FLEX-MONROE: A Unified Platform for Experiments under Controlled and Operational LTE Settings

Mah-Rukh Fida[1], Konstantinos Kousias[2], Andra Lutu[2], Mohammad Rajiullah[3], Özgü Alay[2], Anna Brunstrom[3], Antonios Argyriou[4]

[1] University of Edinburgh, UK [2] Simula Research Laboratory, Norway
[3] Karlstad University, Sweden [4] University of Thessaly, Greece

ABSTRACT

This paper presents FLEX-MONROE, a unique platform that facilitates achieving a thorough understanding of LTE networks, one that captures the status of current operational MBB networks and that also enables LTE performance improvements by allowing experimentation in an environment with controllable LTE parameters. Using this platform, we propose to investigate how variations in the LTE network parameters influence the network characteristics, which, in turn, translate to application performance metrics that represent the end-user experience. We argue that the FLEX-MONROE platform is crucial to understand, validate and ultimately improve how current operational MBB networks perform, towards providing guidelines for designing future 5G architectures. Furthermore, understanding the effects of low-level tweaks in network parameters in the LTE infrastructure on the application performance is critical to provide guidelines on how to improve the application performance in the current but also future MBB networks.

1 INTRODUCTION

Open experimentation with operational mobile broadband (MBB) networks is currently a fundamental requirement of the research community in its strive to innovate mobile communications. At the time of writing, despite many individual efforts to offer access to infrastructure and technologies [7, 14], there is still a compelling need for building a truly open MBB testbed that can offer the service and experience of an operational commercial network with the unique perk of controlling different elements in the network. In this paper, we attempt to address this very ambitious goal. We introduce FLEX-MONROE, a unique platform for open experimentation with LTE networks both in the wild with commercial operational networks with MONROE [3] and in controlled settings with configurable equipment with FLEX [17].

We present our approach for the integration of two measurement platforms (i.e., FLEX and MONROE) to build a one-of-a-kind testbed, FLEX-MONROE. We designed MONROE to advance our understanding of today's operational MBB ecosystem from the end-user's perspective. Due to the lack of control of the experimenter over commercial networks, MONROE cannot extract and, more importantly, validate the root-cause of MBB networks performance problems, especially the ones related to the network configuration and characteristics. In order to shed more light on the effect of different network-specific parameters on the application performance, there is a dire need for complementing the measurements we run in the wild with similar measurements in an experimental LTE network. FLEX offers the ideal solution for experimentation with LTE networks by providing a combination of configurable commercial equipment, core network software and open source components. The FLEX testbeds enable the identification of how different LTE network parameters impact the performance of services and applications over a LTE infrastructure. The FLEX-MONROE integration results in a platform that captures the status of current operational MBB networks and that also enables LTE performance exploration through experimentation in an environment with controllable LTE parameters.

To demonstrate the benefits of FLEX-MONROE, we show the impact of twisting the LTE network parameters on the network quality of service characteristics, which, in turn, translate to application performance metrics that impact the end-user experience. We use the particular case of web browsing to showcase the potential of performing such analysis in FLEX-MONROE. We explore the mapping between web browsing performance metrics and the set of features that capture the network characteristics of operational LTE

WiNTECH'17, October 20, 2017, Snowbird, UT, USA.
 ISBN 978-1-4503-5147-8/17/10…$15.00
DOI: http://dx.doi.org/10.1145/3131473.3131477

networks we measure in MONROE. We also check a similar mapping between the LTE network parameters and the networks characteristics using FLEX. The merged analysis captures how the underlying network parameters in the LTE infrastructure ripple through all the layers until the end-user and impact the web browsing performance.

2 PLATFORMS OVERVIEW

All major mobile network operators are in the process of adopting LTE/4G cellular networks. These are expected to rule the cellular landscape at least for the current decade, while also forming the starting point for further progress, ushering in the 5G era. The lack of open or at least openly configurable cellular equipment and core network software is a major limiting factor for applied research in this field outside of the boundaries of vendor and operator R&D groups. MONROE and FLEX are two approaches that aim to address this gap. In the following, we give a brief description of these two measurement platforms.

2.1 FLEX

The FIRE initiative targets the creation of a multidisciplinary research platform for investigation and experimental evaluation of innovative ideas in networking and services. FLEX (FIRE LTE Testbeds for Open Experimentation) contributes a crucial missing piece in FIRE's infrastructure puzzle: configurable cellular access technologies and LTE in particular. FLEX builds programmable LTE components as extensions to existing European testbeds, thus providing an open and remotely accessible platform for experimentation with LTE. The deployment of LTE components took place on three existing and well-established FIRE wireless testbeds, thus enabling these facilities with novel capabilities for LTE experimentation-driven research. Among different facillities, in this paper we focus on the NITOS [17] indoor testbed. The integration with other facilities is part of our ongoing effort.

