



Cellular Uplink Bandwidth Prediction Based on Radio Measurements

Imane Oussakel, Philippe Owezarski, Pascal Berthou

► **To cite this version:**

Imane Oussakel, Philippe Owezarski, Pascal Berthou. Cellular Uplink Bandwidth Prediction Based on Radio Measurements. MobiWAC 2019, Nov 2019, Miami beach, United States. hal-02315083v2

HAL Id: hal-02315083

<https://hal.laas.fr/hal-02315083v2>

Submitted on 6 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Cellular Uplink Bandwidth Prediction Based on Radio Measurements

Imane Oussakel, Philippe Owezarski, Pascal Berthou
LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France
{imane.oussakel,philippe.owezarski,pascal.berthou}@laas.fr

Abstract—In 4G 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 phenomenon, 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 radio metrics are key features whose evolutions can help predicting 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 study the correlation between radio noise and throughput in uplink transmissions. Based on radio measurements, the main supervised machine learning algorithms are used, such as Random Forest and Support Vector Machine to predict the uplink received bandwidth. For a specific user service, we are able to predict the end-to-end received bandwidth, i.e. the amount of received data on the server side during a specific period at a very low scale of 100 ms. Results also prove that uplink bandwidth predictions are less accurate compared to bandwidth prediction for downlink based on radio measurements.

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, the Intelligent Transport System - ITS - emphasizes many services relying on car communication system. In such system, the vehicle is connected to pedestrians, other vehicles and the ITS servers via wireless networks. Many car companies have been engaged in a research effort to tackle the upcoming ITS issues. For instance, Continental Digital Service France (CDSF) and LAAS-CNRS have launched a collaboration in the framework of the eHorizon project (2017-2021) to cope with ITS systems [6]. As cellular networks are well known for their high throughput and low latency, especially the 4G standard, recent researches endorse the use of LTE-A for the middle term services of connected cars. In such communication, most applications send data (i.e. data uploading) to remote internet servers through LTE-A interface. Thus, contrary to the Human-to-Human applications where downlink (DL) transmission dominates, cars services use mostly uplink (UL) communication. For instance, automotive data analytics [13] forecast a production of up to 30 terabytes of data per day by an average car. Therefore, the conventional roles are inverted: when all the vehicles are sending their information, a large amount of data must be received by the

ITS server with a high quality of service (QoS). Eventhough, 4G is expected to satisfy the high demand of high throughput, several researches show that LTE suffers from congestion [15], [21], especially in UL with the growth of machine traffic. This implies a QoS degradation, and sometimes, the inaccessibility to the network, that limits the envisaged applications and services.

Over years, researchers were focused on enhancing downlink (DL) transmission as only DL was challenging. They improved DL schedulers and congestion control mechanisms while taking predicted throughput as an input parameter [1], [2]. For that, DL LTE throughput prediction has been the ultimate goal of research, as it improves many use cases, such as video streaming quality. Although schedulers and congestion control mechanisms are different in UL and DL, we forecast an improvement of UL LTE-A in a similar way as DL, by incorporating UL predicted throughput. Machine application servers may avoid saturation based on the forecasted bandwidth, hence reducing UL QoS degradation. Thereby, UL bandwidth prediction is valuable in the machine-to-machine context.

In wireless networks, throughput is one of the high performance metrics that changes rapidly depending on the environmental situations. However, radio phenomena affect the different performance metrics. It is increasingly hard to understand how such phenomena degrade the QoS, as they are uncontrolled. In this article, we tackle one radio phenomenon, i.e. noise, to understand how it influences the throughput performance in uplink 4G. In the framework of eHorizon project, we study the case of a connected car transmitting a fixed amount of data to a specific server based on LTE-A network. The aim is to estimate/predict the uplink-received bandwidth by the server. To that end, a 4G testbed is deployed, where the radio propagation is controlled in an Anechoic Room (AR) while injecting a specific noise profile. We investigate our work on UL bandwidth prediction based on radio measurements over a small time granularity, from 100 ms to 1 s scales. Different machine-learning (ML) algorithms are tested to validate the impact of wireless channel on bandwidth prediction in UL transmissions.

The remaining of this paper is organized as follows. Section II points out the 4G 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.

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.

