# Experimental Estimation of LTE-A Performance

Imane Oussakel
*LAAS-CNRS, Université de Toulouse*
*CNRS*
Toulouse, France
imane.oussakel@laas.fr

Philippe Owezarski
*LAAS-CNRS, Université de Toulouse*
*CNRS*
Toulouse, France
philippe.owezarski@laas.fr

Pascal Berthou
*LAAS-CNRS, Université de Toulouse*
*CNRS*
Toulouse, France
pascal.berthou@laas.fr

*Abstract*—In cellular networks, the emergence of machine communications such as connected vehicles increases the high demand of uplink transmissions, thus, degrading the quality of service per user equipment. Enforcing quality-of-service in such cellular network is challenging, as radio phenomena, as well as user (and their devices) mobility and dynamics, are uncontrolled. To solve this issue, estimating what the quality of transmissions will be in a short future for a connected user is essential. For that purpose, we argue that lower layer metrics are a key feature whose evolution can help predict the bandwidth that the considered connections can take advantage of in the following hundreds of milliseconds. The paper then describes how a 4G testbed has been deployed in order to investigate throughput prediction in uplink transmissions at a small time granularity of 100 ms. Based on lower layer metrics (physical and mac layers), the main supervised machine learning algorithms are used, such as Linear Regressor and Random Forest to predict the uplink received bandwidth in different radio phenomena environment. Hence, a deep investigation of the impact of radio issues on bandwidth prediction is conducted. Further, our evaluation shows that the prediction is highly accurate: at the time granularity of 100 ms, the average prediction error is in the range of 6% to 12% for all the scenarios we explored.

*Index Terms*—QoS monitoring, cellular networks, bandwidth prediction, machine learning, software defined radio.

## I. INTRODUCTION

In the scope of smart cities, where everything is connected at anytime and anywhere, a tremendous growth of services and/or applications demand is forecasted. Most applications rely on wireless networks to send their data to the corresponding remote internet servers. For instance, the Intelligent Transport System -ITS- emphasizes many services relying on car communication system. In such system, vehicular video conferencing, traffic monitoring and data uploading are handled by cellular network, thus, leading to a high amount of data to be managed by such networks. For instance, automotive data analytics [10] forecasts a production of up to 30 terabytes of data per day by an average connected car to be sent to ITS servers. Therefore, contrary to the Human-to-Human applications where downlink transmission dominates, smart cities services use mostly uplink communication. Hence, the conventional roles are inverted: when all the connected nodes are sending their information, a large amount of data shall be received by their corresponding server with a high quality of service.

In wireless networks, the high performance metrics, such as throughput, change rapidly depending on the environmental situations. However, radio phenomena such as noise, pathloss and multipath fading affect the different performance metrics. It is increasingly hard to understand how such phenomena degrade the Quality of service, as they are uncontrolled. In a more general context, the QoS degradation in wireless network is mainly caused either by radio phenomena variation or when the total requested bandwidth exceeds the cells capacity [26]. In this article we focus mainly on radio issues: environment issues and radio congestion. Hence, the middle nodes are well provisioned to avoid any transport congestion in the network. 4G and 5G networks are expected to satisfy the high demand of high QoS, i.e. high throughput and low latency, not only for legacy, but also smart cities applications. Nevertheless, with the introduction of machine communication, the growth of machine traffic outpaces that of cellular networks, resulting in a quality of service degradation. In fact, several researches show that LTE suffers from congestion [13], [14], especially in uplink. This congestion implies a QoS degradation and sometimes the inaccessibility to the network, which limits the envisaged applications and services.

Over years researchers were focused on enhancing downlink transmission as only downlink was challenging. They improved congestion control mechanisms while taking predicted throughput as an input parameter [1], [12]. For that, downlink cellular throughput prediction has been the ultimate goal of research, as it improves many use cases, such as video streaming quality.

Many dissimilarities between uplink and downlink lead to the insufficiency of the proposed techniques for downlink bandwidth prediction to be applied in uplink. In fact, contrary to downlink (DL) transmissions where each user predicts the incoming throughput from one base station (eNB) at a time, during uplink (UL) transmissions, the eNB will need to predict the incoming throughput of multiples users simultaneously. Moreover, the scheduling and the access techniques of DL and UL are different. For instance, the UL scheduler is synchronous while the DL is asynchronous. The UL uses SC-FDMA (Single-Carrier Frequency Division Multiple Access) as a multiple access scheme while the Orthogonal Frequency Division Multiple Access (OFDMA) is used for DL. Hence, the UL user data is affected differently by the channel variation compared to DL and the signal treatment is also different in the receiver (eNB). Such differences with downlink motivate our work on uplink performance study to investigate the possibility

of estimating/predicting the bandwidth for uplink transmissions, that could be different from related work focusing on downlink.

Even though uplink and downlink are different, we forecast an improvement of uplink LTE-A schedulers in a similar way as downlink [9], by incorporating the predicted throughput. Furthermore, a cell radio congestion avoidance can be achieved by comparing the estimated coming data rate and the base station capacity. Machine application servers may avoid saturation based on the forecasted bandwidth, hence, reducing uplink quality of service degradation. Thereby, uplink bandwidth prediction is valuable in the machine-to-machine context.

In this article, several scenarios are tested in order to understand how each radio phenomenon influences the bandwidth in cellular networks, aiming to estimate/predict the bandwidth at small time granularity. To that end, an LTE-A testbed is deployed, where the radio propagation is controlled in an anechoic room. The middle nodes are well provisioned to avoid throughput degradation caused by network bottleneck. Therefore, throughput deterioration is mainly caused by environment phenomena. We investigate our work on bandwidth prediction based on lower layers measurements/metrics over a small time granularity, from 100 ms to 1 s scales. Different machine-learning algorithms are tested to validate the impact of radio issues on bandwidth prediction.

