network measurements so as to help us calibrating better and more realistic models. For this calibration, we are currently performing extensive, efficient and automatic measurements in the laboratory, we will count on end users to help us completing this dataset with further applications and more realistic network and user conditions.

ACQUA is mainly meant for end users, but it is also of interest to (mobile) network operators and to content providers to estimate the QoE of their customers and their networks without each time having to run expensive application-level traffic and to involve real users.

Functional Description: An application in ACQUA is a function, or a model, that links the network-level and device-level measurements to the expected Quality of Experience. Supervised machine learning techniques are used to establish such link between measurements both at the network level and the device level, and estimations of the Quality of Experience for different Internet applications. The required data for such learning can be obtained either by controlled experiments as we did in a prior work on Skype and YouTube Quality of Experience, or by soliciting the crowd (i.e. crowdsourcing) for combinations (i.e. tuples) of network- and application-level measurements and corresponding user-level Quality of Experience. Our current work is concentrating on using the ACQUA principle in the estimation and prediction of the Quality of Experience for main user's applications. We refer to the web site of the project for further details.

Assessment: Audience = 3, Software Originality = 4, Software Maturity = 3, Evolution and Maintenance = 3, Software Distribution and Licensing = 5.

URL: http://project.inria.fr/acqua/

Authors: Thierry Spetebroot, Chadi Barakat

Contact: Chadi Barakat

7.1.2 ElectroSmart

Keywords: Crowd-sourcing, UMTS, GSM, Bluetooth, Wi-Fi, 4G, 3G, 2G, Electromagnetic waves, Android, LTE

Functional Description: The Internet and new devices such as smartphones have changed fundamentally the way people communicate, but this technological revolution comes at the price of a higher exposition of the general population to microwave electromagnetic fields (EMF). This exposition is a concern for health agencies and epidemiologists who want to understand the impact of such an exposition on health, for the general public who wants a higher transparency on its exposition and the health hazard it might represent, but also for cellular operators and regulation authorities who want to improve the cellular coverage while limiting the exposition, and for computer scientists who want to better understand the network connectivity in order to optimize communication protocols. Despite the fundamental importance to understand the exposition of the general public to EMF, it is poorly understood because of the formidable difficulty to measure, model, and analyze this exposition.

The goal of the ElectroSmart project is to develop the instrument, methods, and models to compute the exposition of the general public to microwave electromagnetic fields used by wireless protocols and infrastructures such as Wi-Fi, Bluetooth, or cellular. Using a pluri-disciplinary approach combining crowd-based measurements, in-lab experiments, and modeling using sparse and noisy data, we address challenges such as designing and implementing a measuring instrument leveraging on crowd-based measurements from mobile devices such as smartphones, modeling the exposition of the general public to EMF to compute the most accurate estimation of the exposition, and analyzing the evolution of the exposition to EMF with time. This technological breakthrough will have scientific, technical, and societal applications, notably on public health politics, by providing the scientific community and potential users with a unique measuring instrument, methods, and models to exploit the invaluable data gathered by the instrument.

This project is supported by the UCN@Sophia Labex in 2016/2017/2018 (funding the engineer Mondi Ravi), by an Inria ADT (funding the engineer Abdelhakim Akodadi) 2017/2018, by and Inria ATT (funding the business developer David Migliacci) in 2017/2018, and by the academy 1 of UCAJedi (funding a Ph.D. student Yanis Boussad) 2017/2020.

In August 2016, we released the first stable public release of ElectroSmart. On the 16th December 2020 the app has been downloaded 760 000 time, we have 200 000 active users and a score of 4,5/5 on Google Play. We collected 8 billions measurements in 150 countries.

Assessment: A-5, SO-4, SM-4, EM-3-up4, SDL-1

We are in a process of creating a startup to commercialize the exposition maps we can build with the data we are collecting.

URL: https://electrosmart.app

Authors: Arnaud Legout, Hackob Melconian, Mondi Ravi, Abdelhakim Akodadi

Contact: Arnaud Legout

Participants: Arnaud Legout, Mondi Ravi

7.1.3 nepi-ng

Keywords: Wireless network, Experimentation

Functional Description: In the specific context of R2lab, we have created a tool suite for orchestrating network experiments, that for historical reasons we refer to collectively as nepi-ng, for NEPI new generation. An umbrella website is available at https://nepi-ng.inria.fr/.

At this point, nepi-ng has a much smaller scope than its NEPI ancestor used to have, in that it only supports remote control of network experiments over ssh. As a matter of fact, in practice, this is the only access mechanism that we need to have for running experiments on both R2lab, and PlanetLab Europe.

The design of nepi-ng of course is modular, so that it will be perfectly possible to add other control mechanisms to this core if and when it becomes necessary.

- asynciojobs:
  - URL: http://asynciojobs.readthedocs.io/en/latest/
  - Version: asynciojobs v0.5.4
  - Keywords: networking experimentation, orchestration
  - License: CC BY-SA 4.0
  - Type of human computer interaction: python library
  - OS/Middleware: Linux
  - Required library or software: python-3.5 / asyncio
  - Programming language: python3
- apssh:
  - URL: http://apssh.readthedocs.io/en/latest/
  - Version: apssh v0.7.1
  - Keywords: networking experimentation, orchestration
  - License: CC BY-SA 4.0
  - Type of human computer interaction: python library
  - OS/Middleware: Linux
  - Required library or software: python-3.5 / asyncio
  - Programming language: python3

URL: http://nepi-ng.inria.fr

Contacts: Thierry Parmentelat, Thierry Turletti

7.1.4 Distrinet

Name: Distrinet

Keywords: SDN, Emulation, Large-scale Emulators, Network simulator

Scientific Description: Networks have become complex systems that combine various concepts, techniques, and technologies. As a consequence, modelling or simulating them now is extremely complicated and researchers massively resort to prototyping techniques. Mininet is the most popular tool when it comes to evaluate SDN propositions. Mininet allows to emulate SDN networks on a single computer but shows its limitations with resource intensive experiments as the emulating host may become overloaded. To tackle this issue, we propose Distrinet, a distributed implementation of Mininet over multiple hosts, based on LXD/LXC, Ansible, and VXLAN tunnels. Distrinet uses the same API than Mininet, meaning that it is compatible with Mininet programs. It is generic and can deploy experiments on Linux clusters (e.g., Grid'5000), as well as on the Amazon EC2 cloud platform.

Assessment: A5, SO3, SM2, EM2-down, SDL4

Functional Description: Distrinet is an extension of Mininet that relies on LXC to be distributed in the cloud, and particularly in Amazon. The extension has been designed to be fully compatible with Mininet. As using Distrinet potentially involves the collaboration of multiple machines we focused on guaranteeing the correctness (in a sense that results are trustworthy) of simulations when running on multiple machines. To speedup deployments, loading and unloading operations have been parallelised with asynchronous calls. The pool of machines used for simulations is automatically provisioned thanks to Ansible.

Release Contributions: First release

URL: https://distrinet-emu.github.io

Publication: hal-03000617v1

Authors: Giuseppe Di Lena, Damien Saucez, Andrea Tomassilli

Contacts: Damien Saucez, Thierry Turletti, Walid Dabbous, Frédéric Giroire

Participants: Damien Saucez, Giuseppe Di Lena, Andrea Tomassilli, Frédéric Giroire, Thierry Turletti

Partner: Orange Labs

7.2 New platforms

7.2.1 Reproducible research Lab - R2lab

Scientific work around network protocols and related software stacks requires experiments, hence experimental conditions, to be reproducible. This is a particularly challenging requirement in the wireless networking area, where characteristics of wireless channels are known to be variable, unpredictable and hardly controllable.

