Learning from Large-scale Commercial Networks: Challenges and Knowledge Extraction towards Machine Learning Use Cases

Roman Zhohov
Ericsson Research
Linköping, Sweden
roman.zhohov@ericsson.com

Alexandros Palaios
Ericsson Research
Aachen, Germany
alex.palaios@ericsson.com

Philipp Geuer
Ericsson Research
Aachen, Germany
philipp.geuer@ericsson.com

ABSTRACT

Machine Learning (ML) algorithms are proposed to replace conventional algorithms in the area of wireless networking. Many of the suggested algorithms are often based on the data acquired from simulators or small-scale test-beds. We provide a study based on a dataset collected over a large commercial network, and highlight some of the real network dynamics that learning agents need to cope with. Our dataset includes not only measurements from the User Equipment (UE) but also integrates information from the network. Based on the collected data, we highlight some of the aspects that are important for the design of learning agents and discuss potential dataset characteristics that might hinder the learning process. Then we discuss what dataset characteristics can facilitate the deployment of ML algorithms in the real networks. Finally, we showcase how throughput prediction can be implemented by using ML techniques and provide some examples and insights on feature engineering and the training process.

KEYWORDS

Performance Evaluation, QoS, QoE, Artificial Intelligence, Machine Learning, Network Dynamics, Dataset Characteristics

1 INTRODUCTION

In recent years learning agents are being suggested [15] to replace or substantially improve traditional algorithms in the area of wireless communications. Many of those proposals are based on successfully implemented prototypes that are utilizing data collected from simulators [13] or testbeds [7] of limited scale (i.e. limited number of network nodes and user terminals). Very often these datasets are collected only from the UE-side without or with limited information from the network[5],[20],[16], [6],[8]. Such datasets come with inherent limitations on how accurately they can resemble the real network dynamics [10], [14], [17]. Any disparities with real measurements further limit the performance of the proposed algorithms when deployed in commercial networks. Similarly, understanding the real dynamics of commercial networks[21], [11] can potentially improve the robustness of proposed solutions. Having more accurate ML models not only provides better run-time performance but also significantly reduces the re-training periods and removes the overhead of the extra data collection procedures.

ML algorithms strongly benefit from specific dataset characteristics, for example, consistent statistics over different periods, that imply some notion of time stationarity and repeatability in the measurements. In this work, we highlight which of those features can be further utilized to improve the adoption of ML models.

In this paper, we present insights and results that are coming from a long measurement period, extending over several days, and thus providing a more consistent overview of the network dynamics. We describe important dataset characteristics and how ML algorithms shall be used and evaluated in conjunction with use cases.

The remainder of the paper is organized as follows. Section 2 describes the network setup and data collection procedure. In Section 3, we provide an overview of the dataset and present important insights on the network dynamics. In Section 4, we describe how ML was applied for the throughput prediction problem. Section 5 concludes the paper and discusses future research goals.

2 LARGE-SCALE MOBILE NETWORK

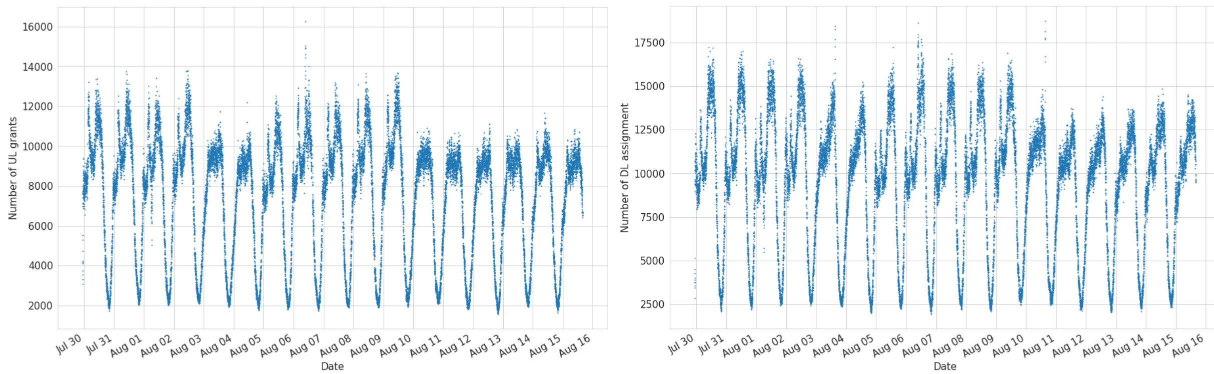
In this section, we describe the network and data collection procedure. Measurements were conducted in the commercial network of one of the biggest operators in Asia. The network covers a large geographical area and consists of more than 100 base stations. The data collection procedure took place for 18 days. Such large-scale data collection is a non-trivial task since it creates a high signaling overhead in the network. As was mentioned earlier, the collected dataset is unique since it includes information not only from the user terminals but also from the base stations. The UE-side measurements include radio measurements, such as RSRP, RSRQ, TA values and traffic reports such as UL/DL (uplink/downlink) throughput and session sizes. Radio reports are triggered based on A4 events. The A4 event is triggered when a neighboring cell becomes stronger than the specified threshold, the detailed description of the A-events is given in the respective technical specifications [1], [2], [3]. The traffic reports are collected at the end of the RRC session. The cell logs from the base stations are collected periodically (every 60 seconds) and they include various cell counters, such as the PRB (Physical Resource Block) utilization, the number of active UEs, the number of UL scheduling grants, the DL scheduling assignments, etc. In total, the dataset includes measurements from more than 400 unique cells with more than 2 million unique UEs.