The remaining of this paper is organized as follows. Section II points out the LTE-A background of the experimentation. The related work is reviewed in section III. Section IV describes the testbed deployment inside an anechoic room. Section V covers the prediction methodology used in this article. It includes data collection procedure, data analysis and the developed model for the prediction evaluation. Section VI presents and discusses the obtained results for uplink bandwidth prediction. Finally, section VII concludes this paper.

## II. LTE-A BACKGROUND

Our work is based on 4G communications, as uplink cellular traffic has become dense with the introduction of M2M communications, especially connected cars. This section presents a quick overview of the main 4G metrics and techniques that our study is based on, motivating our focus on uplink transmissions.

Fig. 1 shows a basic architecture of the 4G network when a User Equipment (UE) is connected via the LTE access network to the Evolved Packet Core network (EPC). The EPC maintains the sessions, registration procedures and routing of UE IP-packets. The base station for LTE-A radio is named evolved NodeB (eNB); it ensures mainly Radio resource management and scheduling in both uplink and downlink. Further, to ensure the multiple users access, 4G uses OFDMA in downlink and SC-FDMA in uplink. Orthogonal Frequency Division Multiple Access (OFDMA) is a multiple carrier system, where each OFDM data symbol is transmitted over one sub-carrier. In contrast, with single-Carrier Frequency Division Multiple Access (SC-FDMA) in uplink, data symbols are transmitted

in series and each symbol is carried by a wider bandwidth. Hence, contrarily to OFDMA that reduces/vanishes the inter-symbol interference (ISI), SC-FDMA is prone to ISI. ISI reduces data-rate when no compensation is present. For that the eNB deploys as a first step a complex frequency equalizer to mitigate such distortion. Therefore, UL data-symbols are not only affected differently by the channel variation compared to DL, but signal treatment is also different in the eNB. Such dissimilarities with DL motivates our work on uplink performance study, that could be different from related work focusing on downlink only.

Fig. 2 exhibits the frame type used for FDD (Frequency Division Duplex) radio transmission mode, i.e. different frequency bands are used for downlink and uplink. The frame length is fixed to 10 ms, containing 10 sub-frames of 1 ms (two adjacent slots). Each slot is defined as a resource block of 0.5 ms over 12 sub-carriers spaced by 15 kHz. The LTE-A radio resource is then referred as Physical Resource Block (PRB) that forms 180 kHz x 0.5 ms. In time domain, the PRB contains seven SC-FDMA data symbols expanded over 12 sub-carriers in frequency domain.

The eNB allocates for each UE a specific number of resource blocks based on its resource-scheduling algorithm, UE's capability and the channel quality. Although schedulers are vendor specific, it is reported to consider one TTI (Transmission Time Interval) of 1 ms as the smallest scheduling time unit [15]. Thus, based on the used bandwidth, a defined number of useful resource blocks is available per slot. For example, for a 5 MHz bandwidth, 25 physical resource blocks of 180 kHz could be assigned to the users attached in the cell. In order to handle the communication between the UE and eNB, the use of these PRBs is standardized by using channels formats in the lower layers. For the uplink transmissions mainly three channels are used [27]. However, the UE activates the PRACH, Physical Random Access Channel to demand access to the network; it communicates the control signaling information using the Physical Uplink Control Channel (PUCCH). PUCCH is mainly located in the bounds of the used bandwidth in frequency domain as illustrated on fig. 2. The UE transmits its data on the Physical Uplink Shared Channel (PUSCH), which uses the remaining PRBs.

In the wireless system, channel is changing rapidly, resulting in bandwidth variation. In order to adapt to channel's change and guarantee a high quality of service over time in downlink and uplink communications, the UE and eNB accomplish many measurements and reports. The 3GPP has normalized the main measures and mechanisms for each side of the network. As we are interested in uplink communication, we cover in
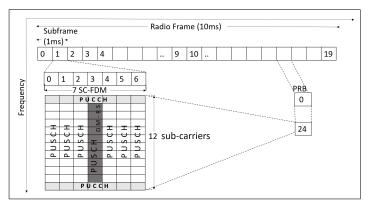


Fig. 1: Basic 4G architecture.

Fig. 2: Uplink resource grid for one slot.



Fig. 3: LTE-A Testbed deployment.

more details the main uplink measurements/metrics used for prediction in section V - part V.A.

## III. RELATED WORK

Over years, bandwidth estimation and prediction has been widely studied in wired networks and WLAN. D. Koutsonikolas et al. [2] reveal the ineffectiveness of using those techniques in cellular networks as they are characterized by the large short-scale fluctuation of bandwidth. Although, [4]- [6] show the possibility of throughput estimation and prediction in wireless networks essentially in Downlink. Their proposed methods are based on higher layer measurements. In cellular networks, based on radio measurements, multiple performance tasks are introduced at different layers of the protocol stack. For instance, when a bad signal is received HARQ (Hybrid automatic repeat request) is triggered in the second layer. Given this, we are interested in using the lower layers measurements/metrics to predict bandwidth instead of higher layer measurements.