The R2lab wireless testbed was designed with reproducibility as its central characteristics; it is built around an isolated and anechoic chamber, featuring RF absorbers that prevent radio waves reflections, and a Faraday cage blocking external interferences. R2lab thus provides an ideal environment for running reproducible wireless experiments.

R2lab has been operated since december 2015, in the context of the FIT (Future Internet of Things) Equipment of Excellence project, and as such, it is now federated with the other testbeds that are part of the FIT initiative. As of early 2019, it has been also federated within the Fed4Fire initiative.

Available toolsets, both hardware and software, are mostly stable apart from low noise marginal deployment of new kinds of radio devices, that now encompass among 5G and LoRa, among others. Our focus at this point of the project is to leverage our initial technical and financial investment, and

to produce scientific work around reproducibility, particularly from a methodological standpoint, as illustrated by various publications liste in the R2lab web site https://r2lab.inria.fr/papers.md.

Worth being mentioned as well, as part of a partnership with the OpenAirInterface initiative, R2lab is used on a daily basis for system-wide regression tests of the OAI stack, which in return allows us to offer up-to-date images for running OAI-based experiments. Emphasis has been put lately on offering tools that leverage kubernetes as the swiss-knife for orchestrating the deployment of a complete 5G infrastructure as an elastic set of microservices. Several research activities by the team this year use R2lab [22, 28, 27]

Access to R2lab is open 24/7. We currently have around 220 active users from all over the world among them 28 new users registered in 2020. For more details see http://r2lab.inria.fr.

7.2.2 Network simulator for aircrafts

- Keywords: network, simulation, real-time
- Functional Description: In collaboration with Safran Electrical and Power we produced a network design tool for aircrafts. This tool simulates aircraft networks. The tool is about 10,000 lines of code, out of which we produced 2,000.
- Assessment: A-2up,SO-3,SM-2up,EM-4,SDL-3,OC-DA-CD-TPM
- Licence: confidential
- URL: confidential
- Contact: Damien Saucez

8 New results

8.1 Service Transparency

8.1.1 When Deep Learning meets Web Measurements to infer Network Performance

Participants: Imane Taibi, Chadi Barakat.

The ability to monitor web and network performance becomes crucial to understand the reasons behind any service degradation. Such monitoring is also helpful to understand the relationship between the quality of experience of end users and the underlying network performance. Many troubleshooting tools have been proposed recently, which mainly consist of conducting active network measurements from within the browser, e.g., NDT, MobiPerf, SpeedTest, Fathom. Yet, these tools are either computationally expensive, less generic or greedy in terms of data consumption. The main purpose of this work funded by the IPL BetterNet is to leverage passive measurements freely available in the browser and machine learning techniques (ML) to infer network performance (e.g., delay, bandwidth and loss rate) without the addition of new measurement overhead. To enable this inference, we propose and implement a framework based on extensive controlled experiments where network configurations are artificially varied and the Web is browsed, then ML is applied to build models that estimate the underlying network performance. In particular, we contrast classical ML techniques (such as random forest) to deep learning models trained using fully connected neural networks and convolutional neural networks (CNN). We also compare the estimation accuracy of our approach with the best known web-based network measurement techniques available nowadays. The results of our study show that our approach can give a very good accuracy compared to other techniques, its accuracy is even higher than most standard techniques, and very close to the rest. Our experimental study also shows that neural networks have a higher accuracy compared to classical ML approaches. Furthermore, the model accuracy improves considerably using CNN. These results were published in [24, 23]. This study was performed in collaboration with Yassine Hadjadj-Aoul from the EPI DIONYSOS of Inria Rennes - Bretagne Atlantique.

8.1.2 Leveraging Website Popularity Differences to Identify Performance Anomalies

Participants: Giulio Grassi, Chadi Barakat.

This study was funded by Inria within IPL BetterNet and was performed in collaboration with the EPI MIMOVE of Inria Paris and the Department of Computer Science of Boston University. Its focus is on the detection of performance anomalies across websites with a large span of popularities. In fact, Web performance anomalies (e.g. time periods when metrics like page load time are abnormally high) have significant impact on user experience and revenue of web service providers. Existing methods to automatically detect web performance anomalies focus on popular websites (e.g. with tens of thousands of visits per minute). However, across a more representative set of websites, passive measurement volume varies enormously, and some sites will only have small numbers of measurements per hour. Low rates of measurement creates gaps and noise that prevent the use of existing methods. This study develops WMF, a web performance anomaly detection method applicable across a range of websites with highly variable measurement volume. To demonstrate our method, we leverage data from a website monitoring company partner of the IPL BetterNet, which allows us to leverage cross-site measurements. WMF uses matrix factorization to mine patterns that emerge from a subset of the websites to "fill in" missing data on other websites. Our validation using both a controlled website and synthetic anomalies shows that WMF's F1-score is more than double that of the state-of-the-art method. We then apply WMF to three months of web performance measurements to shed light on performance anomalies across a variety of 125 small to medium websites. The results of this work will be presented at the IEEE INFOCOM 2021 conference [20].

8.1.3 From Encrypted Video Traces to Viewport Classification

Participants: Othmane Belmoukadam, Chadi Barakat.

The Internet has changed drastically in recent years, multiple novel applications and services have emerged, all about consuming digital content. In parallel, users are no longer satisfied by the Internet's best effort service, instead, they expect a seamless service of high quality from the side of the network. This has increased the pressure on Internet service providers (ISP) in their effort to efficiently engineer their traffic and improve their end-users' experience. Content providers from their side, and to further protect the content of their customers, have shifted towards end-to-end encryption (e.g., TLS/SSL), which has complicated even further the task of ISPs in handling the traffic in their network. The challenge is notable for video streaming traffic which is driving the Internet traffic growth, and which imposes tight constraints on the quality of service provided by the network depending on the content of the video stream and the equipment on the end-user premises. Video streaming relies on the Dynamic Adaptive Streaming over HTTP (DASH) protocol which takes into consideration the underlying network conditions (e.g., delay, loss rate, and throughput) and the viewport capacity (e.g., screen resolution) to improve the experience of the end user in the limit of available resources. Nevertheless, knowing the reality of the encrypted video traffic is of great help to ISPs as it allows taking appropriate network management actions. In this work that fits within the PhD thesis of Othmane Belmoukadam, we propose an experimental framework able to infer fine-grained video flow information such as chunk sizes from encrypted YouTube video traces. We also present a novel technique to separate video and audio chunks from encrypted traces based on Gaussian Mixture Models (GMM). We evaluate our technique with real chunk sizes (Audio/Video) collected through the browser using the Chrome Web Request API. Then, we leverage these results and our dataset to train a model able to predict the class of viewport (either SD or HD) per video session with an average 92% accuracy and 85% F1 score. The results of this work were presented at the 16th International Conference on Network and Service Management (CNSM 2020) and were awarded the Best Paper Award [16].

8.1.4 Impact of browser viewport on video resolution patterns

Participants: Othmane Belmoukadam, Muhammad Jawad Khokhar, Chadi Barakat.