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. 4G is known as a flexible standard as it proposes the use of different bandwidths {3,5Mhz, 4 MHz, 5 MHz, 10 MHz, 20Mhz} and can be deployed on many frequency bands. 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 subcarrier. This transmission of multiple data symbols in a parallel manner leads to a high PAPR (Peak to Average Power Ratio). PAPR causes a high-energy consumption for the transmitter. For uplink transmissions, the efficiency of power amplifier becomes crucial as the UE has a limited battery power. To avoid this OFDMA drawback in uplink, the 4G deploys the single-Carrier Frequency Division Multiple Access (SC-FDMA), where data symbols are transmitted in series and each symbol is carried by a wider bandwidth. Such difference with downlink motivates our investigation on uplink.

The radio transmission is based on either Time Division Duplex (TDD) or FDD (Frequency Division Duplex) mode. Different frequency bands are used for downlink and uplink in FDD; in contrast, the uplink and downlink in TDD mode are separated in time while sharing the same frequency band. For each duplex mode, a frame structure type has been defined. For instance, in FDD, frame type 1 is used in Half and full duplex mode. As shown in fig. 2, the frame-type 1- 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 subcarriers spaced by 15 kHz when a simple Cyclic prefix (CP) is used and spaced by 7.5 kHz in the case of extended cyclic prefix. Then, data symbols are transmitted over one or

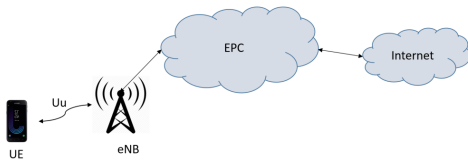


Fig. 1: Basic 4G architecture.

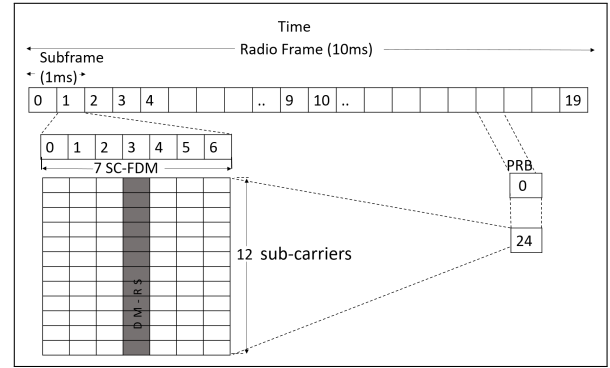


Fig. 2: Uplink resource grid for one slot.

several resource blocks. Each resource block holds six or seven SC-FDMA data symbols depending on the used cyclic prefix. In the physical layer, the LTE-A radio resource is referred as Physical Resource Block (PRB) that forms 180 kHz x 0.5 ms. Our measurements are mainly based on the specified PRB. In time domain, the PRB contains seven SC-FDMA data symbols expanded over 12 subcarriers in frequency domain. The eNB allocates for each user 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 [22]. Thus, based on the used bandwidth, a defined number of useful resource blocks is available. 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 instance, for uplink transmissions at Physical layer level, 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. 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 QoS degradation. 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 for each side of the network. The difference between the UE and eNB composition results in many dissimilarities between downlink and uplink measurements. As we are interested in uplink communication, we cover in more details the main uplink measurements 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. Kout-

sonikolas et al. [3] reveal the ineffectiveness of using those techniques in cellular networks as they are characterized by the large short-scale fluctuation of bandwidth. Although, [5]- [9] 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 radio measurements 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 [16] classifies bandwidth into high and low categories in order to increase user equipment coordination efficiency for transmission in cellular background. Margolies et al. [12] investigate the reproducibility of signal quality over the same path to predict throughput based on user trajectory tracking. Authors of [10], [11] proposed the use of machine learning techniques, especially Random Forest, to predict instantaneous throughput in LTE-A networks. However, A. Samba et al. [11] 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 $\pm 20\%$. C Yue et al. [10] 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 $\pm 10\%$ 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. In fact, contrarily to DL that uses OFDMA reducing/vanishing the inter-symbol interference (ISI), UL 4G uses SC-FDMA, where data symbols are transmitted in series and each symbol is carried by a wider bandwidth. Hence, SC-FDMA is susceptible to ISI, which 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 receiver (eNB). This difference with downlink motivates our work on uplink performance study to investigate the possibility of estimating/predicting the bandwidth in SC-FDMA based systems, that could be different from related work focusing on downlink only. For estimating/predicting UL traffic, we apply mainly Support Vector Machine (SVM) [14] and Random Forest (RF) [8] as machine learning techniques to predict the UL bandwidth based on Radio measurements. Both algorithms provide good predictions in different application domains with

