



Applications and Devices Benchmarking

Paper Title: An End-to-End Testing Ecosystem for 5G the TRIANGLE Testing House Test Bed

Authors: Andrea F. Cattoni¹, German Corrales Madueño¹, Michael Dieudonne², Pedro Merino³, Almudena Diaz Zayas³, Alberto Salmeron³, Frederik Carlier⁴, Bart Saint Germain⁴, Donal Morris⁵, Ricardo Figueiredo⁵, Jeanne Caffrey⁵, Janie Baños⁶, Carlos Cardenas⁶, Niall Roche⁷, Alastair Moore⁷

1. Keysight Technologies Denmark, {andrea.cattoni, german.madueno}@keysight.com
2. Keysight Technologies Belgium
3. University of Malaga, Spain, {pedro, almudiaz, salmeron}@lcc.uma.es
4. Quamotion, Belgium
5. RedZinc, Ireland {dmorris, ricardo}@redzinc.net
6. AT4 wireless, Spain
7. University College London, UK

Published in: Journal of Green Engineering, Special Issue "Green Engineering and Energy Efficiency in the Context of 5G"

DOI (link to publication from Publisher): 10.13052/jge1904-4720.633

Publication date: 2016

Document Version Accepted author manuscript, peer reviewed version.

Citation for published version (IEEE):

Andrea F. Cattoni et al, "An End-to-End Testing Ecosystem for 5G the TRIANGLE Testing House Test Bed" in Journal of Green Engineering, Special Issue "Green Engineering and Energy Efficiency in the Context of 5G", vol. 6, issue 3, pp. 285-316, ©2016 River Publisher. DOI: 10.13052/jge1904-4720.633

Please note that copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us at info@triangle-project.eu providing details, and we will remove access to the work immediately and investigate your claim.

An End-to-End Testing Ecosystem for 5G

The TRIANGLE Testing House Test Bed

Andrea F. Cattoni¹, German Corrales Madueño¹, Michael Dieudonne², Pedro Merino³, Almudena Diaz Zayas³, Alberto Salmeron³, Frederik Carlier⁴, Bart Saint Germain⁴, Donal Morris⁵, Ricardo Figueiredo⁵, Jeanne Caffrey⁵, Janie Baños⁶, Carlos Cardenas⁶, Niall Roche⁷, Alastair Moore⁷

1. Keysight Technologies Denmark, {andrea.cattoni, german.madueno}@keysight.com
2. Keysight Technologies Belgium
3. University of Malaga, Spain, {pedro, almudiaz, salmeron}@lcc.uma.es
4. Quamotion, Belgium
5. RedZinc, Ireland {dmorris, ricardo}@redzinc.net
6. AT4 wireless, Spain
7. University College London, UK

Abstract— Mobile applications will be a dominant element in the 5G domain. Ensuring the correct and efficient behaviour of the applications and devices on which they run becomes a critical factor to ensure the mobile communications market meets the expectations of final users. The EU project TRIANGLE is building a framework to help app developers and device manufacturers in the evolving 5G sector to test and benchmark new mobile applications, devices, and services utilizing existing and extended FIRE testbeds. This innovative framework will facilitate the evaluation of the Quality of Experience and enable certification for new mobile applications and devices.

Keywords—5G systems; testing; Quality of Service; Quality of Experience; Mobile Apps; Software Defined Networks; Business Network Applications

I. INTRODUCTION

Mobile communications is one of the biggest growth areas in the Information and Communication Technologies (ICT) market, and it is expected to reach 4.7 billion mobile devices connected to the Internet by 2020 [1], with an overall economic revenue of \$373 billion in 2017 in the smartphone market alone [2]. Mobile apps will be a major software stakeholder in this huge ecosystem. At the same time 5G will become the de facto connecting technology standard for the majority mobile devices, including cellular-based Internet of Things (IoT) ones. Ensuring a correct and efficient behaviour of the applications and devices becomes a critical factor in meeting the expectations of end users. Extended battery life for any type of device, whether a smartphone or an IoT sensor, is one of those expectations. The TRIANGLE framework represents a great opportunity for European enterprises providing innovative testing tools based on testbeds with 5G capabilities. The purpose is to develop realistic testing methodologies that will complement other 5G research and innovation activities, helping to maintain the traditional European leadership in the mobile communications arena.

While radio related certification of mobile devices already has a strong standards based ecosystem thanks to the collaboration among the European Telecommunication Standards Institute (ETSI), 3rd Generation Partnership Program (3GPP), Global Certification Forum (GCF), vendors and test houses, there is still a lack of consensus on the benchmarking or testing methods at the apps level. The Key Performance

Indicators (KPIs) and techniques to perform such kind of testing for apps could be rather different than the certification methods associated with 3GPP standards for mobile devices (user equipment or UEs), partially due to the proprietary nature of many widely used apps, lack of a testing methodology and lack of standardized interfaces. The actors involved in producing and exploiting mobile apps are different in number and categories compared with those involved in mobile device manufacturing and integration. Thousands of Small and Medium Enterprises (SMEs) build apps compared to the few big vendors that build smartphones. Therefore, existing solutions for testing apps represent a very fragmented market with many ad-hoc tools associated to vendors.

From a technical point of view, the key issue is to base the testing of apps and mobile devices in real but controllable conditions (testbeds) that simulate/emulate realistic 5G scenarios as defined by Next Generation Mobile Networks (NGMN) [3]. Testbeds that combine proper realistic hardware and software, are currently provided in Europe by the FIRE project, but do not cover the new 5G scenarios. It is nowadays clear that 5G will be a very complex system. Such complexity could negatively affect the expected duration of the device battery. While studies of power consumption correlated to device models and network conditions are present in the literature, they are unfortunately very rare and use ad-hoc methods that cannot be directly considered for the implementation of a certification process.

From a market point of view, traditional test houses, that are key actors in the success of mobile technologies, could be playing again a leading role in the definition and exploitation of a new generation of tools for the emergent market of mobile apps plus devices in the 5G ecosystem. Other players like research institutes, innovative SMEs, test equipment industry, device vendors, etc., will also be necessary to develop such solutions.