As stated above, the video streaming transmission is based on DASH which takes into consideration the underlying network conditions and the terminal characteristics to select the video resolution to request from the server. We question in this work the efficiency of this transmission in taking into account the terminal characteristics, the viewport in particular, knowing that requesting a resolution exceeding the viewport results in waste of bandwidth. Such bandwidth waste can either save money when the user is on a pay as you go data plane, or steal bandwidth from other users who are in need for it to further improve their Quality of Experience (QoE). To narrow the stats, we present a controlled experimental framework that leverages the YouTube and Dailymotion video players and the Chrome web request API to assess the impact of browser viewport on the observed video resolution pattern. In a first attempt of kind, we use the observed patterns to quantify the amount of wasted bandwidth. Our data-driven analysis points to high sensitivity of the Dailymotion player toward small viewports (240x144 and 400x225) compared to the YouTube player resulting in 15% and 8% less bandwidth waste respectively. However, as the users shift toward large viewports, the YouTube player becomes more viewport friendly compared to the Dailymotion player with shows an estimated bandwidth waste of 28%. This study has been published in [17].

8.1.5 Missed by Filter Lists: Detecting Unknown Third-Party Trackers with Invisible Pixels

Participants: Imane Fouad, Arnaud Legout.

Web tracking has been extensively studied over the last decade. To detect tracking, previous studies and user tools rely on filter lists. However, it has been shown that filter lists miss trackers. In this paper, we propose an alternative method to detect trackers inspired by analyzing behavior of invisible pixels. By crawling 84,658 webpages from 8,744 domains, we detect that third-party invisible pixels are widely deployed: they are present on more than 94.51% of domains and constitute 35.66% of all third-party images. We propose a fine-grained behavioral classification of tracking based on the analysis of invisible pixels. We use this classification to detect new categories of tracking and uncover new collaborations between domains on the full dataset of 4,216,454 third-party requests. We demonstrate that two popular methods to detect tracking, based on EasyList & EasyPrivacy and on Disconnect lists respectively miss 25.22% and 30.34% of the trackers that we detect. Moreover, we find that if we combine all three lists, 379,245 requests originated from 8,744 domains still track users on 68.70% of websites. This work has been published in PETS 2020 [19]. This study was performed in collaboration with Nataliia Bielova from the EPI PRIVATICS of Inria Grenoble Rhône-Alpes.

8.1.6 Evaluating smartphone performance for cellular power measurement

Participants: Arnaud Legout, Mondi Ravi, Yanis Boussad, Walid Dabbous.

Smartphones are affordable devices nowadays, capable of embedding a large variety of sensors such as magnetometers or orientation sensors, but also the hardware needed to connect them to most wireless communication technologies such as Wi-Fi, Bluetooth, or cellular networks. Therefore, they are handy devices able to perform received signal strength indicator RSSI measurements for a wide variety of applications such as cellular coverage maps, indoor localization, or proximity tracking. However, to the best of our knowledge, the accuracy of such measurements has never been rigorously assessed. The goals of this article are to assess the accuracy of the RSSI measurements made with a commercial off-the-shelf smartphone in a variety of conditions and how possible inaccuracies can be corrected. We primarily focus on the long-term evolution (LTE) RSSI, but we also extend our results to the Bluetooth RSSI. In this article, we build a controlled experimental setup based on commodity hardware and on open-source software. We evaluate the granularity and limitations of the Android application programming interface that returns the RSSI. We explore how reliable the measurements in a controlled environment with a mono-polarized antenna are. We show that the orientation of the smartphone, the position or orientation of the source, and the transmission power have a significant impact on the accuracy of the measurements. We introduce several correction techniques based on radiation matrix manipulations and on machine learning in order to improve measurement accuracy to less than 5 dB Root Mean Square Error, when compared to a professional equipment. We also explore the reliability of measurements

made in an outdoor realistic environment. We show that although transmission diversity available in LTE base stations significantly improves the measured RSSI regardless of the smartphone orientation, the Bluetooth RSSI remains largely sensitive to the smartphone orientation. The experimental set-up is publisehd in [18] and the complete work reported in [25] is published in Transactions on Instrumentation and Measurments [12]

8.2 Open Network Architecture

8.2.1 Efficient Pull-based Mobile Video Streaming leveraging In-Network Functions

Participants: Thierry Turletti.

There has been a considerable increase in the demand for high quality mobile video streaming services, while at the same time, the video traffic volume is expected to grow exponentially. Consequently, maintaining high quality of experience (QoE) and saving network resources are becoming crucial challenges to solve. In this work, we propose a name-based mobile streaming scheme that allows efficient video content delivery by exploiting a smart pulling mechanism designed for information-centric networks (ICNs). The proposed mechanism enables fast packet loss recovery by leveraging in-network caching and coding. Through an experimental evaluation of our mechanism over an open wireless testbed and the Internet, we demonstrate that the proposed scheme leads to higher QoE levels than classical ICN and TCP-based streaming mechanisms. This work has been presented at the IEEE Consumer Communications & Networking Conference (CCNC), in January 2020 at Las Vegas, USA [22]. The following link `https://github.com/fit-r2lab/demo-cefore` includes the artefacts that allows to reproduce performance results shown in the paper.

8.2.2 Quality of Experience-Aware Mobile Edge Caching through a Vehicular Cloud

Participants: Chadi Barakat.

Densification through small cells and caching in base stations have been proposed to deal with the increasing demand for Internet content and the related overload on the cellular infrastructure. However, these solutions are expensive to install and maintain. Instead, using vehicles acting as mobile caches might represent an interesting alternative. In this work, we assume that users can query nearby vehicles for some time, and be redirected to the cellular infrastructure when the deadline expires. Beyond reducing costs, in such an architecture, through vehicle mobility, a user sees a much larger variety of locally accessible content within only few minutes. Unlike most of the related works on delay tolerant access, we consider the impact on the user experience by assigning different retrieval deadlines per content. We provide the following contributions: (i) we model analytically such a scenario; (ii) we formulate an optimization problem to maximize the traffic offloaded while ensuring user experience guarantees; (iii) we propose two variable deadline policies; (iv) we perform realistic trace-based simulations, and we show that, even with low technology penetration rate, more than 60% of the total traffic can be offloaded which is around 20% larger compared to existing allocation policies. These results were published in the IEEE Transactions on Mobile Computing journal [14].

8.2.3 On Accounting for Screen Resolution in Adaptive Video Streaming: QoE driven bandwidth sharing framework

Participants: Othmane Belmoukadam, Muhammad Jawad Khokhar, Chadi Barakat.

Screen resolution along with network conditions are main objective factors impacting the user experience, in particular for video streaming applications. User terminals on their side feature more and more advanced characteristics resulting in different network requirements for good visual experience. Previous studies tried to link MOS (Mean Opinion Score) to video bitrate for different screen types (e.g., Common Intermediate Format (CIF), Quarter Common Intermediate Format (QCIF), and High Definition (HD)). We leverage such studies and formulate a QoE driven resource allocation problem to pinpoint the optimal bandwidth allocation that maximizes the QoE over all users of a network service provider located

behind the same bottleneck link, while accounting for the characteristics of the screens they use for video playout. For our optimization problem, QoE functions are built using curve fitting on datasets capturing the relationship between MOS, screen characteristics, and bandwidth requirements. We propose a simple heuristic based on Lagrangian relaxation and KKT (Karush Kuhn Tucker) conditions to efficiently solve the optimization problem. Our numerical simulations show that the proposed heuristic is able to increase overall QoE up to 20% compared to an allocation with a TCP look-alike strategy implementing max-min fairness. Results of this work were first published in CNSM 2019 then in Journal of Network Management [11].