NITOS indoor tesbed is a fully controllable RF-isolated environment consisting of over 60 operational wireless nodes offering for experimentation on various technologies, including Wi-Fi and LTE. To this extent, an LTE core infrastructure resides in NITOS with two LTE Base Stations (eNBs) and the SIRRAN's LTEnet packet core network. As for now, a third of the available nodes are equipped with LTE dongles and can be scheduled for experimentation with the LTE network.

Limitations: Multiple users can experiment with NITOS, at the same time, as long as they access different resources (i.e., one single user can reserve a node in the platform at a time). For a single experimenter, this might translate into unplanned congestion, making his experimental results unexpected.

2.2 MONROE

The MONROE (Measuring Mobile Broadband Networks in Europe) system is the first open access platform for independent, multihomed and large-scale experimentation on commercial cellular operators. MONROE integrates 150 hardware devices (MONROE nodes) and a software framework that enables the orchestration of experiments and the collection, analysis, visualization and sharing of measurements that run on each hardware device. MONROE allows authenticated external users to access the platform, reserve resources and deploy their own experiments, which run in isolation to minimize bias. Aside from this, MONROE offers a series of ready-to-use experiments[1] offered as Experimentation as a Service (EaaS). Researchers, and other external users, can run measurements according to a pre-defined data quota to guarantee fairness. The software components and the architecture design are open sourced to foster community maintenance and deployability [1, 3].

Each MONROE node is a Linux-based programmable device that is multihomed to three Mobile Broadband operators thanks to a dedicated 3G/4G modem (LTE CAT6) for each carrier [24]. The nodes are deployed in heterogeneous environments including mobile (e.g., public transport vehicles) and stationary ones (e.g., volunteers hosting nodes in their homes). The nodes upload experimental results, together with the node metadata (e.g., geo-location, carrier, low-level aspects of the radio link) to a central back-end for public release. MONROE enables repeatability of measurements in similar conditions (specific uniform hardware and the same conditions/context for the stationary nodes).

Limitations: Though offering access to experiment with 12 mobile carriers operating in 4 European countries, MONROE offers a limited number of measurement vantage points per operator and per country. Thus, experimenters can perform comparisons of operators' services and analysis of measurement results only within the constraints of the geographic footprint of the platform. Moreover, whenever measuring a MBB operator, MONROE uses commercial-grade mobile subscriptions that are compatible with the ones customers can purchase. The differences in commercial offers thus reflect in the different data plans we activate in MONROE and, subsequently, in the datasets we can collect. For example, the number of samples we can collect from measurements on Orange (ES) (10GB) is much smaller than the number of samples we can collect on Telia (SE) (200GB).

3 FLEX-MONROE INTEGRATION

In this section, we give a detailed description of the FLEX-MONROE platform and the configuration involved for the integration of the two testbeds together. We also outline

[1] https://github.com/MONROE-PROJECT/Experiments

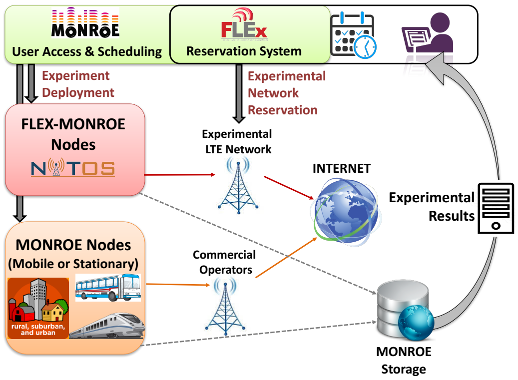

Figure 1: High level design of the FLEX-MONROE platform. The end user needs to connect to both the MONROE User Access and Scheduling and the FLEX NITOS Reservation System. After reserving the LTE experimental network via NITOS, the end-user can deploy experiments on the FLEX-MONROE nodes using the MONROE user interface in the same manner as deploying on normal MONROE nodes connected to commercial operators. The Experimental Results are transferred to the MONROE Storage solution, from where they are linked back to the experimenter via the MONROE User Access web interface.

the workflow an experimenter needs to follow for running measurements in FLEX-MONROE.

3.1 FLEX-MONROE Platform

The base idea for achieving the FLEX-MONROE integration is to have a MONROE node operate in the FLEX premises[2] and run the available MONROE measurements while attaching to the experimental LTE networks. In other words, our aim is to measure the experimental LTE networks in FLEX in the same manner as we measure any other commercial network in MONROE. This allows for a unified approach for running experiments both in commercial and experimental networks. We show in Figure 1 the high-level FLEX-MONROE system architecture.