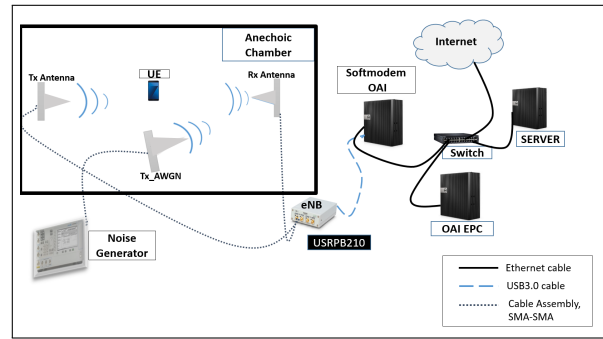


Fig. 3: LTE-A Testbed deployment.

accurate time series and are insensitive to high dimensional feature spaces [19], which corresponds to the case study in this article.

IV. TESTBED DEPLOYMENT

The testbed objective is to limit the uncontrolled effects during a wireless communication, and study the impact of disturbing features on high level QoS metrics. The wireless system is LTE-A as already mentioned in the paper. Fig.3 schematizes the testbed deployment. Openairinterface (OAI) software based platform [4] is implemented. It spans the full LTE 3GPP protocol stack (including features from LTE-Advanced and LTE-Advanced-Pro) for both radio access and core network. For a full access to radio measurements, the 4G network is based on the Software defined radio (SDR). The OAI softmodem is connected with a hardware platform for SDR: USRP B210. The later is connected to a host computer to perform processing, and then connected to a PC running core network, and accessing the internet and a server. Cars are going to be connected using a 4G sim card in their system. To emulate a connected car communication system, a commercial UE (Samsung Galaxy J3 2017) is used and controlled remotely. To simplify the preliminary experimentations, the network transmission mode is Single Input Single Output (SISO). The eNB antennas (Tx and Rx antennas) and the UE are placed inside an AR, where the RF propagation is controlled thanks to the microwave absorbers materiel on the walls, scattering any wireless signal that comes across. Then the room inside is free from any multipath phenomena and radio perturbation or degradation. The testbed works on frequency band 7 (regulated to be used for ITS communications) with 5 MHz bandwidth and uses Frequency Division Duplex (FDD) mode, which corresponds to the traditional and stable version of OAI platform with USRPB210. As our objective is to observe the impact of radio perturbation on the bandwidth variation, the server and the intermediate PCs are well provisioned for not behaving as a bottleneck during communication. Throughput deterioration is mainly caused by environment phenomena.

IPERF3 generates traffic at the UE side and IPERF3-server monitors throughput reception in the server. As TCP (Transport Control Protocol) changes the transmission window

based on the perceived packet losses in the window, we instead use UDP (User Datagram Protocol) in order to have a fixed transmission amount of data during the whole test duration. Given this, any observed bandwidth degradation is essentially due to channel quality variation. Using a speed test throughput application in the UE, the maximum UL data-rate achieved in the testbed is around 8Mbps. Further, the traffic consists of a UDP flow with a constant throughput of 8 Mbps, which allows the use of all the available PRBs. The size of the packets is set to 1350 bytes to avoid any segmentation during transmission.

Noise profile: 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. Thus, in time varying scenarios, the received signal amplitude undergoes rapid fluctuation that is often modeled as a random variable with a particular distribution. In this paper, 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 [18]. 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 developed an assisted labVIEW program on the noise generator. Given an interval of maximum and minimum noise levels, each 10s the noise level takes randomly a value in the specified interval. The programmed step for noise level change (10s) 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. Also, the abrupt changes in noise levels during the transmission tend to reflect the real environments, where the user's mobility across different shadowers leads to aggressive/alleviated signal attenuation. The noise is injected inside the anechoic room using a signal generator with a directional antenna toward the eNB receiver. The bandwidth is fixed to 5Mhz to scramble the full UL bandwidth.