8.2.4 Machine Learning for Next-Generation Intelligent Transportation Systems

Participants: Tingting Yuan, Thierry Turletti, Chadi Barakat.

Intelligent Transportation Systems (ITS) include a variety of services and applications such as road traffic management, traveler information systems, public transit system management , and autonomous vehicles, to name a few. It is expected that ITS will be an integral part of urban planning and future cities as it will contribute to improved road and traffic safety, transportation and transit efficiency. On the other hand, ITS poses a variety of challenges due to its scalability and diverse quality-of-service needs, as well as the massive amounts of data it will generate. We wrote a survey that explores the use of Machine Learning (ML) to enable ITS. In the context of the DrIVE associated team on Distributed Intelligent Vehicular Environments, we explored how Machine Learning (ML) technology can be applied to a broad range of ITS applications and services, such as cooperative driving and road hazard warning and submitted a survey to a journal [29]. We proposed a novel dynamic controller assignment algorithm that targets connected vehicle services and applications. Our approach considers a hierarchically distributed control plane, decoupled from the data plane, and uses vehicle location and control traffic load to perform controller assignment dynamically. We analyzed its performance using simulations with real-world vehicle mobility traces in a TNSM journal paper [15]. Then, we investigated how to leverage multiple 5G unmanned aerial vehicles (UAVs) to enhance network resource allocation among vehicles by positioning UAVs on-demand as a "flying communication infrastructure". We proposed a deep reinforcement learning (DRL) approach to determine the position of UAVs to improve the fairness and efficiency of network resource allocation. Our simulation results are promising and show that our proposed DRL approach to UAV positioning can help improve network resource allocation according to the targeted fairness objective [30].

8.2.5 Cross-layer Loss Discrimination Algorithms for MEC in 4G networks

Participants: Mamoutou Diarra, Thierry Turletti, Walid Dabbous.

Traditional loss-based Congestion Control Algorithms (CCAs) suffer from performance issues over wireless networks mostly because they fail to distinguish wireless random losses from congestion losses. Different loss discrimination algorithms have been proposed to tackle this issue but they are not efficient for 4G networks since they do not consider the impact of various link layer mechanisms such as adaptive modulation and coding and retransmission techniques on congestion in LTE Radio Access Networks (RANs). We propose MELD (MEC-based Edge Loss Discrimination), a novel server-side loss discrimination mechanism that leverages recent advancements in Multi-access Edge Computing (MEC) services to discriminate packet losses based on real-time RAN statistics. Our approach collects the relevant radio information via MEC's Radio Network Information Service and uses it to correctly distinguish random losses from congestion losses. Our experimental study made with the QUIC transport protocol shows over 80% higher goodput when MELD is used with New Reno and 8% higher goodput when used with Cubic. This work is described in a research report [27] and is currently under submission.

8.3 Experimental Evaluation

8.3.1 Time Sensitive Networks Evaluation

Participants: Damien Saucez.

Since he's back from Safran (October 2020), Damien Saucez has been working on developing a Time Sensitive Networking (TSN) simulation module for OMNeT++. This simulation module is designed for industrial usage and thus closely follows the standards. The idea here is that most potential industrial actors that would use the simulator have no real developper or networking simulation teams. Instead they have large pool of engineers used to read standards and evaluate appliances provided by third parties. Therefore it is essential to provide a solution that is as close as possible to what their day-to-day business is. As a consequence, some choices for the implementation of the simulation would appear inefficient for the ones used to write network simulators but they are not the target of the tool. Our focus on being usable directly by R&D teams in industries explains why exceptionally were are not using ns-3. In addition, as most corporate environments wouldn't allow to install specific development tools, we are working on providing a cloud based support for simulation that we call Simulation-as-a-Service. The backend is an Omnest (the commercial version of OMNeT++) engine while the frontend is purely javascript (REACT) such that simulations can be designed and analysed directly from a web browser. As OMNeT++ relies on a relatively old C++ technology writing a purely reactive javascript simulation environment for OMNeT++ is somehow challenging since nothing as really been thought for such usage.

In parallel, Damien Saucez is currently establishing new collaboration with Naval Group with the target to propose optimised virtual networks for the naval industry.

8.3.2 Distrinet: Distributed Network Emulation

Participants: Giuseppe Di Lena , Damien Saucez, Thierry Turletti.

Regarding the Networking Experimentation Reproducibility research axis, we have pursued on efforts on the Distrinet tool, a distributed implementation of Mininet over multiple hosts. We have analyzed its performance on Linux clusters as well as on the Amazon EC2 could platform, compared them with the ones of Mininet, Maxinet and Mininet Cluster Edition, and reported the results in the ACM CCR journal [13]. We also implemented a new placement module with different algorithms that take into account both link and node resources and minimize the number of physical hosts needed to carry out the emulation. We showed that our placement methods outperform existing ones and analyzed their performance in a ICC'21 conference paper [21]. An extended version of this work [26] is expected to be submitted to a journal. This study was performed in collaboration with Frédéric Giroire from the EP COATI of Inria Sophia Antipolis - Méditerranée.

8.3.3 Passive Delay Measurement for Fidelity Monitoring of Distributed Network Emulation

Participants: Houssam Elbouanani, Chadi Barakat, Walid Dabbous, Thierry Turletti.

Emulation has become an excellent approach for the validation and evaluation of network research. It provides researchers with a contained, customizable, and scalable testing environment, which can be easily packaged and published for potential readers to reproduce its results. However, as the network components are only virtual, it lacks the inherent realism of physical testbeds. In light of this, monitoring specific metrics of the emulated network has been proposed as a solution to mitigate to some degree inaccuracies caused by emulation. While this is not difficult to implement in single-machine settings (e.g. with Mininet), in scenarios where the emulation is distributed over multiple physical machines (e.g. Distrinet) any monitoring is limited by their asynchrony. In this work that fits within the PhD thesis of Houssam Elbouanani, we tackle the case of packet delay monitoring, to which we propose a methodology for passively measuring delays with or without underlying assumptions about time synchronization. We implement and evaluate our proposed methodology in an open testbed and show that it can reach results within few microseconds of perfect accuracy. This work is currently under submission [28].

8.3.4 Reproducible 5G network automation with Kube5G in R2lab

Participants: Thierry Parmentelat, Thierry Turletti, Walid Dabbous.

We have deployed a disaggregated 5G network using OAI Mosaic5G snaps and Kube5G in R2lab. We also produced a demonstration video, in which we showcase how we use kube5g (https://mosaic5g.io/kube5g/) to deploy a fully functional LTE network inside of R2lab, inside a k8s cloud running entirely inside the R2lab anechoic chamber. The video has been published on YouTube and the work was presented at the Mosaic 2020 Workshop online.

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

Collaboration with Safran
Participant: Damien Saucez.

Last year, the research collaboration with Safran on Constrained Software Defined Networks has evolved into a new stage: Damien Saucez took a one year secondment from Inria to join Safran and further develop this activity from "inside". After spending one year at Safran working on the next generation of digital networks for aircrafts, Damien Saucez came back to Inria. Given the crisis linked to the pandemic, at the current time the collaboration is on hold.