The FLEX-MONROE node can be either the hardware native MONROE node running the MONROE software (hardware integration) or can be a hardware native FLEX node running the MONROE node software (software integration). For hardware integration, we installed within the FLEX NITOS testbed two physical MONROE nodes equipped with custom FLEX SIM cards that enables communication with the NITOS LTE network. For software integration, the MONROE node software image is installed in any compatible NITOS native node, following the normal steps for experimenting with the NITOS testbed[3]. This integration required support in the MONROE system, of the UEs currently deployed in the FLEX NITOS testbed, e.g Huawei E392 and ZTE MF831.

[2]We further refer to these measurement nodes as FLEX-MONROE nodes

[3]http://nitlab.inf.uth.gr/doc/load_saveOMF.html

Once activated, the FLEX node running the MONROE software image becomes available in the MONROE system just as any other native MONROE node. In both options for integration, the resulting FLEX-MONROE nodes connect to the NITOS experimental LTE network as they would connect to any other commercial LTE network within the MONROE system. The nodes run the native MONROE software without any customization and communicate with the back-end infrastructure in the MONROE system, similar to any native MONROE node. The FLEX-MONROE nodes (once connected to the NITOS LTE network) are thus capable to run any experiments MONROE offers as a service to the community and run these experiments in isolation while controlling the FLEX LTE network (we further exemplify this in Section 4). Apart from the existing MONROE EaaS, the experimenter has the freedom to design, implement and run its custom measurements on FLEX-MONROE.

In order to interact with the platform, the FLEX-MONROE user needs to access both the standalone MONROE system[4], as well as the standalone FLEX NITOS system[5]. Then, a normal workflow in order to deploy experiments on FLEX-MONROE requires the synchronization of the reservations in both systems and follows a number of steps, which we explain below.

3.2 Resource Reservation

Running an experiment in FLEX-MONROE entails a preliminary phase where the user reserves the required resources. This is a two-step process as resources from both FLEX and MONROE testbeds need to be accessed (Figure 1). However, the only legal combination in terms of reservation steps requires the experimenter to first access FLEX http://nitos.inf.uth.gr/reservation and ensure that the LTE network is exclusively under his/her control. To ensure fair access to the FLEX resources, there is a maximum limit for the reservation period of 4 hours per day and per resource. Although theoretically possible, deploying experiments on a FLEX-MONROE nodes without previously verifying the status of the eNodeB is not recommended since the network resources can be shared with other users or the LTE network might not even be active. Within the FLEX reservation period, the experimenter can use the MONROE User Interface and Scheduling system for deploying measurements in FLEX-MONROE (further discussed in Section 3.4). The experimenter then inputs the duration of the experiment he/she means to deploy, which implicitly delimits the reservation period for the FLEX-MORNOE node. The experimenter can further coordinate these experiments with similar ones in

[4]https://www.monroe-system.eu/

[5]http://nitos.inf.uth.gr/

Table 1: LTE parameters that the experimenter can modify in the NITOS experimental network.

Parameter	Description	Range
DL BW	Downlink bandwidth	5/10MHz
UL BW	Uplink bandwidth	5/10MHz
Power	Signal transmit power	-15dbm to -26dBm
Tx Mode	Enabled antennas	1/2
MCS DL	Downlink MCS profile	0-28
MCS UL	Uplink MCS profile	0-26
FQ Band	LTE band	7/13

the MONROE platform, for comparison with operations in commercial networks.

3.3 NITOS Configuration

Assuming the experimenter successfully completed the reservation phase in FLEX-MONROE, the next step is to access and configure the NITOS LTE experimental network for further experimentation. To access the LTE network, the experimenter first connects to the NITOS server using a certified NITOS slice[6]. Using the OMF Aggregate Manager service *LTErf*, the experimenter can then control the LTE base station parameters. A dedicated LTErf server runs on LTE base station that allows for configuration of the LTE parameters. The experimenter can set both the eNB and EPC to their default values to ensure that they are on a 'clean' state or recovery from possible changes that took place from prior users. After having the above steps completed, the experimenter can check the state of the SIRRAN's LTE core network. This entails verifying the connection status of the components comprising the architecture of the LTE core network (i.e., the Home Subscriber Server, the Packet GateWay, the Serving GateWay and the Mobility Management Entity). When all elements report the *connected* state, LTE network is ready for experimentation. We summarize in Table 1 a list of configurable parameters with a short description and the range of possible values. The experimenter can both query and replace the value of each parameter using the LTErf service. Also, one can both infer and repeat the configuration settings of commercial network in FLEX-MONROE under controlled settings. The value field should be restrained within the boundaries specified in the LTErf documentation. Restarting NITOS base station is required for the changes to take place. After having configured all the parameters, the experimenter can save the configuration for future use.