Even though, both client and network throughput are important in cellular networks, link bandwidth prediction related work considered only Downlink channel. F. Lu et al. [1] use CQI (Channel quality indicator) and DRX (Discontinuous Transmission) to predict instantaneous downlink throughput in 3G networks. The study in [18] classifies bandwidth into high and low categories in order to increase user equipment coordination efficiency for transmission in cellular background. Margolies et al. [9] investigate the reproducibility of signal quality over the same path to predict throughput based on user trajectory tracking. Authors of [7], [8] proposed the use of machine learning techniques, especially Random Forest, to predict instantaneous throughput in LTE-A networks. However, A. Samba et al. [8] predict throughput based on both eNB and UE information before the connection is established for content providers. They conclude that applying RF on radio measurements leads to a promising prediction: 52% of the relative prediction errors are within ±20%. C Yue et al. [7] consider a set of lower layer measurements and historical throughput to predict real-time LTE-A throughput. Accurate predictions were obtained: 69% of the relative prediction errors are within ±10% for walking scenario.

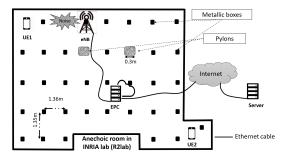All of the above studies focus on predicting bandwidth for only DL transmissions. Our work differs in that we are interested in UL that takes advantage of a different transmitter and receiver composition.

For estimating/predicting UL traffic, we apply mainly Linear Regressor (LR) [11] and Random Forest (RF) [23] as machine learning techniques to predict the UL bandwidth based on lower layers metrics/measurements. The main reason of using LR is to investigate the linearity between the lower layers metrics related radio channel and the high QoS metric, i.e. bandwidth. In contrast, RF is well known to provide good predictions in different application domains with accurate time series and high dimensional feature spaces [16], which corresponds to the case study in this article. Hence, a comparative study of the two ML techniques performance in cellular communication domain is conducted.

## IV. TESTBED DEPLOYMENT

The testbed objective is to limit the uncontrolled effects during a wireless communication, and study the impact of disturbing low level features on bandwidth metric. The wireless system is LTE-A as already mentioned in the paper. R2lab [22] is used as an underlying testbed, which offers a wireless network with multiple Software defined radio (SDR) nodes inside an anechoic room. The RF propagation inside the room is controlled thanks to the microwave absorbers materiel on the walls, scattering any wireless signal that comes across. The SDR based nodes allow a full access to realistic physical metrics/measurements. With remote access to the indoor nodes, the deployment of LTE-A network is accomplished. Fig. 3 schematizes the testbed. Openairinterface (OAI) software based platform [3] is implemented, which spans the full LTE 3GPP protocol stack (including features from LTE-Advanced and LTE-Advanced-Pro) for both radio access network (E-UTRAN) and core network (EPC). The openairinterface softmodem is connected with a hardware platform for SDR: USRP B210 (Universal Software Radio Peripheral) attached to two antennas, receiver and transmitter antennas. The USRP B210 is connected to a host computer to perform processing, and then connected to a PC running EPC network, and accessing the internet toward a distant server. The testbed works on frequency band 7 with 5 MHz bandwidth and uses FDD mode, which corresponds to the traditional and

stable version of openairinterface platform with USRPB210. Further, as our objective is to observe the impact of radio phenomena on the bandwidth variation, the server and the intermediate PCs are well provisioned for not behaving as a bottleneck during communication.

To emulate 4G connected nodes, two commercial UEs are used and placed inside the anechoic room. The hardware metallic enclosure boxes scattered inside the room are considered as fixed multipath sources, which we take advantage of in investigating the impact of multipath phenomenon on bandwidth variation.

In real environments, multiple radio phenomena are scrambling the communications. Example phenomena include multipath fading leading to InterSymbol Interference (ISI) noise, pathloss and random processes such as AWGN (Additive White Gaussian Noise). These phenomena tend to attenuate aggressively the transmitted signal, which causes a significant amount of signal strength reduction and hence QoS degradation. The upcoming scenarios tend to investigate the impact of each phenomenon on the bandwidth prediction. For each test scenario, a radio phenomenon is added in the system to complexify the tests.

*a) Scenario 1: Multipath fading:* For the first scenario, we investigate the impact of multipath fading and pathloss on bandwidth prediction/estimation. In other words, the two UEs used in the experimentation are on different distance from the eNB, i.e. each UE has a different pathloss. Over the room, multiple metallic boxes surround the eNB (fig. 3). They introduce multipath fading leading to ISI. The two pylons, covered with absorbers are considered as shadowers for the transmissions, especially for UE2. Therefore, only pathloss, shadowing and multipath fading are present in the system as scramblers for the transmission, i.e the anechoic room is isolated from any other radio phenomena. In order to avoid any degradation/losses due to the insufficiency of radio resources the total transmitted bandwidth by the two users is inferior to the maximum network capacity. In fact, after several tests the maximum uplink data-rate achieved in the testbed, by 1 connected UE, is around 8Mbps. Hence UE 1 transmits during the whole test duration a fixed amount of data of 2Mbps and UE 2 transmits 3Mbps. A benchmark of lower layers metrics is performed to construct datasets for ML algorithms as explained further.

*b) Scenario 2: Noise & Multipath fading:* In this scenario, another radio phenomenon is added in the anechoic room to investigate its presence on the bandwidth prediction, i.e, noise. For that we generate in a controlled manner a specific noise profile to scramble the communication.