3 DATA PROCESSING AND ANALYSIS

As mentioned earlier, measurement patterns for calculating the base station counters are periodic, while UE-side measurements

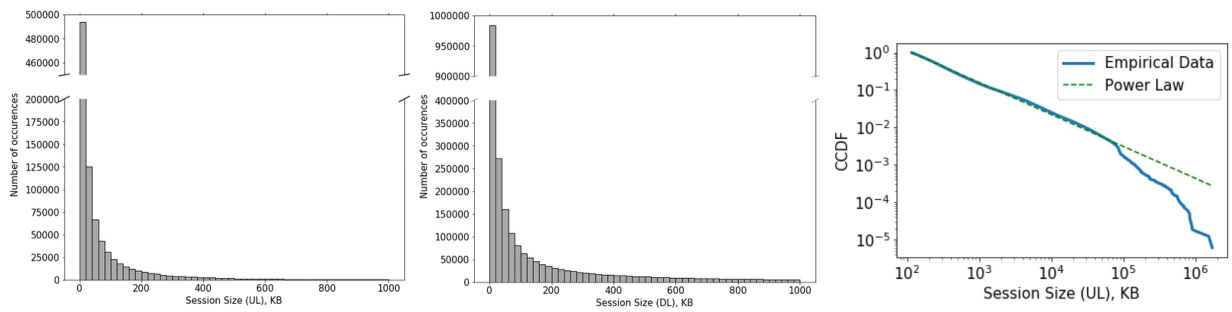
5G-MeMU'21, August 23, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8636-4/21/08…$15.00
https://doi.org/10.1145/3472771.3472773



(a) **(b)**

Figure 1: The number of uplink grants (a) and downlink assignments (b), during the course of the 18 measurements days.



(a) **(b)** **(c)**

Figure 2: The distribution of session sizes for the uplink (a) and the downlink (b). Note that the very long tails have been truncated to allow a more informative plotting. Figure (c) is the CCDF of the session size with both axes in log scale.

are aperiodic (i.e event or session-based). Collected data was preprocessed in order to align cell- and UE-based measurements for each cell. The resulting set of features is given in Table 1.

Table 1: The collected features from the network.

	Features
UE-side features	RSRP, RSRQ, TA
Network-side features	Cell Bitrate in Uplink and Downlink Number of active UEs in Uplink and Downlink Number of Uplink Grants Number of Downlink Scheduling Assignments Scheduling Activity for Uplink and Downlink

Next, we provide some insights on utilization patterns of a large-scale cellular network. Fig. 1 depicts the number of UL grants for all cells which is a good indicator of end-users activity and requests to the network. Specifically, 18 consecutive days of measurements are plotted, which provides good overview of the evolution of dynamics in the real network. The users' activity follows the diurnal cycle of the day and patterns seem to remain constant over the total number of days. The total number of UL grants changes slowly over time, which already hints that reported aggregated values from the network, even with slow sampling, can potentially provide enough information for specific ML tasks. This also confirms previous assumptions in the literature that overall statistics from large operators remain constant if not very slowly evolving [12]. Such slowly evolving patterns and repeatability are important data characteristics that can enable faster adoption of ML algorithms. For example, instead of measuring the network load constantly, one can easily accurately estimate it. Estimating such features can simplify considerably the implementation of ML algorithms since there is no need for storing or sharing data over extended periods.