Collaboration with Ekinops
Participants: Thierry Turletti, Walid Dabbous.

We have started a collaboration with EKINOPS on the topic of Multi-access Edge Computing. The activity started with a CIFRE thesis. The PhD student Mamoutou Diarra started his PhD on this topic on November 2019. Currently, he is working on efficient congestion control mechanisms for 5G scenarios in mutli-access edge environments.

Collaboration with Orange
Participant: Thierry Turletti, Damien Saucez.

We have a collaboration with Orange on the topic of Network Function Virtualization. The activity includes the CIFRE PhD thesis of Giuseppe Di Lena that started his PhD on resilient NFV/SDN environments on April 2018 co-supervised by Thierry Turletti and Frédéric Giroire from the COATI project-team.

9.2 Bilateral grants with industry

QWANT
Participant: Arnaud Legout.

The PIA ANSWER project is led by the QWANT search engine and the Inria Sophia Antipolis Méditerranée research center. This proposal is the winner of the "Grand Challenges du Numérique" (BPI) and aims to develop the new version of the search engine http://www.qwant.com with radical innovations in terms of search criteria, indexed content and privacy of users. In the context of this project, we got with Nataliia Bielova from the INDES project-team a funding for a 3 years Ph.D. working on Web tracking technologies and privacy protection.

10 Partnerships and cooperations

10.1 Inria Internal Funding

The DIANA team was part of the Inria Project Lab BetterNet (http://project.inria.fr/betternet/) that ended in November 2020. Within this lab, Inria funded the PhD of Imane Taibi who is hosted by the DINOYSOS team in Rennes and is co-supervied by Chadi Barakat from the DIANA project-team and Gerardo Rubino and Yassine Hadjadj-Aoul from the DINOYSOS project-team. The PhD of Imane

Taibi started on the 1st of November 2017. Further in 2018, Inria funded a PostDoc position to supervise the experiments planned within the IPL and develop the data analysis part. This PostDoc position was occupied in 2020 by Giulio Grassi who was co-supervised by Chadi Barakat from the Diana project-team and Renata Teixeira from the MIMOVE project-team. Giulio Grassi started on October 1st, 2018 and left Inria in June 2020.

10.2 International initiatives

10.2.1 Inria Associated Teams involved in an Inria International Lab

DrIVE

Title: Distributed Intelligent Vehicular Environment - Enabling ITS through programmable networks

Duration: 2018 - 2021

Coordinator: Thierry Turletti

Partners: UniCamp (Brazil) - Department of Computer Engineering and Industrial Automation (Mateus Augusto Silva Santos and Christian Esteve Rothenberg) and UC Santa Cruz (USA) - Department of Computer Science and Engineering (Katia Obraczka).

Inria contact: Thierry Turletti

Summary: Transportation systems are part of our society's critical infrastructure and are expected to experience transformative changes as the Internet revolution unfolds. The automotive industry is a notable example: it has been undergoing disruptive transformations as vehicles transition from traditional unassisted driving to fully automated driving, and eventually to the self-driving model. Communication technology advancements such as support for vehicle-to-infrastructure (V2I) and vehicle-to-vehicle (V2V) communication have been one of the key enablers of next generation transportation services, also known as Intelligent Transport Systems (ITS). However, ITS services and applications pose significant challenges to the underlying communication and network infrastructure due to their stringent low latency, reliability, scalability, and geographic decentralization requirements. The DrIVE associated team proposal aims at addressing such challenges by: (1) developing a programmable network control plane that will dynamically adjust to current environment conditions and network characteristics to support ITS' scalability, quality of service, and decentralization requirements, and (2) applying the proposed distributed network control plane framework to ITS applications, such as road hazard warning, autonomous- and self-driving vehicles, and passenger-centric services (e.g., infotainment and video streaming).

10.3 International research visitors

- Wilson Borba Da Rocha Neto
  - Date: Feb 2020 - Mar 2020 (shortened because of COVID19 pandemy)
  - Institution: Univ Campinas
  - Subject: Emulated mobility for fixed-grid testbeds
  - Supervisor: Thierry Turletti

10.4 European initiatives

10.4.1 FP7 & H2020 Projects

Fed4FIREplus

Title: Federation for FIRE Plus

Duration: Jan 2017 - Dec 2021

Coordinator: iMinds - Belgium

Partners: 20 european partners including IMEC (Belgium), UPMC (Fr), Fraunhofer (Germany), TUB (Germany), etc.

Inria contact: Thierry Parmentelat

Summary: The Fed4FIRE+ project has the objective to run and further improve Fed4FIRE as best-in-town federation of experimentation facilities for the Future Internet Research and Experimentation initiative. Federating a heterogeneous set of facilities covering technologies ranging from wireless, wired, cloud services and open flow, and making them accessible through common frameworks and tools suddenly opens new possibilities, supporting a broad range of experimenter communities covering a wide variety of Internet infrastructures, services and applications. Fed4FIRE+ will continuously upgrade and improve the facilities and include technical innovations, focused towards increased user satisfaction (user-friendly tools, privacy-oriented data management, testbed SLA and reputation, experiment reproducibility, service-level experiment orchestration, federation ontologies, etc.). It will open this federation to the whole FIRE community and beyond, for experimentation by industry and research organisations, through the organization of Open Calls and Open Access mechanisms. The project will also establish a flexible, demand-driven framework which allows test facilities to join during the course of its lifetime by defining a set of entry requirements for new facilities to join and to comply with the federation. FIRE Experimental Facilities generate an ever increasing amount of research data that provides the foundation for new knowledge and insight into the behaviour of FI systems. Fed4FIRE+ will participate in the Pilot on Open Research Data in Horizon 2020 to offer open access to its scientific results, to the relevant scientific data and to data generated throughout the project's lifetime. Fed4FIRE+ will finally build on the existing community of experimenters, testbeds and tool developers and bring them together regularly (two times a year) in engineering conferences to have maximal interaction between the different stakeholders involved.

SLICES-SC

Title: Scientific Large-scale Infrastructure for Computing/Communication Experimental Studies– Starting Community

Duration: Mar 2021 - Feb 2024

Coordinator: Sorbonne Univesité - France

Partners: UTH, Mandat International, PSNC, IMDEA, CNR, EURECOM, COSM, IoTLab, University of Oulu, INRIA, Imec, SZTAKI, TUM.

Inria contact: Walid Dabbous

Summary: Today we are experiencing the digital transformation happening with an unprecedented pace, with the community constantly researching on new solutions to support this transformation with ample computational power and connectivity. Towards addressing such research efforts, Research Infrastructure (RI) specific to addressing Digital Sciences research efforts have been deployed worldwide, towards trying to address key aspects contrary to off- the-shelf commercial infrastructure: 1) Full control over the parameters of an experiment, 2) Repeatable experiments regardless of the physical infrastructure, 3) Valid experimental results, which are easy to cross-reference and replicate. As such, several RIs have emerged, offering experimentation services with bleeding edge resources, that otherwise are only offered only in industrial R&D laboratories, with limited functionality. Towards combating these issues, SLICES Research Infrastructure is about to be deployed, aiming to provide high quality experimentation services with emerging technologies around the area of digital sciences (5G/6G, NFV, IoT and Cloud Computing), in an Internet-scale setup. With SLICES-SC, we aspire to foster the community of researchers around this ecosystem, create and strengthen necessary links with relevant industrial stakeholders for the exploitation of the infrastructure, advance existing methods for research reproducibility and

experiment repeatability, and design and deploy the necessary solutions for providing SLICES-RI with an easy to access scheme for users from different disciplines. A set of detailed research activities has been designed to materialize these efforts in tools for providing transnational (remote and physical) access to the facility, as well as virtual access to the data produced over the facilities. The respective networking activities of the project aspire in fostering the community around these infrastructures, as well as open up to new disciplines and industrial stakeholders.