3.4 Experiment Deployment

As a final act, the FLEX-MONROE experimenter can use the MONROE User Access and Scheduling system to deploy a measurement campaign. MONROE uses Linux Docker containers to host and deploy a variety of user-specific experiments. As part of the EaaS initiative, MONROE users can already configure and deploy measurements on any MONROE node, both native MORNOE nodes or FLEX-MONROE nodes. The experimenter can also prepare its own custom experiments. All experimental results are stored within the MONROE back-end, using the same solution we offer to all MONROE authenticated users. From there, results link back to the user, who can download them by accessing at any time the MONROE user access and scheduling system.

[6]http://nitlab.inf.uth.gr/doc/lte.html

4 WEBWORKS IN FLEX-MONROE

In this section, we demonstrate how to use the FLEX-MONROE platform to run measurement campaigns using MONROE EaaS. In particular, we direct our attention to web browsing. The end-user experience depends on many factors dictated by the MBB ecosystem, including radio network parameters (e.g., MCS and transmission mode settings), wireless channel conditions (e.g., congestion and interference), transport settings (e.g., protocol parameter and congestion control algorithms) and application layer settings (e.g., caching, distance to CDNs, size, type and number of page objects). FLEX-MONROE offers the ideal setup for such research, since it provides all the pieces in this chain. We further leverage `WebWorks`[7], the EaaS MONROE offers for testing web performance by monitoring the Page Load Time (PLT) metric towards particular target websites. In the following subsections, we describe the web-performance evaluation experiments we run in FLEX-MONROE.

4.1 WebWorks Design

`WebWorks` enables the collection of multiple web performance metrics while visiting a target webpage using Firefox in headless mode. We leverage the Selenium web automation framework [21] to simulate web surfing and collect web performance metrics in MONROE. Among the several tools that the framework provides, the Selenium webdriver offers a large set of APIs to interact with a given web browser in the same way as a regular user would. For example, we can use the APIs to click on links, buttons or to enter text in input forms. We enable `WebWorks` to use Selenium with Firefox: upon invoking the webdriver, `WebWorks` launches the native Firefox browser in the MONROE nodes to visit any target input webpage. We set the user agent string in Firefox as to retrieve mobile versions of the pages from the web servers. MONROE nodes are not equipped with displays that GUI-based programs like Firefox require to render the output. We thus use the X virtual framebuffer (Xvfb [2]) to mimic the missing display and enable the browser to behave normally.

[7]https://github.com/MONROE-PROJECT/Experiments/tree/master/experiments/WebWorks

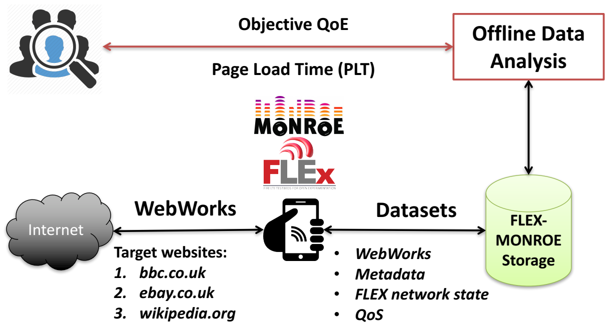

Figure 2: The analysis of web performance in FLEX-MONROE using MONROE `WebWorks`.

During each experiment run we use the HAR export trigger add-on [28] to log Firefox's interactions with the visited pages in a JSON-formatted archive file called HAR (HTTP Archive). We then use the HAR file to derive a number of parameters and metrics, including PLT, size of web pages in bytes, number of objects, size of each objects, number of domains. `WebWorks` tracks as main performance metric the PLT, which is primarily based on OnLoad event triggered by the browser. This event is fired when all objects on a page are loaded. Additional parameters such as object types (javascript, css, image etc.), object load time including DNS resolution time, TCP connection time, object receive timings are also available from the HAR files. Apart from the HAR, we collect RTT statistics during each experiment run.

4.2 Methodology and Experiment Setup

In Figure 2, we depict the workflow for the FLEX-MONROE web experiments. The methodology we put forward consists of a hybrid measurement approach in FLEX-MONROE, where we deploy the `WebWorks` module first in operational networks through MONROE platform and then in a controlled experimental environment through the FLEX-MONROE nodes in the NITOS testbed. This approach enables us to achieve a cross-layer empirical understanding of how varying the LTE network configuration influences the end-user sessions.

Target Websites: Aside from the impact of the network conditions, the content in the website is an important factor impacting an end-user. The category of the website hints towards its complexity in terms of content [10]. Thus, to identify the factors that affect PLT, we analyze complete web page loads of three popular target web pages from different categories that capture the different interests of end-users, namely wiki portal (wikipedia.org), news portal (www.bbc.co.uk) and on-line shopping (ebay.co.uk). We limit the number of websites to only three in order to comply with the data quotas available for each experimenter on the MONROE platform. All nodes resolve the target websites using Google's public DNS resolver; not the mobile carrier's default resolver.