**Noise profile:** In time varying scenarios, the received signal amplitude undergoes rapid fluctuation that is often modeled as a random variable with a particular distribution. For that we consider a Gaussian distribution, AWGN, which is characterized by its amplitude that affects the signal strength. Moreover, noise (AWGN) is introduced as it causes transmission errors and may disrupt the communication with ISI production for high power noise [24]. Contrarily to work in literature where

AWGN is often taken with constant attenuation, we introduce randomness in the attenuation in order to have attenuation fluctuations of the signal over time. For that we define noise level as noise with a given gain and amplitude. We used mainly GNUradio on USRP B210. Given a list of gain levels and an interval of maximum and minimum amplitude levels, each 10 seconds a random value of amplitude and gain are chosen. In fact, the amplitude value affects the statistical characteristics of the noise source, i.e, the standard deviation of the Gaussian noise. The gain affects the transmitted signal power. The programmed step for noise level change (10 s) is chosen as to have sufficient samples for each noise level. Therefore, low noise level values keep the channel flat, while high noise level disrupt totally the communication, with the probability of introducing ISI, not only with noise but also with metallic boxes scattered in the room. Furthermore, the abrupt changes in noise levels during the transmission tend to reflect the real complex radio environments, where the user's mobility across different shadowers leads to aggressive/alleviated signal attenuation. The bandwidth is fixed to 5Mhz to scramble the full UL bandwidth. Both users transmit at same data-rate as in the scenario 1, and same benchmark of lower layer metrics is performed.

*c) Scenario 3: Radio congestion & Noise & Multipath fading:* During the previous scenarios, the total transmitted bandwidth by the two UEs is much lower than the maximum available bandwidth. In this part, another radio phenomenon leading to bandwidth degradation is introduced in the testbed, i.e. radio congestion. The later occurs when the total bandwidth required by the connected UEs outpaces the maximum eNB bandwidth. Mobile application based speedtest is tested in scenario 1 for UE 1. It reaches the maximum throughput of 8 Mbps. Therefore, in order to realize radio congestion, UE1 transmits its data at a fixed amount of 4 Mbps and UE2 at 5 Mbps. Noise is introduced as described in scenario 2 to complexify the test.

For all the above scenarios, IPERF3 generates traffic at the UE side and IPERF3-server monitors throughput reception in the server. In order to have a fixed transmission amount of data during the whole test duration, UDP (User Datagram Protocol) is used as a transport protocol. In fact, TCP (Transport Control Protocol) changes the transmission window based on the perceived packet losses in the window. Given this, any observed bandwidth degradation is essentially due to radio environment variation.

## V. Prediction Methodology

Given the LTE-A testbed in the anechoic room, we are able to generate 4G traffic in different scenarios as described in the previous part. In this section, we present the methodology aiming at building our datasets.

### A. Data collection

The eNB performs different measurements in order to decode the received data and adapt to channel variation. With SDR at the eNB side, we are able to collect all the performed

eNB measurements and metrics, especially from the lower layers (physical and mac layer). As the connected UEs communicates their control signaling information using PUCCH, and transmit their data on PUSCH channels, we perform a deep benchmark of the main metrics/measurements linked with the two channels. The OAI eNB metrics/measurements are performed as depicted by the 3GPP standard. The collected metrics are mainly extracted during the lower layers data processing.

On one hand, radio measurements are collected as they are crucial for higher layer mechanisms and reflect the channel quality. The main measures are:

- Timing Advance (TA): As the propagation delay of different UEs is different based on their positions, TA is introduced to ensure synchronization between uplink and downlink at the eNB side. TA is a negative offset at the UE. Based on the UE transmitted PRACH, the eNB estimates the initial TA. Once the UE is connected, the eNB keep estimating TA and adjusting it by transmitting to the UE the required value, TA_update.
- SNR: SNR: Signal to Noise Ratio compares the level of the desired received signal to the level of noise. Taking $P_{signal}$ and $P_{noise}$ as the average power of the received signal and noise respectively, SNR is defined (in decibels) as follows: $SNR_{dB} = 10.log_{10}(P_{signal}/P_{noise})$. It is measured for each received PUSCH holding UE's data.
- received UL_CQI: Uplink Channel Quality Indicator. It is computed at the eNB based on the observed SNR.
- PUCCH received power and noise power: the two measures are estimated principally for each PUCCH handling a scheduling request.
- PUCCH threshold: It is the threshold to detect the pucch format1.

On the other hand, during decoding/demodulating the received channels, multiple metrics are extracted. In fact, when the UE is transmitting data via PUSCH, its data is transmitted using a Modulation Coding Scheme (MCS). The MCS is related to the modulation order Qm, e.g, QPSK (Quadrature Phase Shift Keying), 16 QAM (Quadrature Amplitude Modulation) and 64 QAM. The modulation scheme has a direct impact on the data-rate. However, when using a higher-order QAM scheme such as 64 QAM, each symbol represents a group of bits to be transmitted, i.e. each symbol has a 6-bit signature. The MCS is chosen based on the measurements exchange, reflecting the channel quality and the network capacity, between the eNB and the UE. Further, the transport block size (TBS) could be determined based on MCS value and the number of used resource blocks. TBS refers to data in the physical layer [27].

During the UE data transmission toward the network, eNB checks for data correctness using CRC (Cyclic Redundancy Check) and sends ACK/NACK to the UE. If the channel is good, errors are detected and corrected. On the other hand, when the channel quality is bad, the CRC might be insufficient. Further, LTE implements synchronous HARQ (contrarily to the asynchronous HARQ in DL) which combines ARQ error-control and a high-rate forward error-correcting coding (FEC). This way, the eNB requests a retransmission in case of error correctness inability. Such mechanism is characterized by multiple parameters, such as the round and redundancy version (rv). Is it possible to send ACK/NACK by the UE during the transmission, hence the length of ACK information in bits is considered. Moreover, the length of each received SDU/PDU (Service Data Unit/Protocol Data Unit) on the MAC layer is extracted with the buffer size when data is handled.