10.5 National initiatives

- **ANR FIT** (2011-2020): FIT (Future Internet of Things) aims at developing an experimental facility, a federated and competitive infrastructure with international visibility and a broad panel of customers. It will provide this facility with a set of complementary components that enable experimentation on innovative services for academic and industrial users. The project will give French Internet stakeholders a means to experiment on mobile wireless communications at the network and application layers thereby accelerating the design of advanced networking technologies for the Future Internet. FIT is one of 52 winning projects from the first wave of the French Ministry of Higher Education and Research's Equipements of Excellence (Equipex) research grant programme. The project will benefit from a 5.8 million euro grant from the French government. Other partners are UPMC, IT, Strasbourg University and CNRS. The project ended in 2020. See also https://fit-equipex.fr/.

- **ANR BOTTLENET** (2016-2021): BottleNet aims to deliver methods, algorithms, and software systems to measure Internet Quality of Experience (QoE) and diagnose the root cause of poor Internet QoE. This goal calls for tools that run directly at users' devices. The plan is to collect network and application performance metrics directly at users' devices and correlate it with user perception to model Internet QoE, and to correlate measurements across users and devices to diagnose poor Internet QoE. This data-driven approach is essential to address the challenging problem of modeling user perception and of diagnosing sources of bottlenecks in complex Internet services. ANR BottleNet will lead to new solutions to assist users, network and service operators as well as regulators in understanding Internet QoE and the sources of performance bottleneck. This national project got extended until March 2021 because of the COVID pandemic, and is now coordinated, starting from November 2020, by Chadi Barakat from the Diana team.

11 Dissemination

11.1 Promoting scientific activities

Chadi Barakat is on the editorial board of the Computer Networks journal, and was/is on the Technical Program Committee for the International Teletraffic Congress (ITC 33), the Network Traffic Measurement and Analysis Conference (TMA 2020 and 2021), the International Conference on Network and Service Management (CNSM 2020), the 19th Mediterranean Communication and Computer Networking Conference (MedComNet 2021), and the first International Workshop on Machine Learning in Networking (MaLeNe 2021). He is currently in charge of international affairs at Inria Sophia Antipolis and is member of the COST-GTRI of Inria.

Walid Dabbous is Director of the Academy of Excellence NIDS (Networks, Information and Digital Society). He is also member of the scientific committee of the DS4H Graduate school and member of the Ubinet International Master program steering committee. He was also member of the CRCN/ISFP Inria researchers recruitment jury in 2020.

Arnaud Legout is member of the scienfic board of the Regalia project whose aim is to build a software environment for testing and regulation support to deal with the risks of bias and disloyalty generated by the algorithms of digital platforms. He was also member of the CRCN/ISFP jury in 2020. He served for ten years (2010-2020) as the head of the CUMI (Commission des Utilisateurs des Moyens), the local IT users commission at Inria Sophia-Antipolis.

Thierry Turletti is/was in the program committees of the following workshops and conferences: 23rd Algotel Conference, 2021, IEEE International Conference on Communications (ICC'21), IFIP/IEEE Interna-

tional Symposium on Integrated Network Management (IM'21), 11-12th Workshops on ns-3 (WNS'20-21) and 17th IFIP/IEEE International Symposium on Integrated Network Management (NOMS'20). He is still on the editorial boards of the "Wireless Networks" journal published by Springer Science and of the "Advances in Multimedia" Journal published by Hindawi Publishing Corporation. Since June 2020, Thierry Turletti became a member of the *Comité de Suivi Doctoral (CSD)* and left his post as *President of the Commission of Development and Technology (CDT)* for the INRIA Sophia Antipolis Méditerranée Research Center which he had held since November 2010.

11.2 Teaching - Supervision - Juries

Teaching

- Master 2 Ubinet: Chadi Barakat and Walid Dabbous, Evolving Internet, 31.5 hours, M2, Univ. Côte d'Azur, France.
- Master 2 Ubinet: Chadi Barakat and Walid Dabbous, Internet Measurements and New Architectures, 31.5 hours, M2, Univ. Côte d'Azur, France.
- Master 1 in Computer Science: Chadi Barakat, Computer Networks, 15 hours, M1, Univ. Côte d'Azur, France.
- Master 1 in Computer Science: Chadi Barakat, Internet of the Future, 15 hours, M1, Univ. Côte d'Azur, France
- Master Estel: Chadi Barakat, Voice over IP, 9 hours, Univ. Côte d'Azur, France.
- Master 2 Ubinet: Arnaud Legout, From BitTorrent to Privacy, 22.5 hours, M2, Univ. Côte d'Azur, France.
- Master IUP GMI: Damien Saucez, Security and privacy in networks, 38h, M2, University of Avignon, France.
- **E-learning**
  - Python: Arnaud Legout and Thierry Parmentelat are co-authors of the UCA MOOC Python 3 : "Python 3 : des fondamentaux aux concepts avancés du langage" that lasts 9 weeks on FUN (https://www.france-universite-numerique-mooc.fr/). For the second session there were 49 369 registered persons. In total, this MOOC on all its editions has been followed by 94 559 persons.

Supervision

PhD students

- PhD in progress: Othmane Belmoukadam works on "QoE aware content management in the Internet caching and transport". He is supervised by Chadi Barakat and funded by the doctoral school EDSTIC of Université Côte d'Azur (UCA).
- PhD in progress: Yanis Boussad works on "Large scale characterization of the exposition to microwaves". He is co-supervised with Leonardo Lizzi, LEAT.
- PhD in progress: Giuseppe Di Lena works on "Building a resilience methodology for NFV/SDN" in Apr 2018. His PhD is co-supervised by Thierry Turletti, Damien Saucez and Frédéric Giroire from the Coati project-team .
- PhD in progress: Mamoutou Diarra started his PhD on "Multi-access Edge Computing" in november 2019. He is co-supervised by Thierry Turletti, Walid Dabbous and Amine Ismail from Ekinops.

- PhD in progress: Houssam Elbouanani started his PhD on “Experiment control for reproducible research” in december 2019. He is co-supervised by Walid Dabbous, Chadi Barakat and Thierry Turletti.
- PhD in progress: Iman Fouad started her PhD on on Web tracking technologies and privacy protection in november 2017. Her thesis is co-supervised by Arnaud Legout and Nataliia Bielova (Indes).
- PhD in progress: Imane Taibi works on “Big data analysis for network monitoring and troubleshooting”. She is co-supervised by Gerardo Rubino, Yassine Hadjadj-Aoul from the Dionysos project-team and Chadi Barakat.
- PhD defended: Mathieu Thiery finsihed his PhD on “Data protection of connected objects and smartphones” in December 2020. He was co-supervised by Vincent Roca from the Privatics project-team and Arnaud Legout.