II. THE TRIANGLE PROJECT

The “5G Applications and Devices Benchmarking” (TRIANGLE) project [4] is a European Union (EU) funded project within the framework of the “Horizon 2020” (H2020) initiative under FIRE/FIRE+ objective devoted to Future Internet Research and Experimentation. The project started in

January 2016 and has an expected duration of three years. Among its partners there are laboratories, SMEs, test equipment suppliers and academia. The primary objective of the TRIANGLE project is to promote the testing and benchmarking of mobile applications and devices in Europe as the industry moves towards 5G and to provide a pathway towards certification in order to support qualified apps and mobile developments using FIRE testbeds as testing framework.

TRIANGLE is developing a framework that facilitates the evaluation of the Quality of Experience (QoE) of new mobile applications, services and devices designed to operate in future 5G mobile broadband networks. The framework will exploit existing FIRE facilities adding new facilities when necessary. The project will identify reference deployment scenarios, will define new KPIs and QoE metrics, will develop new testing methodologies and tools, and will design a complete evaluation scheme. The project will develop a framework to ensure users QoE in new challenging situations, especially those due to heterogeneous networks and considering the role software will have in the new 5G ecosystem. The framework, will also provide the means to allow certification and obtain a quality mark for the applications, services and devices which are compliant with the requirements and test specifications developed in the project. This will allow vendor differentiation, especially among startups and SMEs, in the current globalized and competitive. The framework, methods and tools developed during the project will focus on providing the mechanisms to incorporate new wireless technologies and topologies envisaged in 5G and contribute to the new ecosystem.

A. Why TRIANGLE?

During recent years the research activity growth has resulted in several wireless testbeds. But although the mobile traffic estimates grow year after year it has not reflected in the number of FIRE wireless testbed users. The main reason for this situation is the design of the FIRE testbeds themselves, which are network

centric. FIRE testbeds are normally focused on network configuration and have very complex and sophisticated configuration mechanisms, while the experimenters are not familiarized with the complex setup of the network resources and most of the time end up using only the default configuration. Based on previous experience running the federation of PerformLTE testbed [5], it has been identified that testbed users spend most of their time gaining access to all the low level parameters that impact the user traffic transport performance as they lack enough know-how on setting these parameters to generate a consistent experimentation scenario. This can be frustrating for them because it prevents them from obtaining quick insights on the performance of their products. This is particularly critical for power consumption tests, because an erroneous network configuration can negatively impact the performance of the device [6].

The main idea underpinning the methodology to be used in the design and development of the TRIANGLE test framework is to ensure that the end user is not overwhelmed by the complexity of the overall testbed by not being exposed to all the details, which requires multi-disciplinary knowledge (protocols, radio propagation, software, etc.) to fully understand. To overcome the need for multi-disciplinary knowledge, TRIANGLE will ensure an abstraction of the underlying networking technologies by offering:

- a high level configuration layer (personality) which calls on detailed scenario definition; and
- a flexible framework architecture to incorporate new 5G networking topologies as they become available

The project will design a set of high level configurable scenarios that can be reproduced and whose final testing output is a behaviour score or TRIANGLE-MARK, which defines how well the product (application or device) behaves when used in a

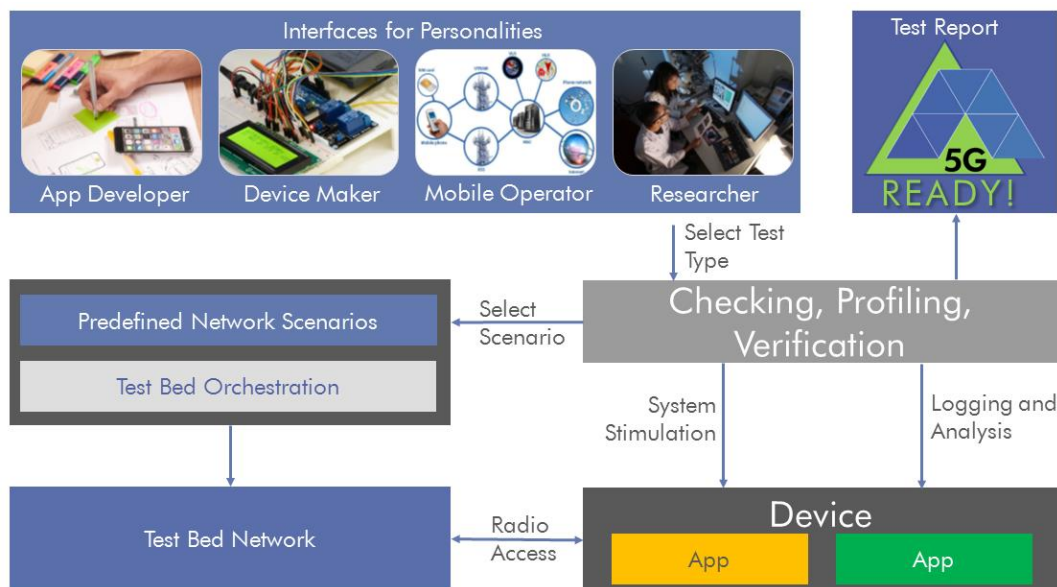


Figure 1 - The TRIANGLE approach for providing a unique certification grade

realistic network, including energy consumption and model-based runtime checking of the apps and devices. It will be possible to modify and setup very complex scenarios by means of a scenario editor, with an Application Protocol Interface (API) and a Graphical User Interface (GUI). This scenario editor will be able to load pre-defined scenarios and generate scenarios based on 5G use cases covering KPIs of interest for each one of them.

The project will provide a framework with different layers of abstraction using, when possible, commercial configuration interfaces as well as experimentation standards, including those provided by Fed4FIRE [7]. Where needed, advanced users will be exposed to deeper configuration details and flexibility.

B. Benchmarking-as-a-Service

Being ready for the 5G rollout in 2020 means in part understanding the technical behaviour of a new 5G product or application as well as the network capabilities. TRIANGLE can provide Benchmark-as-a-Service support to help new entrepreneurs and established suppliers 'tool-up' for 5G. Using the TRIANGLE framework a device or application developer can tune the parameters of its system to optimize performance under different conditions. Such conditions could include network load or other factors affecting radio power that could be emulated on the TRIANGLE testbed.