Table 2: Distribution of samples per website and operator.

Operator	bbc.co.uk	ebay.co.uk	wikipedia.org
Orange (ES)	63	113	44
Yoigo (ES)	45	32	33
Wind (IT)	33	20	32
TIM (IT)	42	45	39
Vodafone (IT)	50	46	39
3 (SE)	102	196	77
Telenor (SE)	251	88	93
Telia (SE)	150	169	67
Telia (NO)	60	121	12
Telenor (NO)	131	183	60

MONROE Campaign: We ran `WebWorks` against the three target web pages using HTTP1.1 [12] in 39 MONROE nodes over a period of 2 months. The content of popular websites such as the ones we mention above changes dynamically. Also, the location of the server(s) delivering the web content is not fixed. It is beneficial, in this case, to run the experiment over a longer period in time and analyze web performance when the web content varies. In total, the MONROE nodes reported web-browsing performance from 10 different operators[8] (Table 2). The purpose of running web-based experiments with different operators is to determine the variation of PLTs over different configurations in the wild.

FLEX-MONROE Campaign: To validate the impact of changes in the LTE network, we exploit the opportunities provided by the experimental FLEX platform for tweaking the values of the network-side parameters, such as the Modulation and Coding Scheme (MCS) or the transmission power level from the eNodeB. In particular, we aim to capture how network configurations impact the experience of the end user through the PLT metric. We investigate whether a raise in transmission power level from eNodeB impacts the PLT positively. This is expected since the corresponding raise in signal quality results in higher Channel Quality Indicator (CQI) reporting from LTE UE to its serving eNodeB. Higher CQI means selection of higher MCS for downlink transmission, which is the deciding factor for the data rate.

Datasets: The metrics we retrieve in the **`WebWorks` dataset** include PLT, size of web-page in bytes, number of objects, types of objects, size of each object, per-object download time (including DNS resolution, TCP connect, send, wait and request timings) and different numbers of domains that are accessed during the experiment. In the same time as running `WebWorks`, we retrieve network-side context features (**Metadata**) from all the MONROE node(s) we instrument, both native MONROE and FLEX-MONROE nodes. The list of parameters is extensive and includes the radio access technology (RAT), connection mode, signal strength metrics,

[8]Roaming operators were excluded from the measurement campaign.

RAT-specific parameters, sensor data (e.g., node temperature), RTT and number of hops to the primary web-page server. While experimenting with FLEX-MONROE, we register the state of the LTE network, which we log in the **FLEX network state dataset**. The FLEX network state dataset collects the state of various wireless parameters (e.g., frequency band indicator, downlink bandwidth, uplink bandwidth, modulation and coding profiles, etc.), the power control parameters (e.g., PUCCH SINR Target, PDCCH Power Control, etc.), and CQI report. We also leverage other MORNOE EaaS (i.e., HTTP download) and collect **QoS parameters** in FLEX-MONROE. This allows us to observe how variations in the QoS relate with the network context data and further impact the web performance. We store all this data in the MONROE storage solution, where we define a special repository for FLEX-MONROE (Figure 2). We then use this data for further offline analysis (Section 5). We offer these datasets publicly to the community[9].

5 WEB PERFORMANCE EVALUATION

As mentioned in the previous section, we run a series of web-browsing experiments in FLEX-MONROE using three target web-pages: *bbc.co.uk*, *ebay.co.uk* and *wikipedia.org*. We first investigate the commercial networks, examining the performance of different operators. Figure 3 reveals that the operators differ in their PLT performance both across websites and countries. We observe that Telia (SE) has significantly better performance across all websites with, 3 (SE) and Telenor (SE) falling equally behind. To the contrary, Wind (IT) seems to underperform while showing large variance in PLT between multiple experiments. When it comes to Norway and Spain, there is no clear pattern and web performance varies across different web pages.

All three websites are dynamic with most of the PLT raised by JavaScript objects and images. The *wikipedia.org* webpage has the lowest number of objects, with almost all objects retrieved from the same domain (Table 3). In our analysis of the MONROE web results, regarding factors impacting PLT of a web-page we found that PLT is highly impacted by the websites' own features, namely number of objects downloaded, number of different domains that a page access, and statistics related to objects' sizes.

We find that all the `WebWorks` experiments ran in LTE networks with frequencies of 800, 900, 1800 and 2600 MHz and bands from 508, 1551, 430, 46 and 344. Among the network parameters we monitor in the metadata, we find that the number of times the RSRP value changes during an experiment has the highest impact on the PLT, followed by the LTE band and the average latency values we measured with ping.