V. PREDICTION METHODOLOGY

Given the LTE testbed in the anechoic room, we are able to generate 4G traffic with noise perturbation. During the whole test duration, noise level changes by the programmed step (10s). This way, samples of same duration for every noise level are generated. In this section, we present the prediction methodology.

A. Data collection

The eNB performs different radio measures in order to decode the received data and adapt to channel variation. With SDR at eNB side, we are able to collect all the performed eNB measurements (as depicted by the 3GPP standard), especially

from the lower layers. After a benchmark of physical layer measurements, we present in the following the main radio measures having a minimum correlation with bandwidth variation:

- RIP: Received Interference Power measured within the bandwidth of each PRB. The eNB measures the noise power over the PRBs each 1 ms.
- UL_RSSI: Uplink Received Signal Strength Indicator. It measures the total wideband received power over the full bandwidth - 5 MHz - including noise and interference. UL_RSSI states the quality in the cell. Too low RSSI reflects the inability of the cell to communicate with any UE, while too high value indicates a high level of interference in the cell.
- 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: $SNR_{dB} = 10 \cdot \log_{10}(P_{signal}/P_{noise})$. It is measured for each received PUSCH holding UE's data and PUCCH containing the UE's control signaling information. Let's denote PUSCH_SNR and PUCCH_SNR the corresponding metrics.
- Rx_power: The received power. It measures the received power in the eNB based on the demodulation reference signal (DMRS), which is used to get a coherent detection and demodulation of UL channels.

One of our objectives is to compare the bandwidth predictions over different time granularity. The bandwidth measurements are performed in a discrete time manner, each δt . As the minimum time report interval in IPERF3 is 100 ms, we fix $\delta t=100ms$. The predictions are then made every δt . The eNB measurements are performed per subframe scale (1ms). In order to have a representative measure per δt , we compute the maximum, minimum, and the mean of each measure per δt to construct the datasets. Table I depicts the constitution of datasets used as input for learning algorithms. The first column represents the datasets labels, and the second column explicits the constitution of datasets in the form of samples. We built 5 datasets based on the aforementioned metrics. Four of them are based on a single feature and the last one combines all the features. For instance, dataset_RSSI contains the single UL_RSSI feature. Each sample in dataset_RSSI is composed of the maximum, minimum and mean measured UL_RSSI over δt , i.e. $\{UL_RSSI_min, UL_RSSI_max, UL_RSSI_mean\}$, dataset_All is a combination of the four features.

B. Methodology

In this article, Python scikit-learn library is used for all the tests. GridsearchCV [23] 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

TABLE I: Constitutions of datasets.

Dataset	Sample notation
Dataset_RIP	{RIP_min, RIP_max, RIP_mean}
Dataset_SNR	{PUSCH_SNR_min, PUSCH_SNR_max, PUSCH_SNR_mean, PUCCH_SNR_min, PUCCH_SNR_max, PUCCH_SNR_mean}
Dataset_RSSI	{UL_RSSI_min, UL_RSSI_max, UL_RSSI_mean}
Dataset_Rx_power	{Rx_power_min, Rx_power_max, Rx_power_mean}
Data_ALL	{Dataset_RIP, Dataset_SNR, Dataset_RSSI, Dataset_Rx_power}

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 [7]. Varma and Simon [17] 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, 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. For a given dataset (table I), 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 randomly 25% of the dataset for testing which we denote $X_{test}=\{X_{test}^1, \dots, X_{test}^m\}$ with size m, and 75% for the training set denoted $X_{train}=\{X_{train}^1, \dots, X_{train}^n\}$ with size n, where X_{train}^p and X_{test}^p are the p^{th} samples in their corresponding datasets. GridsearchCV 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 $i \in \{1\delta t \dots 10\delta t\}$, 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 measures $(X_{test}^t, X_{test}^{t-1}, \dots, X_{test}^{t-w})$ to predict for each sample X_{test}^t from X_{test} the corresponding bandwidth $\hat{B}W_{(i,w)}^t$, $i \in \{1..10\}$. For example, using Dataset_RIP with $i=1$ and $w=2$, M uses the current and