TRIANGLE can enable comparative studies of behaviour in different conditions. Profiles can be established and system behaviour can be measured, analysed and compared to reference applications on the TRIANGLE testbed. This gives a benchmark and an understanding of product position to the tester. Such comparative studies could help device or application development for the 5G market. New product entrants are often in a race to be early into the market to get an established position and early customer wins. TRIANGLE can help with this process and reduce the risk in bringing new products to market.

C. The Triangle-Mark

TRIANGLE offers a TRIANGLE-MARK as a pre-standards pre-certification tool providing a relevant score based on a set of indicators. The TRIANGLE-MARK ensures that a mobile application or device operates in the future 5G mobile broadband networks with a guaranteed QoE by assessing and comparing an application or device capability in a 5G network emulation environment against a set of criteria or KPIs.

As an example, an application with the TRIANGLE-MARK, will have been tested in different hardware platforms (as defined by the TRIANGLE Testbed) to verify very different performance indicators such as stability, security, reliability or the performance, both in terms of time behaviour and in the use of resources. The application will get a final score that allows the users to compare its performance with that of similar applications in the market

When a product gets the TRIANGLE-MARK, it will interwork and perform satisfactorily in the growing ecosystem of 5G networks, devices, services and applications.

D. The TRIANGLE Approach [KEYD]

TRIANGLE intends to address the needs of four test bed user profiles within a single test framework: an advanced Researcher,

a Device Manufacturer, an App Developer, and a Mobile Network Operator (MNO). The project's testing approach is portrayed in Figure 1.

1) Researcher Profile

The most extended personality on FIRE wireless testbeds today is the Researcher personality, especially researchers who are working on protocols and network technologies. This personality is interesting but only represents a small group of possible testbed users. These users will be looking for the ability to fine tune low level parameters; test adaptation within the network layers; create new use case scenarios, etc.; in general to verify the functionality of lower communication layers.

2) Device Manufacturers Profile

Device manufacturers will have a more general system performance level interest in the testbed to understand the capability of the overall device compared to other reference

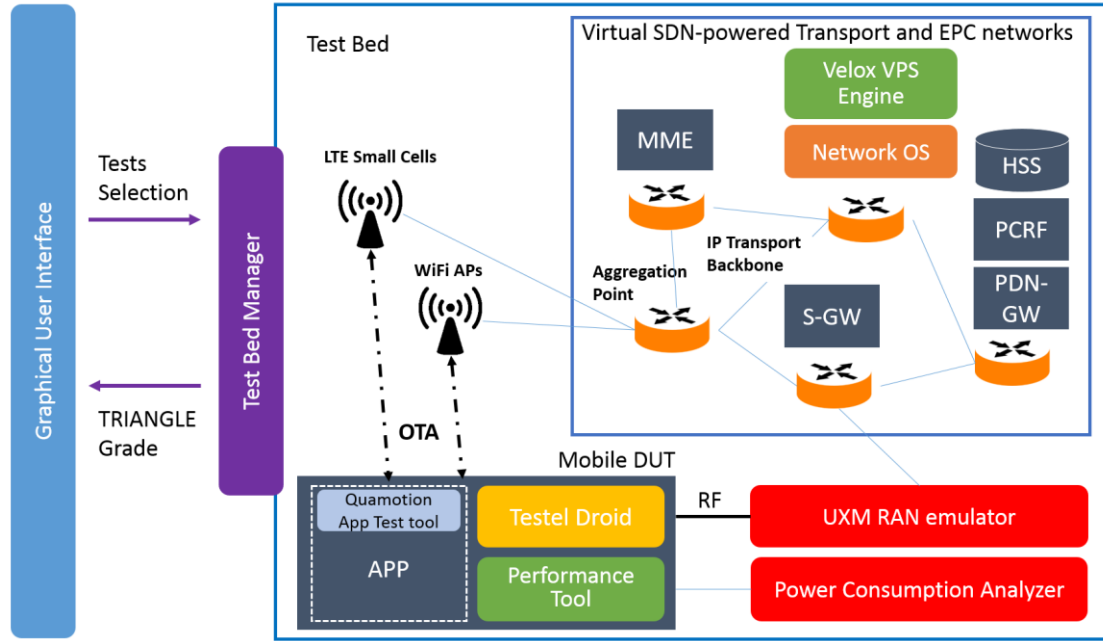


Figure 2 - The TRIANGLE test bed architecture

devices. Today, the number of device manufacturers is rather limited but, with the upcoming Internet of Things (IoT) and device variety, the numbers will grow rapidly. The skill and knowledge of these device makers is also changing from being a fully integrated in-house developer (with all components in house) towards a more system level integrator type that relies on modules and chipsets from external suppliers. This type of outsourcing, combined with the strict energy requirements of IoT, make TRIANGLE a powerful ally for understanding the real capabilities of the devices in terms of energy saving.

3) App Developer Profile

These pure software personalities are the largest target population for TRIANGLE. These are usually start-up companies who wish to develop an app based on a great idea and need platforms to validate its functionality and its performance under challenging radio conditions. The app developer, not being a telecommunication engineer, typically does not care about the network details and radio propagation challenges. The developer typically assumes the wireless connection is a reliable data pipe with characteristics similar to a wired one. TRIANGLE will provide a framework to test the impact of realistic data network capability on the app as well as the app impact on the network resource usage, making possible to provide insight on how the App uses the device resources and battery. Developers will be able to have clear insights on the energy performance of their own code.

4) Mobile Network Operator Profile

While primary targets for testing and certification are the device manufacturers and App developers, the project also wants to provide MNOs with a fully controlled, realistic sandbox where to test advanced network components and configurations. This type of setup is not new to MNOs. As a matter of fact some major operators have similar testbeds in their labs. But although there is wide interest among operators to understand the impact

of network configurations on the performance of mobile devices and the QoE of mobile apps, only few major actors can afford it.

E. Open to experimenters

The test framework developed within the project will be accessible to the developer community, initially to the European community and internationally later on. During the project, access to the test framework will be organized via open calls, where TRIANGLE will seed 15 Open Call Experiments focused on applications and devices mostly from innovative SMEs and entrepreneurs. These Open Call Experiments will be used to fuel the market for benchmarking and certification towards 5G. The TRIANGLE-MARK will be given with the benchmark measurements or score obtained.