All of the aforementioned parameters are collected while decoding/demodulating the received PUSCH/PUCCH channels. In addition to other metrics usefull for PUSCH and UL-SCH (UpLink Shared Channel) related functions, such as decoding time, decoding iteration, beta_offset, number of CRC bits and the cyclic shift. A total of 43 metric is collected during the whole test duration, i.e. 400 seconds.

One of our objectives is to compare the bandwidth predictions over different small time granularity. The bandwidth measurements are performed in a discrete time manner, each $\delta t$. As the minimum time report interval in IPERF3 is 100 ms, we fix $\delta t = 100ms$. The predictions are then made every $\delta t$. The eNB metrics are collected per subframe scale (1 ms). In order to have a representative measure per $\delta t$, we compute the maximum, minimum, mean, median and the standard deviation of each metric per $\delta t$ to construct the datasets. Therefore, each metric $\gamma$ is represented in the dataset as follows: $\{\gamma_{min}, \gamma_{max}, \gamma_{mean}, \gamma_{median}, \gamma_{std}\}$. For each UE $u$, $u \in \{1,2\}$ and each scenario $s$, $s \in \{1,2,3\}$, a dataset is constructed, noted $dataset\_s\_u$. Each $dataset\_s\_u$ contains all the lower layers metrics collected for the given UE $u$ during the scenario $s$, including historical received bandwidth.

### B. Methodology

In this article, Python scikit-learn library is used for all the tests. GridsearchCV [17] is applied to choose the optimal hyper-parameters for each estimator. It combines both gridsearch and cross-validation methods. Gridsearch consists of an exhaustive search over subset values of parameters for a given estimator and cross-validation (CV) technique [20] estimates the prediction error of a model. CV is categorized into exhaustive and non-exhaustive categories, the former is more computational for high dimensional datasets. For that, the non-exhaustive cross-validation is chosen, mainly the recommended K-fold method, with K=10 to have a good compromise between variance and bias of the model [19]. Varma and Simon [21] report that the estimated prediction error from the cross-validation used to tune hyper-parameters is biased, and recommend the use of nested cross-validation instead, where an inner CV is used to select the optimized model (executed with GridsearchCV) and an outer CV to estimate the prediction error. Let denote $K_1$-fold and $K_2$-fold the inner and the outer CV respectively. Given an input dataset, a random split is performed to construct training and test sets. In fact, the dataset is split into $K_2$-folds, one fold is used for testing and the others $K_2$-1 folds constitute the training set. For each hyper-parameter combination from the gridsearch,

**Algorithm 1** Bandwidth prediction from $dataset\_s\_u$

---

**Input:**          Historical metrics/measures with lag size $w$, historical received bandwidth, M: optimized prediction model.
**Output:** $\hat{BW}$=$[(BW^1_{(1\delta t,w)},..,BW^1_{(i,w)}),..,(BW^m_{(1\delta t,w)},..,BW^m_{(i,w)})]$: the predicted bandwidth.
  **for** each(w,i) **do**
    **for** t in range (1, m) **do**
      $(BW^t_{(1\delta t,w)},BW^t_{(2\delta t,w)},..,BW^t_{(i,w)}))$=$M(X^t_{test},X^{t-1}_{test},..,X^{t-w}_{test})$
      $X^t_{test}$ is the value of a feature in time t.
    **end for**
  **end for**
  **return** $\hat{BW}$

---

$K_1$-fold is applied. It divides the training set into $K_1$ equal folds; $K_1$-1 folds are used for training and the remaining fold for evaluation. It computes the prediction error and iterates until all the folds are used for both training and validation, then the prediction error is averaged over all the $K_1$ cases of CV. The hyper-parameter combination achieving a minimized prediction error is selected as the best optimized model. In order to generalize the selected model, the outer loop CV is used where the model is tested $K_2$ times on unseen data, i.e. the test set. Then, the generalized prediction error is the average of the estimated prediction error over the tested sets.

Algorithm 1 represents the prediction model for bandwidth. Given the dataset, we apply the algorithm summarized in the following. Let N be the size of the considered dataset. $K_2$-fold CV is chosen for having 25% of the dataset for testing which we denote $X_{test}$={$X^1_{test},..,X^m_{test}$} with size m, and 75% for the training set denoted $X_{train}$={$X^1_{train},..,X^n_{train}$} with size n, where $X^p_{train}$ and $X^p_{test}$ are the $p^{th}$ samples in their corresponding datasets. GridsearshCV is applied on the $X_{train}$ to select the optimized model, referred by $M$. Let $w$ be the rolling window, it represents the past $w$ time units, and $i$ denotes the forecast window, where $i \in \{1\delta t...10\delta t\}$, with $\delta t$=100 ms. The maximum forecast window is then 1 second. For each lag $w$ and forecast window $i$, $M$ uses the historical metrics $(X^t_{test},X^{t-1}_{test},..,X^{t-w}_{test})$ to predict for each sample $X^t_{test}$ from $X_{test}$ the corresponding bandwidth $BW^t_{(i,w)}$, $i \in \{1..10\}$. For example, with $i$=1 and $w$=2, $M$ uses the current and two past time units of $X^t_{test}$: [$X^t_{test}$, $X^{t-1}_{test}$, $X^{t-2}_{test}$] to predict the upcoming bandwidth in 100ms.

Therefore, for predictions evaluation, we compare the predicted bandwidth $BW^t_{(i,w)}$ with the received $BW^t_{(i,w)}$ bandwidth, based on RMSE (Root Mean Squared Error) metric. RMSE is attractive from a statistical and scientific perspective. It represents the average error prediction in the model, expressed in the units of the variable of interest. It is computed as follows: $RMSE = \sqrt{(1/n * \sum^n_i (y_i - \hat{y_i})^2)}$ where $y_1...y_n$ are the actual values and $\hat{y_1}...\hat{y_n}$ the predicted ones. By squaring the error, a high weight is given to the large errors. RMSE score is negatively oriented, hence lower values are better.