Algorithm 1 Bandwidth prediction from $dataset_i$

Input: Historical radio measurements with lag size w , M : optimized prediction model.
Output: $\hat{B}W=[(\hat{B}W_{(1\delta t,w)}^1, \dots, \hat{B}W_{(i,w)}^1), \dots, (\hat{B}W_{(1\delta t,w)}^m, \dots, \hat{B}W_{(i,w)}^m)]$: the predicted bandwidth.
for t in range $(1, m)$ **do**
 for i in range $(1, m)$ **do**
 $(\hat{B}W_{(1\delta t,w)}^t, \hat{B}W_{(2\delta t,w)}^t, \dots, \hat{B}W_{(i,w)}^t)=M(X_{test}^t, X_{test}^{t-1}, \dots, X_{test}^{t-w})$
 X_{test}^t is the value of a feature in time t .
 end for
end for
return $\hat{B}W$

two past time units of $\{RIP_min, RIP_max, RIP_mean\}_{test}^t$: $[\{RIP_min, RIP_max, RIP_mean\}_{test}^t, \{RIP_min, RIP_max, RIP_mean\}_{test}^{t-1}, \{RIP_min, RIP_max, RIP_mean\}_{test}^{t-2}]$ to predict the upcoming bandwidth in 100ms.

Therefore, for predictions evaluation, we compare the predicted bandwidth $\hat{B}W_{(i,w)}^t$ with the received $BW_{(i,w)}^t$ 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: $RMSE = \sqrt{(1/n * \sum_i^n (y_i - \hat{y}_i)^2)}$ where $y_1 \dots y_n$ are the actual values and $\hat{y}_1 \dots \hat{y}_n$ 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.

VI. UPLINK BANDWIDTH PREDICTION EVALUATION

This part presents and discusses the results obtained on bandwidth prediction using machine learning techniques. Five datasets are tested with two ML techniques RF and SVM. Because of place constraint, we only present the results with RF model; the results for SVM are pretty similar. The evaluation is based on the performance metric RMSE.

A. Sensitivity to forecast window

In order to exhibit the influence of window forecast size on predictions, each dataset from table I is evaluated over $i\delta t$, with $i \in \{1, \dots, 10\}$ and $w = 0$. That is, we predict bandwidth based on the instantaneous radio measurements. The forecast window changes from $1\delta t$ to $10\delta t$. The maximum, minimum and mean observed RMSE with K_2 -fold CV is represented in fig. 4. For each dataset, dashes represent the maximum and minimum of the observed RMSE per δt , and the line links the mean of the observed metric per δt . The predictions errors range from 29 *Kbytes* to 36 *Kbytes*. In fact, the lower RMSE value is obtained for predictions based on dataset_Rx_power, which doesn't exceed 31 *Kbytes*. Dataset_SNR reaches 36.5 *Kbytes* of RMSE for a window forecast of $1\delta t$. We point that the minimum, maximum and mean received bandwidth during the experimentation are 0 *Kbytes*, 128 *Kbytes* and 51.1 *Kbytes* respectively. When combining all the features in dataset_all, the RMSE is not reduced compared with dataset_Rx_power. It is surprising that the SNR achieves the lower scores over the tested features, eventhough it combines the strength of the received signal and the noise measure.

In order to figure out whether the cause is the length of training samples per noise level, another test is performed,

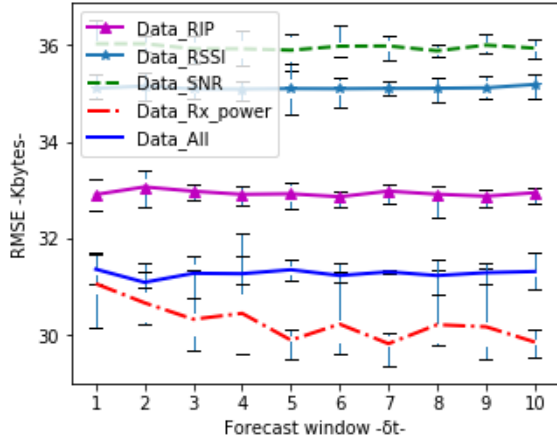


Fig. 4: Impact of forecast window.

where the time step is 60s instead of 10s. Hence, 600 samples are available for each noise level, and a total of 26 noise levels are tested. The same methodology is applied, and similar results of RMSE are obtained (fig.4). It clearly exhibits that the cause behind the observed values of RMSE is not the low number of training samples per noise level. Further, the overall received signal power could be considered the most relevant metric of channel quality in an environment with random noise variation. It is worth noting that the obtained UL predictions are less accurate compared with DL predictions based radio measurements, already mentioned in part III. Over the different forecasted windows, there is no major change in terms of RMSE, thus exhibiting the insensitivity of the model to the forecast window.