[9]https://www.monroe-project.eu/datasets/wintech17_flex-monroe/

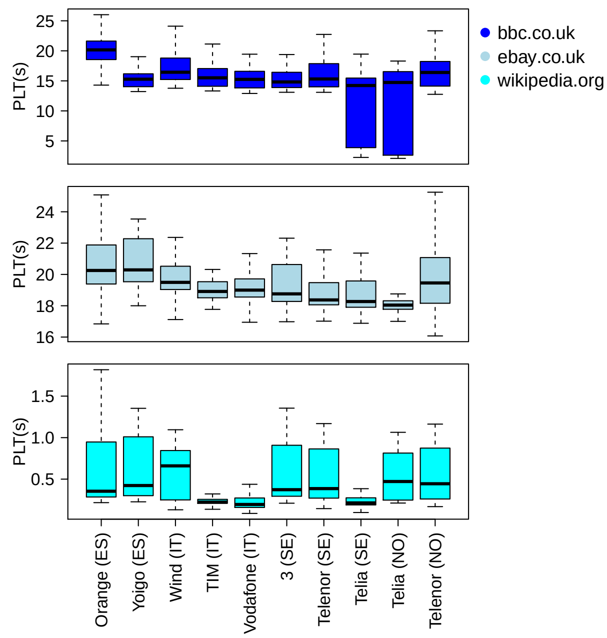

Figure 3: Differences in PLT of target web pages across 10 commercial operators.

Table 3: Features of target web-pages. For each feature we show the average value and in parenthesis the standard deviation. For the JS objects, images, CSS and HTML objects we show the load time reported to the total PLT.

Features mean(std)	bbc.co.uk	ebay.co.uk	wikipedia.org
PLT	19s (6s)	22s (5s)	0.8s (0.7s)
# objects	93 (10)	178 (6)	6 (1)
# domains	32 (2)	24 (1.6)	1.07 (0.2)
JS / PLT	0.41 (0.08)	0.13 (0.08)	0.31 (0.11)
Images / PLT	0.42 (0.08)	0.78 (0.11)	0.49 (0.14)
CSS / PLT	0.06 (0.03)	0.009 (0.005)	0 (0)
HTML / PLT	0.07 (0.03)	0.05 (0.03)	0.19 (0.18)

While analyzing closer the impact of RSRP variation per node on the webpage PLT, we observed that for signal strengths that are close in values – the PLT does not show direct impact. However, by generating bins of signal strength values that are within a distance of 10 units to each other, the bar charts show a tendency of slightly decreasing PLT and less variability with better signal quality. Due to space limitation we have shown results (RSRP vs PLT) from a single MONROE node in Figure 4.

Observing the impact of signal quality on the PLT in the operational networks, we next configure the LTE network in FLEX-MONROE to analyze the same phenomena under the controlled settings. To achieve this, we run the same Web-Works experiments for different transmission power levels at eNodeB. Note that the power level changes at eNodeB in

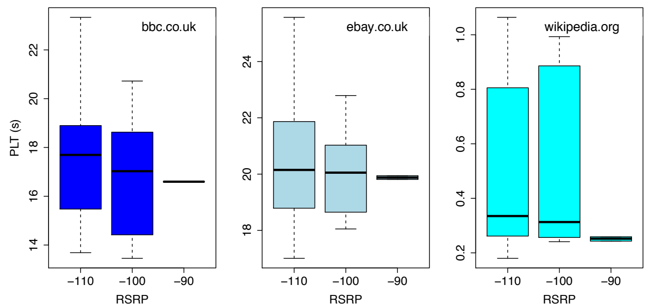

Figure 4: PLT at different signal quality levels for a particular MONROE node.

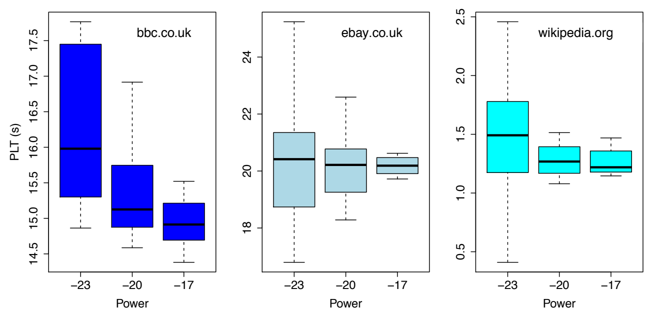

Figure 5: PLT at different power levels

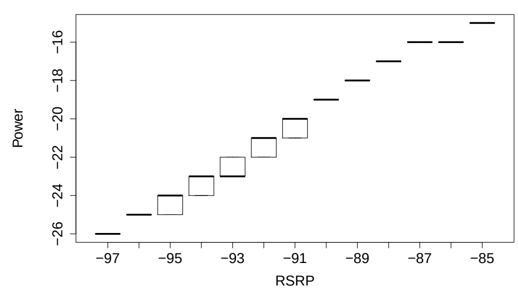

Figure 6: RSRP at different power levels

return affects the received signal strength values. We observe that PLT improves with raise in RFSignalPower (Figure 5); the change in PLT is small though even after a gap of 2 to 3 power levels.