The Open Calls will be used also for improving the testbed capabilities with advanced technologies that are relevant towards the 5G evolution. The project will seed at least two to four technology providers that will include their products in the testbed, making them available to other experimenters.

III. TRIANGLE TEST BED

The main core of the TRIANGLE project is the testbed that will allow all the different users to run tests in a controlled environment that encompasses all the elements of the telco chain. An overview of the main elements that compose the test bed is shown in Figure 2.

A. Mobile Device Monitoring

In order to measure the KPIs perceived by the end device and the end user, several tools are needed on the smartphones to be tested. Such Monitoring tools are being internally developed by TRIANGLE partners and will be available to experimenters.

1) Performance Tool

5G test scenarios will require high resolution for reporting target QoS KPIs. In order to extend the measurement capabilities of other tools like iperf, TRIANGLE will provide up to OSI layer 7 SDU packet resolution in the computation of data performance KPI thanks to the integration of the AT4 wireless Performance Tool. This tool is composed of two components, Controller and Agents (data endpoints), and uses proprietary mechanisms to synchronize the Agents and provide accurate one-way measurements. This tool includes a built-in traffic generator with the capability of generating constant rates, ramps, loops and statistical traffic patterns which is something of utmost importance for setting up the desired environment in terms of varying traffic loads (e.g., for measuring LTE-U impact on Wi-Fi networks). Additionally, this tool has the ability to automate some mobile Apps on Android devices and measure relevant QoE KPI such as YouTube buffering occurrences.

2) *TestelDroid Mobile Monitoring App*

TestelDroid [8] is a software tool that enables passive monitoring of radio parameters and data traffic in Android-based devices. Logging is implemented as an Android service that can be running in the background logging all the information while the application under test is being executed. This functionality enables monitoring of the traffic information generated by any application, which extends the testing to a very wide range of use cases. The parameters to be logged (network, neighbour cells, GPS, traffic) can be flexibly configured using the SCPI interface.

3) *Power Consumption Analyser*

In order to provide a complete overview of the performance of the App or the Device under test, it is nowadays unthinkable not to include an analysis of the energy consumed by the mobile platform. For this reason TRIANGLE plans to provide accurate measurement consumptions thanks to the Keysight N6705B Power Analyser. The instrument allows the provision of DC voltage to the device and accurately measures the drawn current to verify how much power is required by each app and network operation over time.

B. *Network Emulation*

The test bed also comprises a network emulation that spans from the Radio Access Network (RAN) to the core of the LTE system. All the available tools are orchestrated as well, to provide the highest degree of experimentation flexibility to the test bed user.

1) *UXM RAN emulator*

An essential component in the testbed is an instrument capable of emulating multiple cellular networks in a controlled manner. To that end, TRIANGLE envisions the usage of the UXM Wireless Test Platform device by Keysight Technologies, which supports multiple radio access technologies (multi-RAT), including GSM/GPRS, UMTS and LTE-Advanced networks (i.e., 2G, 3G and 4G). The UXM features include intra-RAT and inter-RAT handovers, protocol debugging, IP end-to-end delay and throughput measurements, and performing RF conformance tests. Finally, it should be noted that the UXM also features an advanced fading engine with the main channels models defined by 3GPP [9].

2) *Over-The-Air LTE Small Cells*

UMA lab currently offers an indoor deployment of three LTE 2,6GHz FDD cells, one pico-cell by Alcatel-Lucent and two small cells from Athena wireless. In the context of TRIANGLE this deployment will be updated with new small cells supporting more flexible configurations, including Carrier Aggregation. The use of attenuators and switches will allow controlled experimentation in this commercial based setup, including intra-LTE handovers and LTE-Wi-Fi offloading. Scenarios with several devices can be also reproduced if required to test some apps.

3) *EPC Emulator*

The Enhanced Packet Core (EPC) is the core of LTE networks, and it will also play a central role in pre-5G deployments that include for instance LTE to Wi-Fi offloading. The EPC is a component with functions like controlling which devices are allowed to attach to the network, supporting mobility (handover) or routing traffic to/from the Internet. The UMX emulator already integrates a subset of this functionality. UMA lab brings Polaris Networks EPC to the project. Thus, the Project will support some scenarios with small cells, that provide real but non fully controllable radio propagation conditions, and others based on the UXM emulator, which provides fully controllable radio propagation conditions, both connected to Polaris EPC through standard 3GPP interfaces.

4) *Virtual Path Slicing Engine*

Software Defined Networks (SDN), Network Function Virtualization (NFV) [10] and OpenFlow [11] are techniques that provide high flexibility and can cut costs in packet data networks. Virtual Path Slicing (VPS) is a method that has evolved from studies like [12] on how to implement end to end QoS in packet networks. RedZinc has implemented a VPS engine that can steer traffic on the edge of the network based on the request of the application and coordinate this traffic steering with other autonomous systems. In traditional 3GPP EPC the VPS engine can provide services to the application layer and invoke them using the Rx interface in the Policy Charging Rules Function (PCRF) of the EPC where differentiated QoS Class Identifier can be used for different service levels.

C. *App testing*

In order to test the apps running on the mobile devices thoroughly, the user interaction needs to be taken into account. TRIANGLE takes advantage of several approaches in this respect, considering both real and simulated user interactions. These approaches are complementary and result in a greater coverage of tests and app-related KPIs. These tools are of particular importance to provide a realistic usage pattern of the apps whenever power consumption is measured.