B. Prediction bias

In order to generalize the estimated error predictions and be sure that the 10-fold nested CV doesn't introduce any bias during the prediction stage, we study CV bias with varying folds. Different K values for the K-fold method are tested. A small value of K forms folds containing multiple noise level transitions, while higher values of K forms folds with a low number of samples per noise level. For $K=26$, the number of folds is exactly the number of noise levels. In this part $K_2=K_1$ for the nested CV. Fig. 5 shows the distribution of the observed RMSE per CV size for a forecast window of $1\delta t$ with dataset_SNR. For each K-fold, a vertical boxplot is drawn. It consists of a box from the lower quartile of the observed RMSE to the upper quartile, with a crossbar in the mean of RMSE. The upper and lower fences outside the box represent respectively 95% and 5% of the estimated RMSE. For all the tested K-folds, the mean RMSE is around 36Kbytes. Then, the generalized prediction error for dataset_SNR in an environment with random noise is approximately 36Kbytes.

C. Sensitivity to lag size

In this part, the importance of having numerous past radio measurements for good predictions is analyzed. The main lag

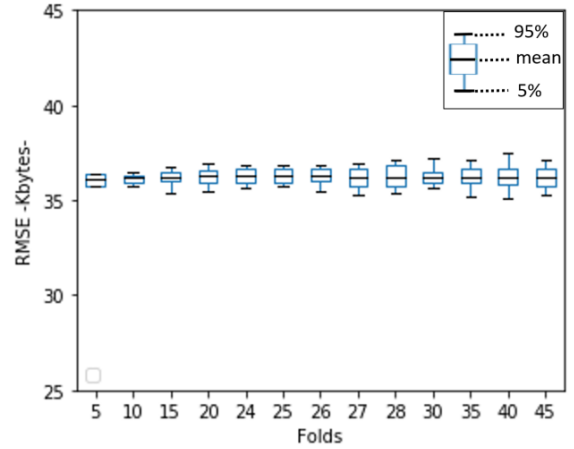


Fig. 5: Cross-validation bias.

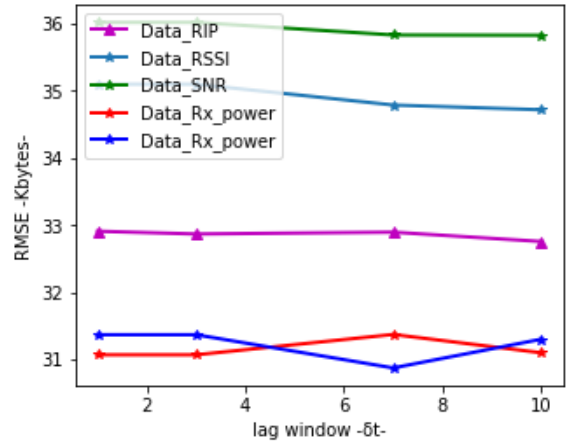


Fig. 6: Impact of lag size w .

sizes w tested are 3, 7 or 10, i.e. using the past three, seven and ten δt measurements. Fig. 6 shows the obtained RMSE for a window forecast of δt with the three values of w for all the datasets. Adding the past measures improves smoothly the predictions based on dataset_SNR and dataset_RSSI. On the other hand, historical information doesn't necessary improve the performance, however, it might degrade it. The reason behind that might be the random and brutal change of the signal attenuation caused by the generated noise profile. In fact, it reflects real radio environment, where the upcoming attenuation is uncontrolled. Hence, forthwith the past measurements become obsolete. Therefore, in the rest of the paper, we fix $w = 0$.