Internes

- Chap Chanpiseth
  - Date: Mar 2020 - Aug 2020
  - Institution: Université Côte D’Azur
  - Mining the ACQUA dataset: understanding mobile user Quality of Experience in the wild
  - Supervisor: Chadi Barakat
- Bernard Tamba Sandouno
  - Date: Mar 2020 - Aug 2020
  - Institution: Université Côte D’Azur
  - Subject: Confronting different network measurements datasets for a better understand-ing and troubleshooting of User Quality of Experience
  - Supervisor: Chadi Barakat
- Angelo Rodio
  - Date: Mar 2020 - Aug 2020
  - Institution: Université Côte D’Azur
  - Subject: Data Augmentation for Wi-Fi Fingerprint-Based Localization in Indoor Environments
  - Supervisor: Arnaud Legout
- Amal Krimi
  - Date: Nov 2020 - Dec 2020
  - Institution: Université Côte D’Azur
  - Subject: Cooperative Localization in LoRa Low Power Wide Area Networks
  - Supervisors: Thierry Turletti and Walid Dabbous
- Anass El Boujidi
  - Date: Nov 2020 - Dec 2020
  - Institution: Université Côte D’Azur
  - Subject: QoE-aware bandwidth sharing for video streaming traffic
  - Supervisor: Chadi Barakat

Juries

- Arnaud Legout served as reviewer of Tompoariniaina Andriamilanto PhD thesis, "Leveraging Browser Fingerprinting for Web Authentication" defended on December 17th 2020, at University of Rennes 1.
- Thierry Turletti serves as reviewer of Adrien Wion's PhD thesis, "Control Plane in Dynamic Software Networks", to be defended on March 5, 2021, at Telecom Paris.
- Thierry Turletti served as examiner and president of El Fadel Bonfoh's PhD thesis "VTL: Une Architecture Stable pour la Conception, l'Implémentation, et le Déploiement de Protocoles de Communication d'Internet" defended on January 26, 2021 at INSA Toulouse.
- Thierry Turletti served as reviewer of Salma Matoussi's PhD thesis "User-Centric Slicing with Functional Splits in 5G Cloud-RAN" defended on January 22, 2021 at Sorbonne Université. He served as reviewer of Soufian Toufga's PhD thesis "Vers des réseaux véhiculaires programmables grâce à la technologie SDN" defended on October 29, 2020 at Université Toulouse 3 Paul Sabatier.
- Thierry Turletti served as jury member for the mid-term review of the PhD thesis of Tsu-Han Wang (Eurecom) for her thesis on 5G modem architecture in October 2020.
- Thierry Turletti served as reviewer of Hoang Long Mai's PhD thesis "Towards a Content-Oriented Security Plane for Named Data Networking: Application to Content Poisoning Attack" defended on November 25, 2020 at Université Technologique de Troyes (UTT).
- Walid Dabbous served as a president of Guillaume Doyen HDR thesis "Intégration du Comportement des Entités Terminales dans la Disponibilité des Services à Grande Echelle" defended on January 12, 2021 at Université Technologique de Compiègne.
- Chadi Barakat served as reviewer of Theodoros Giannakas' PhD thesis, "Joint Modeling and Optimization Caching and Recommendation Systems", defended in March 2020, at Eurecom.
- Chadi Barakat served as jury member for the mid-term review of the PhD thesis of Duncan Deveaux (Eurecom) for his thesis titled "Knowledge networking for next generation vehicular networks" in January 2020.

11.3 Popularization

11.3.1 Articles and contents

Walid Dabbous has prepared a large audience video on electromagnetic waves in 5G context. The video has been published on YouTube and on the Université Côte D'Azur web site.

Arnaud Legout has given an interview to radio France Bleu on the risks related to the deplolyment of 5G and the role of the ElectroSmart application in detecting and minimizing our exposure to Electromagnetic field.

12 Scientific production

12.1 Major publications

[1] F. De Pellegrini, L. Maggi, A. Massaro, D. Saucez, J. Leguay and E. Altman. 'Blind, Adaptive and Robust Flow Segmentation in Datacenters'. In: *INFOCOM 2018 - IEEE International Conference on Computer Communications*. Honolulu, United States, Apr. 2018. URL: https://hal.inria.fr/hal-01666905.

[2] M. Flittner, M. N. Mahfoudi, D. Saucez, M. Wählisch, L. Iannone, V. Bajpai and A. Afanasyev. 'A Survey on Artifacts from CoNEXT, ICN, IMC, and SIGCOMM Conferences in 2017'. In: *Computer Communication Review* 48.1 (Apr. 2018), pp. 75–80. URL: https://hal.inria.fr/hal-01968401.

[3] M. Gabielkov, A. Ramachandran, A. Chaintreau and A. Legout. 'Social Clicks: What and Who Gets Read on Twitter?' In: *ACM SIGMETRICS / IFIP Performance 2016*. Antibes Juan-les-Pins, France, June 2016. URL: https://hal.inria.fr/hal-01281190.

[4] M. J. Khokhar, T. Ehlinger and C. Barakat. 'From Network Traffic Measurements to QoE for Internet Video'. In: *IFIP Networking Conference 2019*. Varsovie, Poland, May 2019. DOI: 10.23919/IFIPNetworking.2019.8816854. URL: https://hal.inria.fr/hal-02074570.

[5] M. J. Khokhar, N. A. Saber, T. Spetebroot and C. Barakat. 'An Intelligent Sampling Framework for Controlled Experimentation and QoE Modeling'. In: *Computer Networks* 147 (Dec. 2018), pp. 246–261. DOI: 10.1016/j.comnet.2018.10.011. URL: https://hal.inria.fr/hal-01906145.

[6] M. N. Mahfoudi, T. Parmentelat, T. Turletti, W. Dabbous and R. Knopp. *Deploy a 5G network in less than 5 minutes: Demo Abstract*. ACM SIGCOMM Posters and Demos. Poster. Aug. 2017. URL: https://hal.inria.fr/hal-01580065.

[7] M. N. Mahfoudi, G. Sivadoss, O. Bensouda Korachi, T. Turletti and W. Dabbous. 'Joint range extension and localization for LPWAN'. In: *Internet Technology Letters* (June 2019). DOI: 10.1002/itl2.120. URL: https://hal.archives-ouvertes.fr/hal-02170466.

[8] D. Saucez, L. Iannone, C. Albert and F. Coras. *Locator/ID Separation Protocol (LISP) Impact*. Internet Engineering Task Force (IETF), Request for Comments: 7834. Apr. 2016. URL: https://hal.inria.fr/hal-01423163.

[9] H. Soni, W. Dabbous, T. Turletti and H. Asaeda. 'NFV-based Scalable Guaranteed-Bandwidth Multicast Service for Software Defined ISP networks'. In: *IEEE Transactions on Network and Service Management* 14.4 (Dec. 2017), p. 14. DOI: 10.1109/TNSM.2017.2759167. URL: https://hal.inria.fr/hal-01596488.

[10] L. Vigneri, T. Spyropoulos and C. Barakat. 'Low Cost Video Streaming through Mobile Edge Caching: Modelling and Optimization'. In: *IEEE Transactions on Mobile Computing* (2018). DOI: 10.1109/TMC.2018.2861005. URL: https://hal.inria.fr/hal-01855304.