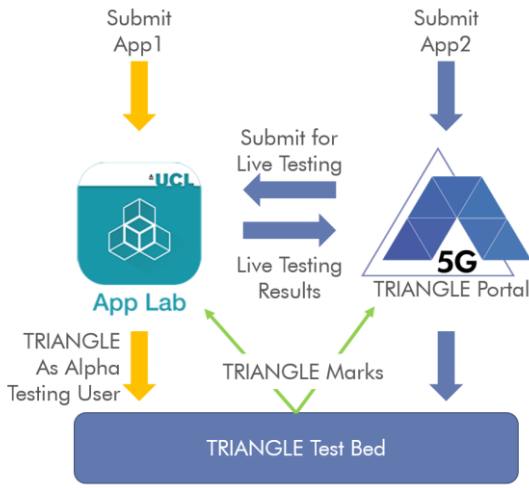


Figure 3 - Live Testing Interaction via App Lab

1) App Lab live environment

TRIANGLE will use App Lab, a high level platform developed by UCL for distributing mobile apps to a large- scale testbed for pre-commercial testing and validation services. App Lab provides a private app store offering in-the-wild rapid user field-testing to UCL staff and students. App Lab is comprised of a server portal for application upload, management and distribution and native mobile client apps for both Android and iOS devices. App Lab uses analytics tools to track how users interact with apps in the real world. By examining the results of app usage, test cases which mimic actual users' behaviour can be derived and executed in a variety of network conditions in the TRIANGLE testbed as shown in Figure 3.

2) Quamotion test automation tool

TRIANGLE will also use Quamotion test automation tool which provides the ability to write and execute automated tests for Android and iOS applications. Execution can be performed without any manual interaction. Not only gestures and keyboard interactions but also the full lifecycle of the application can be automated, i.e. installing, launching, terminating. This set of actions are exposed to the overall test orchestrator such that the end-user actions can be synchronized with various network events, e.g. an unstable network connection in a train. Quamotion will also add recording capability to increase the efficiency in the process of writing test cases. Each step performed on the device is recorded and translated automatically into an action as part of a test script.

3) Model Based Testing

Another alternative considered in the project is model-based testing. Instead of writing test scripts manually with sequences of user actions, e.g. pressing a button, entering some text, app developers can provide a high-level description of the intended workflow. Test cases could be automatically generated from this description, achieving a high coverage of realistic usage scenarios. These models may also be based on crowdsourced data, such as observed user interactions from live tests, including tests run using App Lab.

IV. TEST BED ORCHESTRATION AND MANAGEMENT

The orchestration and management architecture of the testbed will be based on OMF (Orbit Management Framework) [13]. OMF allows the definition of repeatable and automatable experiments using an OEDL definition language. OEDL is a domain-specific language defined for the description of an experiment's execution. OEDL provides a set of experiment-oriented commands and statements which can be used to define the tests, the measurements and the graphical results. These

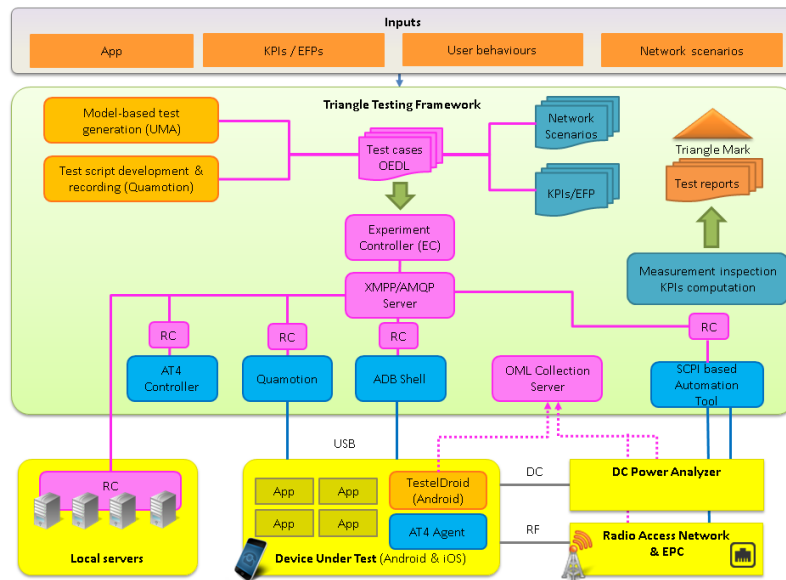


Figure 4 - The TRIANGLE test bed orchestration system [New figure]

OEDL scripts are interpreted by the Experiment Controller (EC), which orchestrates the resources present in the testbed during the execution of experiments. Each resource in the testbed is managed by a Resource Controller (RC), which will translate the actions defined in the experiment into the specific control commands understood by the resources. Measurement data during experiment execution will be collected via OML in a central server. OML provides a programming library for easy application instrumentation and a collection point, a server which stores measurements in an experiment database.

TRIANGLE testbed orchestration based on OMF is shown in Figure 4, starting with the definition of the app test cases by the app developer in the top left and the network scenarios on the top right. Once the test cases and the network scenarios, have been completely defined, they can be executed in the testbed. The RC associated with the device under test interacts directly with the Android and iOS devices over USB in order to install the apps and run the app test cases. In order to coordinate the execution of the different tools involved in the testbed an RC for each one of them will be deployed. These RCs allow the tools to be controlled as part of the test, and receive commands from the experiment controller to fetch and execute a particular test case. During the tests, the monitoring and performance tools will collect all the measurement results from all layers and measurement points present in the testbed. The results will then be passed to the Measurement inspection and KPI calculation module, which outputs the final test report.

V. EXPERIMENTS SETUP

The components provided by some partners have already been integrated into a first version of the TRIANGLE testbed. As an exercise the testbed has been used to evaluate the performance and the power consumption of a YouTube video playing on a mobile device, under different radio channel conditions in a single user scenario. An agent that drives the YouTube API was installed on a Samsung Galaxy S4 connected to the UXM LTE network emulator via a coaxial cable and from there to the YouTube server via internet. The performance tool was controlled remotely to run the agent and record KPIs using the algorithms proposed in [14]. The video used has a length of 5 minutes and resolutions ranging from 1080p to tiny. Power consumption is measured using the DC power analyzer which is directly connected to the power terminals of the mobile. The beginning of the YouTube session is synchronized with the power consumption capture. Once the power consumption measurement is initiated it lasts 120 seconds, independently of the duration of the playback of the video. To determine the total consumption the power measurements have to be extrapolated considering the total playback time.

The scheduler used during the experiments only transmits LTE PDUs to the mobile device (in the downlink) when there is IP data to carry. The UE transmission power is controlled automatically to be kept at 0 dBm in all the test scenarios for better analysis of the impact of the DL test configurations.

In the frequency domain, the full channel bandwidth has been assigned for downlink transmission. This translates into 100 Physical Resource Blocks (PRBs) in a 20 MHz cell.