D. Sensitivity to training data length

Machine learning techniques leverage on 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 various datasets of different lengths. In order to have a significant

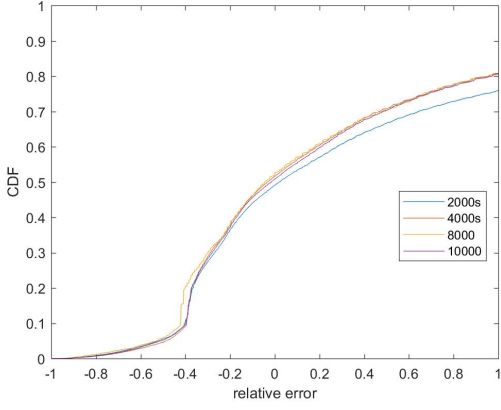


Fig. 7: Impact of training dataset length.

variance in each training dataset, we shuffle the main dataset, i.e. dataset_Rx_Power with bandwidth dataset; 30% is leaved out for test. At the beginning of the evaluation process, the training set consists of 2000 samples. The size is then increased gradually for each test. Based on data_RX_Power, Fig.7 depicts the CDF of the relative prediction error of all the tested sizes with $w = 0$ and $i = \delta t$. The relative error is obtained from the difference between the predicted and actual bandwidth divided by the actual bandwidth. The performance improves when increasing the training dataset size from 2000 to 4000 samples. Larger training dataset sizes don't affect the predictions, and 58% of errors are within $\pm 40\%$. This shows that the model becomes insensitive to training data length when it contains more than 4000 samples.

E. Sensitivity to noise level linearity

In the previous parts, noise level was changing randomly during the test duration. Hence, the selected X_{train} contains a significant amount of random noise level variations. In order to study the impact of the presence of such transitions in the training set, we set up another experimentation process. Keeping the same testbed, we changed the assisted LabVIEW program to control remotely the noise generator. This time, given an averaged granularity of 0.1 dBm as a step for noise level variation, the noise level is increasing every 10 s by the programmed step during the whole test duration (800s). This later experiment is aimed at providing a dataset less complex than the previous one in order to investigate the impairments introduced by the random change of noise levels on the system. The same process of data collection in part V.B is followed, and the two machine learning techniques, RF and SVM, are applied. We evaluate each dataset from table I over $i\delta t$, with $i \in \{1, \dots, 10\}$ and $w = 0$. During bandwidth predictions, the forecast window changes from $1\delta t$ to $10\delta t$. The maximum, minimum and mean observed RMSE with K_2 -fold CV is represented in fig. 8. For each dataset, dashes represent the maximum and minimum of the observed RMSE per δt , and the line links the mean of the observed metric per δt . Based on dataset_SNR and dataset_RIP, the RMSE

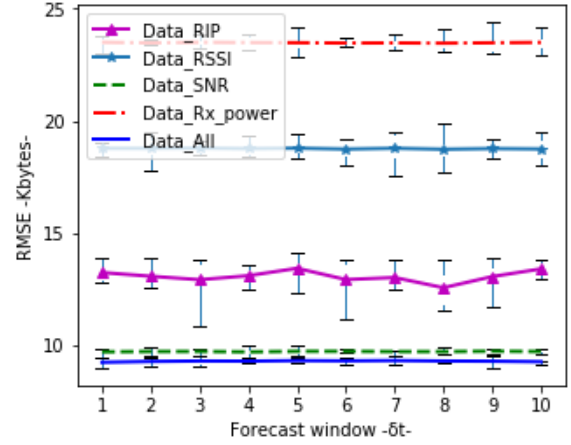


Fig. 8: Impact of noise level variation.

doesn't exceed 9.7 Kbytes and 13.4 Kbytes respectively for all the forecasted windows. It is to be noted that the minimum, maximum and mean received bandwidth are 1.32 Kbytes, 133 Kbytes and 75.04 Kbytes respectively. Comparatively, based only on the received power (dataset_Rx_power), poor predictions are obtained; an RMSE of 23.5 Kbytes is observed for a window forecast of $1\delta t$. Dataset_all that combines all the features, gives pretty similar results to the ones obtained with dataset_SNR. This is explained by the presence of SNR measurements in dataset_all. Therefore, this remarkable decrease of RMSE with the datasets containing linear generated noise levels is essentially due to the absence of randomness in noise level variation. In other words, the lower observed values of RMSE in the linear testbed are essentially due to the consecutive noise level variation in the training set for the model. Hence, the randomness in noise level variation leads to an increase in terms of RMSE, as expected.