In order to understand whether the signal strength variation is large enough to bring greater change in corresponding PLTs of webpages, we ran further experiments. In Figure 6, we observe that though the quality of RSRP reduces with reduction in the RFSignalPower, the overall range of RSRP is quite small, varying only from -85 to -97 dBm. This whole range is considered *Good* for transmission where RSRP values in general can have best quality (around -44 dBm) to worst (around -140 dBm). We conjecture that by varying the distance between the FLEX-MONROE node and the eNodeB, we may observe more notable changes in RSRP and its propagating effects on performance metrics. We leave this for future work.

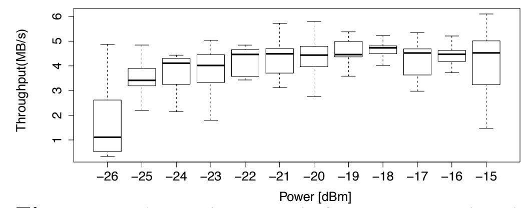

Figure 7: Throughput at different power levels

We further quantify the impact of RFSignalPower changes on the throughput. For throughput analysis we leverage another MONROE EaaS (the HTTP download, in this case) and download 100MBytes of data from the same source in each run. Turning to impact of RFSignalPower changes on throughput, we observe that due to the propagated effect of RSRP to CQI and to MCS, the data rate gradually changes. We show this effect in Figure 7, where the throughput drops from 5 MB/s to 1 MB/s with the fall in power level.

6 RELATED WORK

Evaluation of network performance and assessing the quality experienced by end users require thorough systematic end-to-end measurements. Operators and independent agencies sometimes perform drive tests to identify coverage holes or performance problems in cellular networks. These tests are, however, prohibitively expensive and do not scale well [27]. An alternative common approach is to rely on end users and their devices to run tests [16, 18–20, 25]. The main advantage of crowdsourcing is scalability [13]. However, repeatability is challenging and one can only collect measurements at users' own will, with no possibility of either monitoring or controlling the measurement process. Thus, using dedicated infrastructure [6, 15, 22] tackles many of the limitations of crowdsourcing, at the cost of confining the analysis to a smaller geographical footprint.

The MONROE platform [3, 4] is the first open access hardware-based platform for custom experimentation with commercial mobile carriers in Europe. Even more, the MONROE project offers open datasets on network performance to the community. Several regulators have ongoing nationwide efforts to monitor the broadband networks [11]. Often, the solutions they device are not open to the research community to allow for custom experimentation, nor do they grant free access to the measurement results and methodology. Despite its numerous benefits and exclusive features, MONROE does not allow for experimentation with LTE network configuration changes. Many other individual efforts offer access to infrastructure and technologies [7, 14, 17]. However, no existing testbed can offer the service and experience of an operational commercial network with the unique perk of controlling different elements in the network. FLEX-MONROE fills this gap by unifying the capabilities of MONROE with the advantage of control over the configuration of the LTE experimental networks within FLEX.

We exemplify the research opportunities FLEX-MONROE offers by investigating the interaction between web browsing performance and network configuration. PLT has become the key determinant of user's web browsing experiences [29]. Numerous studies correlate PLT with the end-user QoE [8, 9]. The factors that impact the PLT range from the features of the web-page and the browser to client-device characteristics to the network and physical layer QoS [5, 23, 26, 29]. We focus our study on capturing and validating the impact of LTE network side features on web-browsing performance. For example, with FLEX-MONROE we vary the values for MCS in the NITOS LTE network from 0-28, where 28 indicate maximum data rate. We validate that the MCS impacts the download throughput, which provides an estimate of expected application level performance. Other than the MCS, web-browsing performance is effected by load in terms of number of users as it raises both RTT delay and maximum achieved throughput.

7 CONCLUSIONS

In this paper, we proposed FLEX-MONROE, a unique testbed that offers unified experimentation capabilities in LTE networks both in operational commercial scenarios (MONROE) and in controlled experimental scenarios (FLEX NITOS). We leveraged MONROE EaaS and investigated web-performance QoE features with respect to network side performance parameters in LTE broadband network. This study allowed us to emphasize the potential of FLEX-MONROE and the experimentation opportunities it offers. The dataset we collected is openly available [10]. Also, the testbed is open to authenticated users both in the FLEX and MONROE systems. Users can deploy their own custom experiments or use any of the existing MORNOE EaaS. Our current efforts in terms of testbed integration focus on expanding the FLEX-MONROE platform to also include other FLEX facilities, such are the iMINDS wireless testbed. This latter testbed is particularly interesting for the possibility of experimenting with LTE under mobility scenarios in controlled environments.