12.2 Publications of the year

International journals

[11] O. Belmoukadam, M. J. Khokhar and C. Barakat. 'On Accounting for Screen Resolution in Adaptive Video Streaming: QoE driven bandwidth sharing framework'. In: *International Journal of Network Management* 31.1 (7th Jan. 2021), e2128. DOI: 10.1002/nem.2128. URL: https://hal.inria.fr/hal-02615576.

[12] Y. Boussad, M. N. Mahfoudi, A. Legout, L. Lizzi, F. Ferrero and W. Dabbous. 'Evaluating Smartphone Accuracy for RSSI Measurements'. In: *IEEE Transactions on Instrumentation and Measurement* 70 (5th Jan. 2021), pp. 1–12. DOI: 10.1109/tim.2020.3048776. URL: https://hal.inria.fr/hal-03063997.

[13] G. Di Lena, A. Tomassilli, D. Saucez, F. Giroire, T. Turletti and C. Lac. 'Distrinet: a Mininet Implementation for the Cloud'. In: *Computer Communication Review* (Jan. 2021). URL: https://hal.inria.fr/hal-03000617.

[14] L. Vigneri, T. Spyropoulos and C. Barakat. 'Quality of Experience-Aware Mobile Edge Caching through a Vehicular Cloud'. In: *IEEE Transactions on Mobile Computing* 19.9 (1st Sept. 2020), pp. 2174–2188. DOI: 10.1109/TMC.2019.2921765. URL: https://hal.inria.fr/hal-02145252.

[15] T. Yuan, W. D. Rocha Neto, C. E. Rothenberg, K. Obraczka, C. Barakat and T. Turletti. 'Dynamic Controller Assignment in Software Defined Internet of Vehicles through Multi-Agent Deep Reinforcement Learning'. In: *IEEE Transactions on Network and Service Management* (2021). DOI: 10.1109/TNSM.2020.3047765. URL: https://hal.inria.fr/hal-03000911.

International peer-reviewed conferences

[16] O. Belmoukadam and C. Barakat. 'From Encrypted Video Traces to Viewport Classification'. In: CNSM 2020 - 16th International Conference on Network and Service Management. Virtual Conference, France, 2nd Nov. 2020. DOI: 10.23919/CNSM50824.2020.9269086. URL: https://hal.inria.fr/hal-02947058.

[17] O. Belmoukadam, M. J. Khokhar and C. Barakat. 'On excess bandwidth usage of video streaming: when video resolution mismatches browser viewport'. In: NoF 2020 - 11th IEEE International Conference on Networks of the Future. Bordeaux, France, 12th Oct. 2020. DOI: 10.1109/NoF50125.2020.9249133. URL: https://hal.inria.fr/hal-02883593.

[18] Y. Boussad, A. Legout, W. Dabbous, L. Lizzi, F. Ferrero and M. Naoufal Mahfoudi. 'Open-Source 4G Experimental Setup'. In: 2020 IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting. Montreal, Canada: https://2020apsursi.org/, 5th July 2020. URL: https://hal.archives-ouvertes.fr/hal-02515793.

[19] I. Fouad, N. Bielova, A. Legout and N. Sarafijanovic-Djukic. 'Missed by Filter Lists: Detecting Unknown Third-Party Trackers with Invisible Pixels'. In: PETS 2020 - 20th Privacy Enhancing Technologies Symposium. PETs (Privacy Enhancing Technologies Symposium). Montréal, Canada, 14th July 2020. URL: https://hal.inria.fr/hal-01943496.

[20] G. Grassi, R. Teixeira, C. Barakat and M. Crovella. 'Leveraging Website Popularity Differences to Identify Performance Anomalies'. In: INFOCOM 2021 - IEEE International Conference on Computer Communications. Vancouver / Virtual, Canada, 10th May 2021. URL: https://hal.inria.fr/hal-03109717.

[21] G. D. Lena, A. Tomassilli, F. Giroire, D. Saucez, T. Turletti and C. Lac. 'A Right Placement Makes a Happy Emulator: a Placement Module for Distributed SDN/NFV Emulation'. In: IEEE International Conference on Communications (ICC). Montréal, Canada, 14th June 2021. URL: https://hal.inria.fr/hal-03001913.

[22] K. Matsuzono, H. Asaeda, I. Naladala and T. Turletti. 'Efficient Pull-based Mobile Video Streaming leveraging In-Network Functions'. In: IEEE CCNC 2020 - Consumer Communications & Networking Conference. Las Vegas, United States, 10th Jan. 2020. URL: https://hal.inria.fr/hal-02359569.

[23] I. Taibi, Y. Hadjadj-Aoul and C. Barakat. 'Data Driven Network Performance Inference From Within The Browser'. In: PEDISWESA 2020 - 12th IEEE Workshop on Performance Evaluation of Communications in Distributed Systems and Web based Service Architectures. Rennes, France, 7th July 2020, pp. 1–6. DOI: 10.1109/ISCC50000.2020.9219573. URL: https://hal.inria.fr/hal-02871873.

[24] I. Taibi, Y. Hadjadj-Aoul and C. Barakat. 'When Deep Learning meets Web Measurements to infer Network Performance'. In: CCNC 2020 - IEEE Consumer Communications & Networking Conference. Las Vegas, United States: https://ccnc2020.ieee-ccnc.org/, 10th Jan. 2020, pp. 1–6. DOI: 10.1109/CCNC46108.2020.9045116. URL: https://hal.inria.fr/hal-02358004.

Reports & preprints

[25] Y. Boussad, M. N. Mahfoudi, A. Legout, L. Lizzi, F. Ferrero and W. Dabbous. *Evaluating Smartphone Accuracy for LTE Power Measurement*. 14th Apr. 2020. URL: https://hal.inria.fr/hal-02320342.

[26] G. Di Lena, A. Tomassilli, F. Giroire, D. Saucez, T. Turletti and C. Lac. *Placement Module for Distributed SDN/NFV Network Emulation*. Inria Sophia Antipolis - Méditerranée; I3S, Université Côte d'Azur; Orange Labs R&D [Lannion] (France Télécom), 5th Feb. 2021, p. 32. URL: https://hal.inria.fr/hal-03132873.

[27] M. Diarra, T. Turletti, W. Dabbous and A. Ismail. *Cross-layer Loss Discrimination Algorithms for MEC in 4G networks*. 12th Nov. 2020. URL: https://hal.inria.fr/hal-03001893.

[28] H. Elbouanani, W. Dabbous, C. Barakat and T. Turletti. *Passive Delay Measurement for Fidelity Monitoring of Distributed Network Emulation*. 12th Nov. 2020. URL: https://hal.inria.fr/hal-03001876.

[29] T. Yuan, W. B. Da Rocha Neto, C. E. Rothenberg, K. Obraczka, C. Barakat and T. Turletti. *Machine Learning for Next-Generation Intelligent Transportation Systems: A Survey*. 28th Nov. 2020. URL: https://hal.inria.fr/hal-02284820.

[30] T. Yuan, C. E. Rothenberg, K. Obraczka, C. Barakat and T. Turletti. *Harnessing UAVs for Fair 5G Bandwidth Allocation in Vehicular Communication via Deep Reinforcement Learning*. 12th Nov. 2020. URL: https://hal.inria.fr/hal-03001383.