F. Sensitivity to high noise level

The noise level changes each 10s by a step of 0.1 dBm during the whole test duration. After several tests, the maximum noise level is set to -12 dBm and the minimum to -20 dBm. Low values perturbate smoothly the radio environment, while the higher values, around -12 dBm, result in a packet loss above 90% ($\pm 5\%$) in the cell. In this part, we aim to see whether the prediction in linear testbed is sensitive to the high noise levels. To that end, we split the main dataset, for example, dataset_RIP with bandwidth, into subsets of 1000 samples. Then, each subset contains 10 noise levels of 0.1 dBm difference, i.e. between -17.0 dBm to -16.0 dBm. We apply the methodology in V.B on each subset and plot in fig.9, in 3D, the variation of the perceived RMSE over $i\delta t$ with $i \in \{1..10\}$ and per noise level. The RMSE is almost constant from -19 dBm to -16 dBm, then it increases brutally for the two levels of noise, -15 dBm and -14 dBm. Checking the corresponding bandwidth subsets, a high level of bandwidth variation is observed for these two levels. The increase of RMSE for those levels could be explained by the insufficiency of training samples, as only

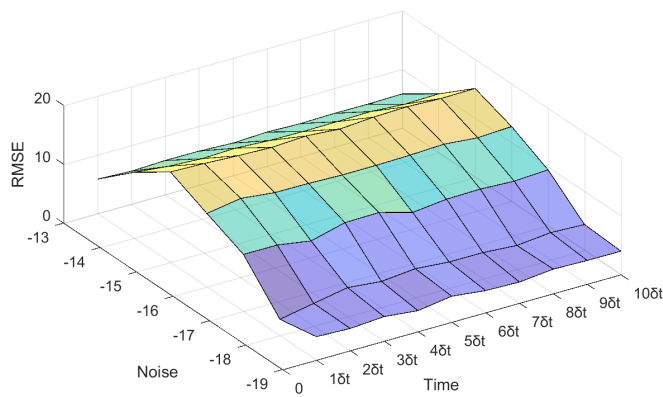


Fig. 9: Sensitivity to high noise level.

700 samples were used. This result points out the necessity of taking advantage of large training sets for each noise level to achieve the insensitivity to high noise levels.

VII. CONCLUSION

In this article, we investigated 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, and then noise is injected in a controlled manner to scramble the uplink transmissions. This allows a clear analysis of encountered behaviors. The noise profile is generated in a manner to emulate the amplitude attenuation variability present in real radio phenomena. Exhaustive radio measurements are performed to collect the main measures reflecting the uplink channel quality. In order to predict the received bandwidth at a granularity of 100 ms, machine learning techniques are used, mainly random forest and support vector machine. Nested cross-validation is used for each case study to generalize the obtained error predictions. As the features choice for machine learning is crucial, the sensitivity of the model to forecast window, lag size, high noise level and training data length are investigated for each feature. The model shows insensitivity in the different cases. We have emulated two types of environments and have shown that in an environment where noise level changes gradually, accurate predictions are achieved. On the other hand, when noise level changes brutally from lower to higher values and vice versa, thus emulating real radio communication systems, high prediction errors are obtained. Therefore, contrarily to downlink, where only radio measurements are sufficient to predict accurately the received bandwidth, uplink bandwidth prediction needs more investigation in a real environment with very versatile devices and users. Hence, cellular uplink is becoming the hotbed of improvements.

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

[3] J. 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.

[4] 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. 33–38, Oct. 2014.

[5] 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 459–472, 2013.

[6] <https://www.continental-automotive.com>

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

[8] L. Breiman, "Random forests," Machine learning, vol. 45, no. 1, pp. 5–32, 2001.

[9] 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 211–222, 2008.

[10] 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

[11] 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.

[12] 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.

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

[14] Smola, A.J., Schölkopf, B.: A tutorial on support vector regression. Statistics and Computing 14(3) (August 2004)

[15] 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.

[16] A. Chakraborty, V. Navda, V. N. Padmanabhan, and R. Ramjee. Coordinating cellular background transfers using LoadSense. In Proc. of ACM Mobicom, pages 63–74, 2013.

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

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

[19] Zekić Sušac, M., Pfeifer, S., Šarlija, N. (2014), "A Comparison of Machine Learning Methods in a High-Dimensional Classification Problem", Business Systems Research, Vol. 5 No. 3, pp. 82-96.

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

[21] 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.

[22] 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.

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