ACKNOWLEDGMENTS

This work is funded by the EU H2020 research and innovation programme under grant agreement No. 644399 (MONROE) and No. 612050 (FLEX), and by the Norwegian Research Council project No. 250679 (MEMBRANE).

REFERENCES

[1] 2017. MONROE. (2017). https://github.com/MONROE-PROJECT.

[2] 2017. Xvfb manula page. (2017). https://goo.gl/EPBtpt.

[3] Özgü Alay et al. 2016. Measuring and assessing mobile broadband networks with MONROE. In *World of Wireless, Mobile and Multimedia Networks (WoWMoM)*. IEEE, 1–3.

[4] Özgü Alay et al. 2017. Experience: An Open Platform for Experimentation with Commercial Mobile Broadband Networks. In *Mobicom*. ACM.

[5] Athula Balachandran et al. 2014. Modeling web-quality of experience on cellular networks. In *MobiCom*. ACM, 213–224.

[6] Džiugas Baltrūnas et al. 2014. Measuring the Reliability of Mobile Broadband Networks. In *Internet Measurment Conference (IMC)*. ACM.

[7] Gautam Bhanage et al. 2010. SplitAP: leveraging wireless network virtualization for flexible sharing of WLANs. In *Global Telecommunications Conference (GLOBECOM)*. IEEE, 1–6.

[8] Nina Bhatti et al. 2000. Integrating user-perceived quality into web server design. *Computer Networks* 33, 1 (2000), 1–16.

[9] Anna Bouch et al. 2000. Quality is in the eye of the beholder: meeting users' requirements for Internet quality of service. In *SIGCHI conference on Human Factors in Computing Systems*. ACM, 297–304.

[10] Michael Butkiewicz et al. 2011. Understanding website complexity: measurements, metrics, and implications. In *Internet Measurement Conference (IMC)*. ACM, 313–328.

[11] FCC. 2013. *2013 Measuring Broadband America February Report*. Technical Report. FCC's Office of Engineering and Technology and Consumer and Governmental Affairs Bureau.

[12] R. Fielding et al. 1999. Hypertext Transfer Protocol – HTTP/1.1. RFC 2616 (Draft Standard). (June 1999).

[13] Matthias Hirth et al. 2015. Crowdsourced network measurements: Benefits and best practices. *Computer Networks* 90 (2015), 85–98.

[14] Mahmoud I Kamel et al. 2014. LTE wireless network virtualization: Dynamic slicing via flexible scheduling. In *Vehicular Technology Conference (VTC Fall)*. IEEE, 1–5.

[15] Zahir Koradia et al. 2013. First Impressions on the State of Cellular Data Connectivity in India. In *ACM DEV-4*. ACM.

[16] Christian Kreibich et al. 2010. Netalyzr: illuminating the edge network. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*. ACM, 246–259.

[17] N. Makris et al. 2015. Enabling open access to LTE network components; the NITOS testbed paradigm. In *1st IEEE Conference on Network Softwarization (NetSoft)*. 1–6.

[18] Ashkan Nikravesh et al. 2015. Mobilyzer: An open platform for controllable mobile network measurements. In *International Conference on Mobile Systems, Applications, and Services*. ACM, 389–404.

[19] OOKLA. 2017. http://www.speedtest.net/. (2017).

[20] John P Rula et al. 2014. No one-size fits all: Towards a principled approach for incentives in mobile crowdsourcing. In *Workshop on Mobile Computing Systems and Applications*. ACM, 3.

[21] Selenium. 2017. (2017). http://www.seleniumhq.org.

[22] S. Sen et al. 2011. Can they hear me now?: A case for a client-assisted approach to monitoring wide-area wireless networks. In *Proc. of IMC*.

[23] Junaid Shaikh et al. 2010. Quality of Experience from user and network perspectives. *Annals of Telecommunications* 65 (2010), 47–57. Issue 1.

[24] Sierra-Wireless. 1997. https://www.sierrawireless.com/products-and-solutions/embedded-solutions/products/mc7455/. (1997).

[25] Joel Sommers and Paul Barford. 2012. Cell vs. WiFi: on the performance of metro area mobile connections. In *Internet measurement conference (IMC)*. ACM, 301–314.

[26] Srikanth Sundaresan et al. 2013. Measuring and migrating web-performance bottleneck in broadband access network. In *IMC*. ACM.

[27] Tektronix. 2009. *Reduce Drive Test Costs and Increase Effectiveness of 3G Network Optimization*. Technical Report. Tektronix Communications.

[28] HAR Export Trigger. 2017. (2017). http://www.softwareishard.com/blog/har-export-trigger.

[29] Xiao Sophia Wang et al. 2013. Demystifying Page Load Performance with WProf.. In *NSDI*. 473–485.

[10] https://www.monroe-project.eu/datasets/wintech17_flex-monroe/