*C. Correlation analysis*

All of the aforementioned metrics/measurements depicted in section V part V-A are combined to constitute the datasets per UE and per scenario. A total of 58 metrics is then used. Each sample of the dataset contains the maximum, minimum,
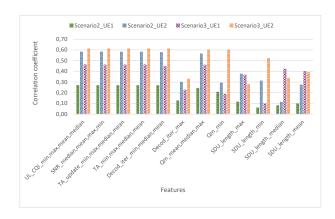


Fig. 4: Features correlation with bandwidth.

mean, median and the standard deviation of each metric. When the past measurements are added ($w = 1$), a high dimensional features space is then considered. In this part the correlation between the collected features and the bandwidth is investigated. Because of place constraint, only some of the most correlated features with bandwidth over all scenarios are presented. Fig. 4 depicts the correlation coefficient between the selected features and bandwidth for both UEs during scenarios 2 and 3. For a given feature, when the correlation level of different statistics is pretty similar, only one bar is presented, i.e. $SNR\_median, mean, max$ indicates that the SNR_median, SNR_mean and SNR_max have pretty similar correlation level with bandwidth.

Based on the figure, the main representative statistics for a given feature are the mean, maximum, minimum and the median. Most of the features mark higher correlation level for UE 2 compared to UE 1. For example, during scenario 2, TA_mean reaches 0.58 for UE 2 and only 0.28 for UE 1. Such results point out the impact of fading multipath on the features variations. It is also remarked that the correlation increases for almost all the features when scenario 3. For instance, the correlation coefficient for Qm_mean (modulation order) for UE 1 increases from 0.25 in scenario 2 to 0.46 in scenario 3. This increase in correlation level sheds light on the radio congestion impact on the lower layer metrics. Further, extreme radio phenomena lead to higher correlation between bandwidth and lower layer metrics/measurements.

*D. Feature selection*

In order to exploit the linear relationship between the collected metrics and the bandwidth, LR is applied as an underlying ML technique. On the other hand, RF has the ability to create the linear and non linear boundaries during the trees building, which leads to accurate predictions. RF is then applied. As many features are available, the selection of the most important ones is crucial, especially for LR. To this end, the statistical technique Principal Component Analysis
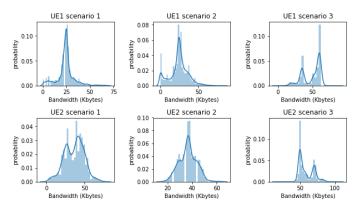
Fig. 5: Bandwidth standard deviation over all scenarios.

(PCA) is used [28]. PCA reduces the high features space dimension with variance maximization of each component. For instance, it uses orthogonal transformation to generate linear combination of orthogonal features vectors. After testing multiple component numbers, using a single component leads to the minimum prediction error. Therefore, the number of component is fixed to 1 for the upcoming tests based LR. In contrast, an approach for features importance determination is implemented in RF models. RF combines multiple decision trees to get a more accurate prediction. For each node split, it looks for the most important feature among a random subset of features. In fact, a given feature is considered as important if its perturbation leads to larger error. Hence, RF relies on features importance for building the forest. Thus, turning PCA unnecessary in the case of RF.

### E. Bandwidth analysis

As the focus of this work is mainly the correlation between radio environment and bandwidth, in this part, an analysis of the impact of radio issues on the received bandwidth is conducted. For that, fig. 5 illustrates the standard deviation of the received bandwidth for each UE during each scenario. Each column is scenario specific. During scenario 1, the two UEs transmit at a fixed low level data-rate. From the figure, UE 1 bandwidth distribution is straight compared to the UE 2 received bandwidth. In fact, the amount of received bandwidth for UE2 is spread out over much wider range. Such issue could be mainly explained by the radio phenomena experienced by UE 2. The UE 2 transmissions face a high level of multipath fading and shadowing compared to UE1 transmissions. UE 1 has a direct line of sight with the eNB. This observation points out the sensitivity of the bandwidth to multipath fading phenomenon.

For scenario 2, noise profile is injected in the testbed. The noise source is placed between the UE1 and the eNB. On the second column of fig. 5, it is clear that noise has added variance to the UE 1 distribution, with the curve becoming right skewed. Such impact is expected as the noise variation introduces errors and ISI, especially for UE1 transmissions. In contrast, surprisingly the UE2 bandwidth variance is decreased

in this scenario. The minimum received bandwidth is increased too. Such result exhibits the positive impact of the noise profile on the unwanted multipath signals. This interesting result need to be investigated deeply in further work.

When the transmitted data outpaces the network capacity in scenario3, two modes are highly distinguishable with low variance for the both users: the maximum received bandwidth is increased. Such result is expected, as during the transmission, a high level of packet loss is observed and retransmission was highly active. It is worth mentioning that multiple metrics values were missed during that scenario. It reflects the real complex wireless systems, where the presence of multiple radio phenomena leads to severe bandwidth degradation.

Overall, this illustration sheds light on the high correlation between radio phenomena and high level QoS, i.e. bandwidth.

## VI. UPLINK BANDWIDTH PREDICTION EVALUATION

This part presents and discusses the results obtained on bandwidth prediction using machine learning techniques, i.e. LR and RF. Six datasets are tested with two ML techniques RF and LR, i.e. one dataset for each UE per given scenario. The evaluation is based on RMSE as a performance metric.