Different levels of Additive White Gaussian Noise (AWGN) have been also added when required to model interference from neighbor cells transmitting at the same frequency. The Signal to Noise Ratio (SNR) is determined by the relation between the power density of the transmitting cell and the interfering noise.

The downlink transmissions follow an open-loop transmit diversity (TM3) MIMO scheme, and the modulation and coding schemes are dynamically adapted to the channel quality and rank reported periodically by the UE.

Channel models emulate the effects of the environment on the radio signal: fading, delay spread and Doppler Effect.

Two different scenarios have been considered: A first scenario which emulates the behavior of a channel while the mobile user is walking in a city (EPA5, 20 dB SNR) and a second scenario which represents a mobile user travelling by bus in a city (EVA5, 10 dB SNR).

The configuration of the experiment includes 4G LTE cells with two transmitting antennas, resulting in a 2x2 MIMO matrix as all LTE devices have minimum two receiving antennas. The user device reports the observed radio channel quality (CQI) and the rank (RI), the later representing the number of independent data flows that can be transmitted on the channel.

VI. RESULTS

The YouTube traffic has a bursty nature. Figure 5 depicts the downlink (DL) traffic captured by TestelDroid during one of the test executions.

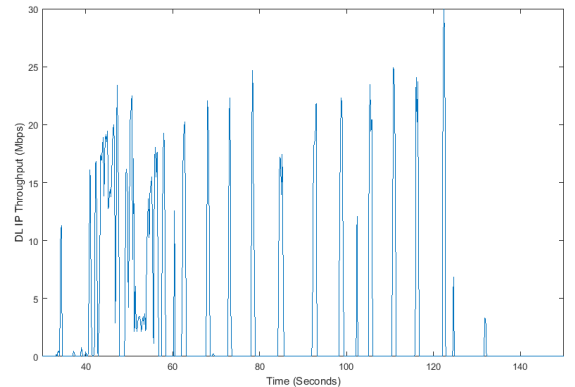


Figure 5 DL IP throughput during a YouTube session

At the beginning of the session, a larger throughput is observed as the application seeks to gather enough data before starting the playback.

Figure 6 displays the CQI and RI histogram recorded during the experiments under the pedestrian scenario. CQI is the mechanism used in LTE by the UE to report what modulation and coding scheme the eNodeB should use to achieve a nominal block error rate in the order of 10% at the link level for initial transmissions as stated in 3GPP TS36213. In terms of RI, both rank 1 (CW0) and rank 2 (CW1) were reported by the UE although it can be observed that rank 1 was the most frequently reported value. The transmission will most of the time just use

one of the flows, which implies lower capacity but less interference on the second flow.

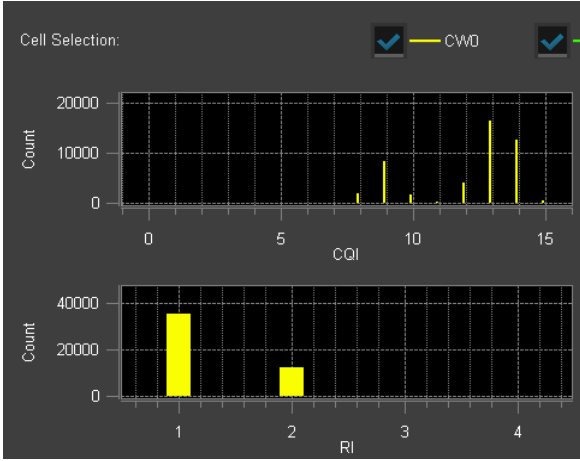


Figure 6 Channel quality and rank reported information in a pedestrian scenario

Table 1 for the pedestrian scenario and table 2 for the vehicular scenario display the number of transmissions successfully acknowledged (ACK) in the first transmission (1st) and, when they occurred, in the retransmissions (2nd, 3rd, 4th). Table 1 differentiates the acknowledged messages for rank 1 (CW0) and rank 2 (CW1).

Table 1 DL retransmissions in a pedestrian environment

	% ACK CW0	% NACK CW0	% ACK CW1	% NACK CW1
1st	87.02	12.97	91.64	8.36
2nd	96.00	1.77	78.72	6.91
3rd	13.95	32.56	10.00	35.00
4th	5.41	43.24	5.56	44.44

In the pedestrian scenario the first transmission has a high success (87,76%), the second transmission (i.e., the first retransmission), has also a high success rate (96%), but in the third and fourth transmissions only a small fraction of the data manages to get through. This is a symptom of bad channel conditions, where the upper layers (RLC and TCP) shall be in charge of ensuring packet delivery without loss of information by increasing transmission at their corresponding levels.

In the vehicular scenario, the selected signal to noise ratio is 10 dB lower than in the pedestrian. This has been chosen to represent the lower strength to be expected in the vehicle. It also has a higher delay spread which is aggressive in terms of multipath. Both factors result in worse propagation conditions being reported by the mobile device under test. As shown in Figure 6, all the recorded rank indications include requests for single layer (non-MIMO) transmissions (rank 1). Additionally, the reported quality (CQI) is lower than in the pedestrian case.

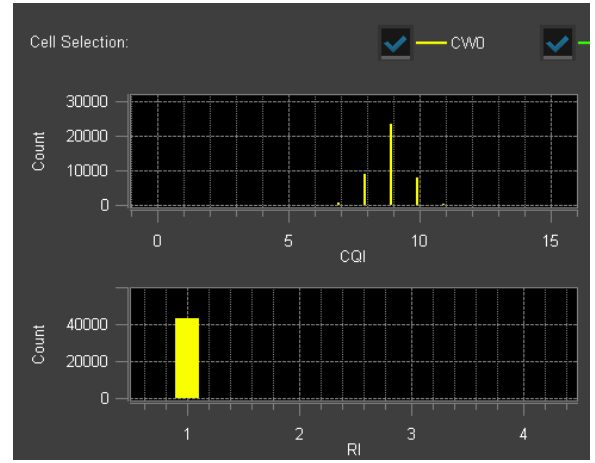


Figure 7 Channel quality and rank reported information in a vehicular scenario

Table 2 DL retransmissions in a vehicular environment

	% ACK CW0	% NACK CW1
1 st	90.83	9.17
2 nd	100	0
3 rd	0	0
4 th	0	0

The different radio channel conditions impact the final QoE of the YouTube sessions. Figure 7 provides a summary of KPIs collected by the Performance Tool and a QoE score according to the method described in [14]. With only small differences in the initial buffering time, low number of re-bufferings (0) and zero re-buffering time (0) the resulting MOS score is high (>3.5). The video quality, not considered in the algorithm used to evaluate the MOS, also remains constant over time (see Figure 8).