### A. Prediction accuracy

In order to compare the prediction performance of the two ML techniques, fig 6 exhibits error prediction based RMSE for each UE over all the scenarios. With the two MLs, the prediction error doesn't exceed 13 Kbytes. In general RF outperforms LR, except for dataset_1_1. For instance, for dataset_2_1 (scenario 2, UE1), the observed RMSE using RF is 6.44 Kbytes and 12.04 Kbytes using LR. It is to be noted that the maximum received bandwidth in scenario 2 is 82.3 Kbytes and 64 Kbytes for UE 1 and UE 2 respectively. Hence, RF leads to accurate predictions, around 7.8% of errors in an environment where only multipath and noise are present. For more complex scenarios such as scenario 3, the maximum received bandwidth is 92 Kbytes and 105 Kbytes for UE 1 and UE 2 respectively. The prediction error in such scenario reaches 13 %, but remain acceptable. Therefore, this results prove the possibility of predicting accurately uplink bandwidth using lower layer metrics/measurements in different radio scenarios.
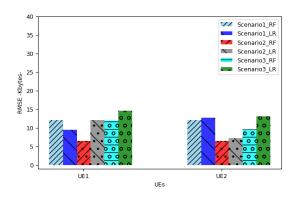


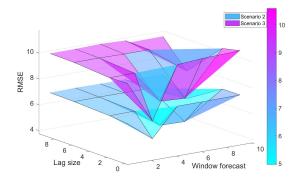Fig. 6: Error prediction for forecast window i=1$\delta t$.

Fig. 7: Sensitivity window forecast and lag size for UE 2.



Fig. 8: Sensitivity to training dataset length.

## B. Window forecast and lag size sensitivity

In order to exhibit the influence of window forecast size on predictions, we evaluate each dataset over $i\delta t$, with $i \in \{1, 3, 5, 7, 10\}$ and $w = 1$. That is, we predict bandwidth based on the instantaneous and the past lower layers measures/metrics, including the past received bandwidth. In addition, the importance of having numerous past radio measurements for good predictions is analyzed. The main lag sizes $w$ tested are 3, 7 or 10, i.e. using the past three, seven and ten $\delta t$ measurements. Fig. 7 plots in 3D the variation of the perceived RMSE over $i\delta t$ with $i \in \{1, 3, 5, 7, 10\}$ and per lag size $w$, $w \in \{1, 3, 7, 10\}$ for UE 2 prediction on dataset_2_2 and dataset_3_2 using RF as an underlying ML for predictions. Based on dataset dataset_2_2, the RMSE doesn't exceed 6.47 Kbytes for all the forecasted windows. The prediction at a granularity of $\delta t = 3$ leads to a smooth amelioration, i.e. the RMSE decreases by 1.49 Kbytes. Generally, when increasing the forecast window, a negligible increase of RMSE is observed based on both datasets. Thus exhibiting the insensitivity of the models to the forecast windows.

The introduction of lag size for predictions at granularity of $1\delta t$ doesn't improve necessarily the performance, i.e. using dataset_3_2, the RMSE increases from 6.46 Kbytes to 6.78 Kbytes when using $w = 10$. Such result points out the insensitivity to lag size when predicting at a granularity of $1\delta t$. On the other hand, a remarkable improvement is observed when predicting larger forecast windows and using lag sizes of $w = 3$ or $w = 7$. This indicates the sensitivity to lag size for predictions of higher forecast windows. Considering that larger $w$ may not necessary improve the prediction accuracy while leading to a higher computation overhead, we fix $w = 1$ in the rest of the paper.

## C. Sensitivity to training data length

Machine learning techniques leverage training sets to give accurate predictions. The size and variance of a given training set should then impact the prediction performance. In this part, we investigate the sensitivity of our model to the training dataset size. For that, we train RF model with different datasets lengths. In order to have a significant variance in each training
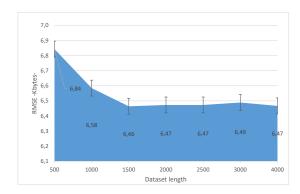
dataset, we shuffle the main dataset, for instance dataset_2_2 with bandwidth dataset; 25 % is leaved out for test. At the beginning of the evaluation process, the dataset consists of 500 samples, i.e. training set length contains 375 samples. The size is then increased gradually for each test. Based on data_2_2, Fig. 8 depicts the prediction errors in terms of RMSE for all the tested sizes with $w = 1$ and $i = \delta t$. The main RMSE for all the tests is inferior to 6.84 Kbytes. The performance improves when increasing the dataset size from 500 to 1500 samples. Then, the prediction error remains between 6.46 Kbytes and 6.49 Kbytes for Larger training dataset sizes. This shows that the model becomes insensitive to training data length when it contains more than 1500 samples.

## VII. CONCLUSION

In this article, we investigate the uplink bandwidth prediction in cellular network and the radio phenomena impact on high-level QoS metrics, i.e. bandwidth. For that, a testbed is deployed in an anechoic room where the radio phenomena are controlled. This allows a clear analysis of encountered behaviors in bandwidth variation. Exhaustive benchmark of lower layers measurements/metrics is performed to constitute dataset reflecting the network response to radio environment variation. In order to predict the received bandwidth at a granularity of 100 ms, machine-learning techniques are used, mainly random forest and linear regression. Nested cross-validation is used for each case study to generalize the obtained error predictions. Different radio scenarios are tested, where the testbed complexity is increased gradually. The analysis of bandwidth variation per radio scenario exhibits the strong correlation between bandwidth and radio environment variation. For each scenario, the sensitivity of the models to forecast window, lag size and training data length are investigated. The model shows insensitivity when predicting. The two machine learning techniques lead to accurate predictions, especially random forest model. Therefore, the obtained results confirm the ability of the developed method to provide accurate uplink predictions in a real environment with very versatile devices and users.