Table 3 YouTube KPIs in single user scenarios

	EPA	EVA
Initial buffering	3.551	2.908
Playback Size (MB)	46.067	45.979
Playback Time (s)	160.638	160.982
MOS	3.89	3.89
% time in 1080p	100	100

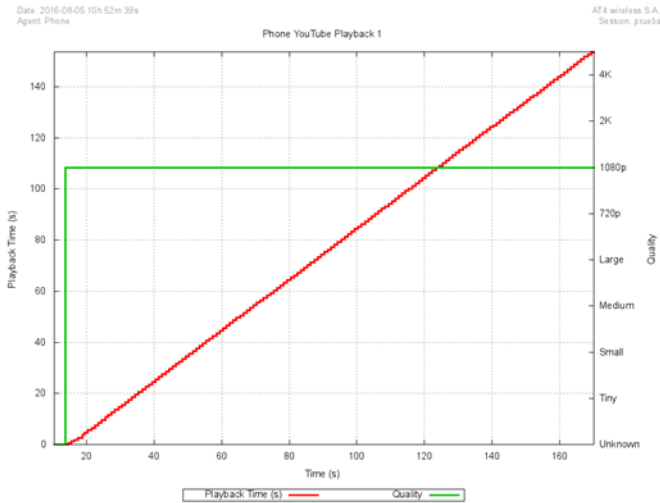


Figure 8 Playback time and video quality versus time

In terms of power consumption, only moderate differences in power consumption have been found. Figure 9 displays the averaged power consumption over time, with a sampling period of 100 ms. Only during the 10 initial seconds, averaging during 5 different sessions per test condition, the power consumption has been observed to be 11% higher in the vehicular scenario than in the pedestrian scenario (as shown in Figure 10). After the initial buffering and processing peak, the power is observed to be similar in both scenarios.

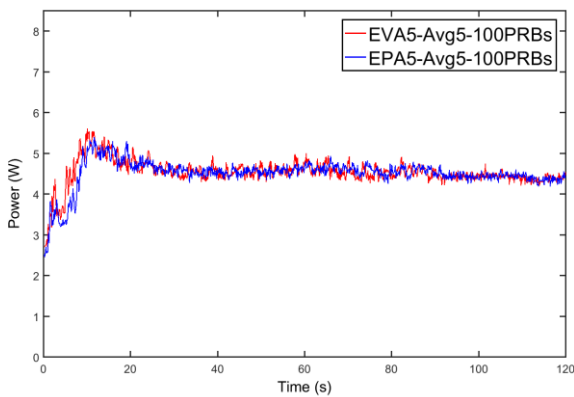


Figure 9 Power consumption over time

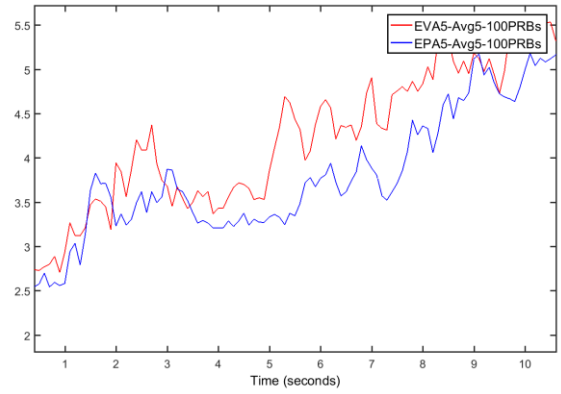


Figure 10 Power consumption during initial buffering in single user scenario

To provide more realistic insight, the experiments have been repeated allocating only 8 PRBs. This resembles a dense environment where the resources of the cell are shared with other users during the session. Under these more constrained conditions, remarkable differences between both scenarios can now be observed.

In the pedestrian scenario, the available capacity does not allow to keep the full 1080p resolution, and YouTube automatically changes the video quality to 720p (Figure 11). In the vehicular scenario, where the radio channel conditions are tougher, YouTube service reacts dropping the video quality down into the large format early in the playback (Figure 12).

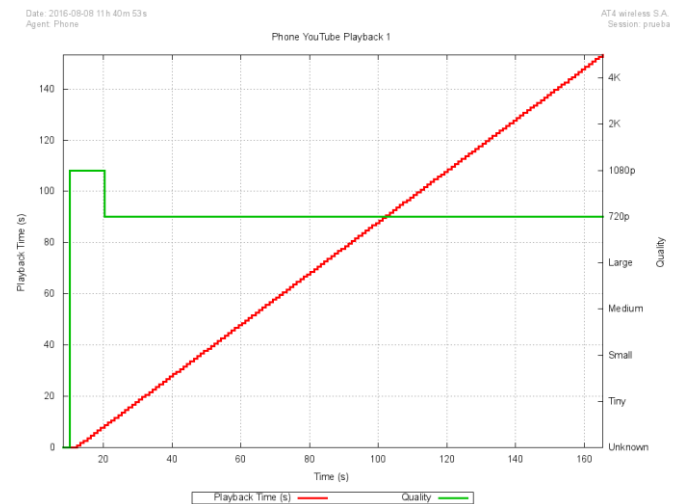


Figure 11 Playback time and video quality versus time in a dense pedestrian scenario