REFERENCES

[1] F. Lu, H. Du, A. Jain, G. M. Voelker, A. C. Snoeren, and A. Terzis. CQIC: revisiting cross-layer congestion control for cellular networks. In Proc. of HotMobile, 2015.

[2] ] D. Koutsonikolas and Y. C. Hu. On the feasibility of bandwidth estimation in 1xEVDO networks. In Proc. of ACM Mobicom International Workshop on Mobile Internet Through Cellular Networks (MICNET), September 2009.

[3] Navid Nikaein, Mahesh K. Marina, Saravana Manickam, Alex Dawson, Raymond Knopp, and Christian Bonnet, Openairinterface: A flexible platform for 5g research, SIGCOMM Comput. Commun. Rev., vol. 44, no. 5, pp. 3338, Oct. 2014.

[4] K. Winstein, A. Sivaraman, and H. Balakrishnan. Stochastic Forecasts Achieve High Throughput and Low Delay over Cellular Networks. In Proc. of Networked Systems Design & Implementation (NSDI), pages 459472, 2013.

[5] https://www.continental-automotive.com

[6] ] X. Liu, A. Sridharan, S. Machiraju, M. Seshadri, and H. Zang. Experiences in a 3G Network: Interplay Between the Wireless Channel and Applications. In Proc. of ACM MobiCom, pages 211222, 2008.

[7] C Yue, R Jin, K Suh, Y Qin, B Wang, W Wei ,LinkForecast: Cellular Link Bandwidth Prediction in LTE Networks- IEEE Transactions on Mobile Computing, 2017

[8] A. Samba, Y. Busnel, A. Blanc, P. Dooze, and G. Simon. Instantaneous throughput prediction in cellular networks: Which information is needed? In IFIP/IEEE International Symposium on Integrated Network Management (IM), May 2017.

[9] R. Margolies, A. Sridharan, V. Aggarwal, R. Jana, N. K. Shankaranarayanan,V. Vaishampayan, and G. Zussman. Exploiting mobility in proportional fair cellular scheduling: Measurements and algorithms. In Proc. of IEEE INFOCOM, 2014.

[10] The Connected Vehicle: Big Data, Big Opportunities: https://www.sas.com/content/dam/SAS/en_us/doc/whitepaper1/connected-vehicle-107832.pdf

[11] WEISBERG, S. (1980). Applied Linear Regression. Wiley, New York.YE, J. (1998). On measuring and correcting the effects of data mining and model selection. J. Amer.Statist. Assoc. 93 120131.

[12] R. Kwan, R. Arnott, R. Trivisonno, and M. Kubota, On pre-emption and congestion control for LTE systems, in Proc. IEEE VTC, Sept. 2010

[13] M.-A. Phan, R. Rembarz, and S. Sories, A capacity analysis for the transmission of event and cooperative awareness messages in LTE networks, 18th World Congress on Intelligent Transport Systems, Orlando, USA, Oct. 2011.

[14] A. Vinel, 3GPP LTE Versus IEEE 802.11p/WAVE: Which Technology Is Able to Support Cooperative Vehicular Safety Applications, IEEE Commun. Letters, vol. 1, no. 2, Apr. 2012, pp. 125-28.

[15] R. Trivedi, M. Patel, "Comparison of Different Scheduling Algorithm for LTE," International Journal of Emerging Technology and Advanced Engineering, Vol. 4, Issue 5, May 2014, pp. 334-339.

[16] Zeki Suac, M., Pfeifer, S., arlija, N. (2014), A Comparison of Machine Learning Methods in a High-Dimensional Classification Problem, Business Systems Research, Vol. 5 No. 3, pp. 82-96.

[17] Geisser, S. (1975). The Predictive Sample Reuse Method with Applications. Journal of the American Statistical Association, 70(350), 320-328. doi:10.2307/2285815

[18] A. Chakraborty, V. Navda, V. N. Padmanabhan, and R. Ramjee. Coordinating cellular background transfers using LoadSense. In Proc. of ACM MobiCom, pages 6374, 2013.

[19] J. Friedman, T. Hastie and R. Tibshirani, The Elements of Statistical Learning, Springer Series in Statistics, 2001.

[20] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In IJCAI, 1995.

[21] S. Varma, R. Simon Bias in error estimation when using cross-validation for model selection. BMC Bioinformatics, 7 (2006), p. 91

[22] Thierry Parmentelat et al. 2018. nepi-ng: an efficient experiment control tool in R2lab. In ACM WiNTECH 2018-12th ACM International Workshop on Wireless Network Testbeds, Experimental evaluation& CHaracterization. 18

[23] L.Breiman,Random forests, Machine learning,vol.45, no.1,pp.532,2001.

[24] Bolat, E (2003). Study of OFDM Performance Over AWGNn Channels. B. Sc. Project, Department of Electrical and Electronic Engineering, Eastern Mediterranean University.

[25] 3GPP, "Evolved Universal Terrestrial Radio Access (EUTRA); Physical layer procedures for control (Release 15),"

[26] S. Singh, Quality of Service Guarantees in Mobile Computing", Journal Computer Communications, Vol 19(4), pp. 359-371 (April 1996). TS 38.213, Dec. 2017.

[27] www.3gpp.org: 3GPP Technical Specification 36.211, 36.213, 36.321.

[28] 1 LT. Jolliffe, Principal Component Analysis (Springer, New York, 1986).