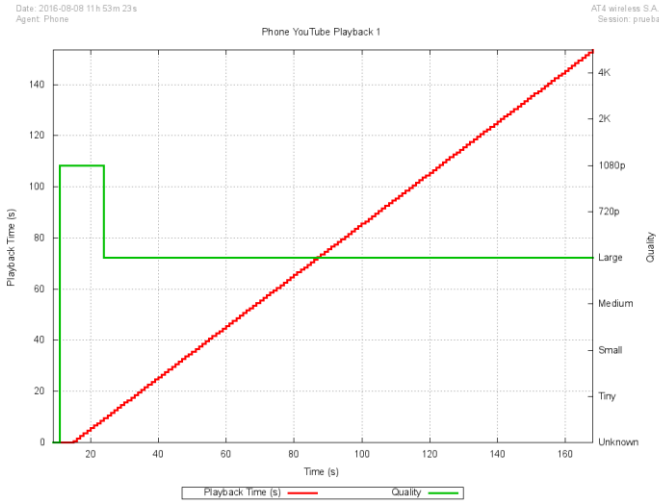


Figure 12 Playback time and video quality versus time in a dense vehicular scenario

Figure 13 shows the power consumption over time. An increment in the power in the pedestrian scenario can be observed. This is caused by the data being fragmented in a larger number of pieces when transmitted over the air because of the smaller amount of allocated PRBs. It must be noted that the mobile device has to process and transmit feedback in fewer occasions in the vehicular scenario because it has a lower resolution. Normalizing the power per Mbit of video the impact becomes clearer. The large resolution requires 17 MB to be transmitted, whereas the 720p video generates only 29 MB of data. Thus, in the pedestrian scenario, the video can be reproduced with the double amount of data at the expense of similar power consumption.

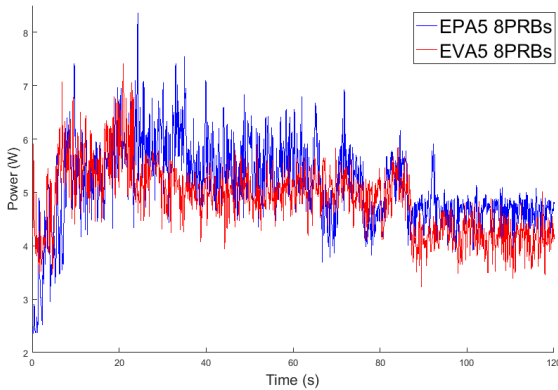


Figure 13 Power consumption in dense scenarios

Figure 13 shows the relevant KPIs and how the MOS is slightly reduced. Because the MOS score algorithm in [13] does not consider the video quality the project has identified an area of research to establish relevant KPIs in dynamic video quality adjustment conditions.

Table 4 YouTube KPIs in dense scenarios

	EPA5	EVA5
Initial buffering (s)	3.325	5.416

Playback Size (MB)	29.425	17.172
Playback time (s)	157.639	159.733
MOS	3.89	3.81
% time in		
1080p	5.23	5.84
720p	94.77	0
Large	0	94.16

1) Importance of KPIs alignment

The synchronization of the different KPIs is very important to enable further analysis and to infer dependencies between different factors. Having synchronized the power monitoring with the application sessions has allowed TRIANGLE to easily characterize the average behaviour of multiple YouTube sessions. Even finer details can be observed within individual sessions, as the instantaneous power consumption is highly correlated with data traffic burstiness.

VII. CONCLUSION

TRIANGLE is a research project that brings together SMEs, labs, test equipment vendors and academia, with the goal of providing an end-to-end testing ecosystem for Apps, devices, and services targeting the future 5G system. The test framework will be based on existing FIRE infrastructure and new elements for completing the ecosystem will be added to expand the testing possibilities offered by the test bed. In order to facilitate the exploitation of the test bed not only by researchers but also by App developers, device vendors, and other players in the telco market, the project will implement abstraction layers fitting the interest and knowledge of each profile. To show the potential of the test framework, some preliminary results of the impact of radio propagation conditions on a well-known service installed in one mobile have been presented.

ACKNOWLEDGMENT

The TRIANGLE project is funded by the European Union's Horizon 2020 research and innovation programme, grant agreement No 688712.

REFERENCES

- [1] "The Rise of the Enterprise in the Smartphone Wars," Goldman Sachs, December 3, 2013.
- [2] "Forecast Analysis: Mobile Phone Production and Semiconductors, Worldwide, 3Q13 Update," Gartner, October 28, 2013.
- [3] NGMN 5G White Paper, NGMN, Feb 2015.
- [4] TRIANGLE project official website www.triangle-project.eu
- [5] A. Díaz-Zayas, C. A. García-Pérez, Álvaro M. Recio-Pérez, and P. Merino, "PerformLTE: A testbed for LTE testing in the Future Internet," in The 13th International Conference on Wired and Wireless Internet Communications, Lecture Notes in Computer Science (LNCS) 9071, Springer, 2015, pp. 313–326.
- [6] A. T. Koc, S. C. Jha, R. Vannithamby and M. Torlak, "Device Power Saving and Latency Optimization in LTE-A Networks Through DRX Configuration," in IEEE Transactions on Wireless Communications, vol. 13, no. 5, pp. 2614-2625, May 2014.

- [7] Wauters, Tim, et al. "Federation of Internet experimentation facilities: architecture and implementation." European Conference on Networks and Communications (EuCNC 2014). 2014.
- [8] Álvarez, A. Díaz-Zayas, P. Merino, F. Javier Rivas Tocado "Field measurements of mobile services with Android Smartphones ," in Consumer Communications and Networking Conference (CCNC), 2012 IEEE , pp.105-109, 14-17 Jan. 2012
- [9] 3GPP TS 36.101. "User Equipment (UE) Radio Transmission and Reception." 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA).
- [10] ETSI Network Functions Virtualisations White Paper #3. October 14 2014
- [11] McKeown, Nick, et al. "OpenFlow: enabling innovation in campus networks." ACM SIGCOMM Computer Communication Review 38.2 (2008): 69-74.
- [12] Braun, Torsten, Michel Diaz, José Enríquez Gabeiras, and Thomas Staub. End-to-end quality of service over heterogeneous networks. Springer Science & Business Media, 2008.pending
- [13] T. Rakotoarivelo, M. Ott, G. Jourjon, and I. Seskar, "OMF: a control and management framework for networking testbeds," ACM SIGOPS Operating Systems Review, vol. 43, no. 4, pp. 54–59, 2010.
- [14] F. Lozano, et al., "Network Performance Testing System Integrating Models for Automatic QoE Evaluation of Popular Services: YouTube and Facebook", Wireless Personal Communications, Springer US (DOI = 10.1007/s11277-015-2479-y).