At the same time, an interesting observation is that there are quite some outliers that can be around 30% higher than the typical peaks, which can cause some issues on ML estimators, if not handled properly [4], [19].

We continue by looking at the UL and DL session sizes. The session size is reported by the UE at the end of each RRC session. Fig. 2 presents an aggregated UE session sizes from all UE-reports. Both figures for the UL and DL show exactly the same trend, an exponentially decaying structure with the large majority of the measurements being very small sessions. Session sizes that are just over 200kB are below than 10% of the collected measurements. The tails of the distribution are relatively long, quite similar to power law

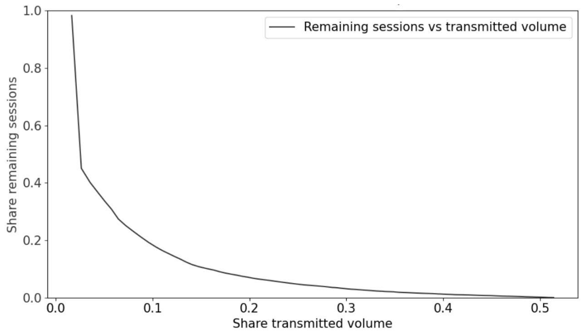


Figure 3: The distribution of the RRC-sessions.

type of models. These models have a probability distribution in the form of $p(x) \propto x^{-a}$ where standard deviation is much larger than the mean. In Fig. 2 we also show the complementary cumulative distribution function (CCDF) of the heavy tailed distribution.

Moreover, the aforementioned session size distribution is found very often in many large-scale cellular networks. We complement the above results presenting the session size distribution from an operator from North America, in Fig. 3. The figure shows the share of transmitted volume over time with respect to share of remaining sessions. For example, 50% of the sessions have a transmitted volume that corresponds to only 5% of the maximum session volume. Fig. 3 confirms the same trend and distribution shape. Not only the session sizes diminish fast but also the session duration. Most of the users' activity is coming in short bursts that are relatively small in size. What seems to be clear though is that large-scale networks exhibit very similar dynamics as the end-users have similar type of activity patterns during the day. The small session sizes are dominating the data collection procedure.

Finally, in this section we present the distribution of users' throughput in UL and DL at the PDCP layer. Fig. 4 shows the UL and DL PDCP throughput distribution. Both figures depict very long distribution tails containing very few measurements even though the "always on" data collection lasted over 18 days. For the DL case, the tail is even longer compared to the UL, but still for both cases, it is clear that the data is highly imbalanced. For higher throughput values there is only a handful of measurements available which makes the learning on such imbalanced data sets very challenging, if not impossible. Longer data collection periods, for more than 18 days, might be needed to compensate to some extent for such an imbalanced data set and provide enough data for all classes. This will reduce the imbalanced dataset problem only to some degree, since capturing largest throughput values leads to unrealistically long capturing periods for the network. For instance, the network collected just 4000 samples for the UL throughput higher than 20Mbps, collecting just 12000 of those will require more than 54 days of collection process running. This is, of course, a good example that getting a good quality data set for ML is not a straightforward nor always a realistic task. Here a smarter data collection procedure might be also needed to tackle the problem.

Table 2: The achieved throughput classes.

Class	Throughput in Mbps
1	0-5
2	5-10
3	10-15
4	15-20
5	20-25
6	25-30
7	30-100

Table 3: The prediction models and the achieved Mean Absolute Error.

Model	MAE
Ordinal Logistic Regression	0.92
Logistic Regression	1.14
Ridge Regression	0.79
Elastic Net	0.80
Random Forest	0.59
Ridge Regression (UE-side features only)	1.09
Random Forest (UE-side features only)	1.18

4 MACHINE LEARNING USE CASE

In this section, we apply ML algorithms in order to predict UE throughput using the collected data. We will specifically focus on the algorithms that have lower complexity. The algorithms' performance will be quantified based on the Mean Absolute Error (MAE). This choice is based on the fact that radio measurements are known to produce strong outliers and few large errors and MAE is less affected from those. In comparison, the Mean Squared Error (MSE) penalizes stronger the large errors [18].

Since collected measurements have large systematic measurement variance (compound effects of the non-directional antennas of the UEs and the fading in the radio environment) to further reduce its impact we reduce the throughput prediction task to a classification problem, predicting a range of achievable rates. We have split the throughput into 5 Mbps wide bins and assigned the label to each class. The last class is larger spanning from 30-100 Mbps. The only exception is for the case of the ordinal regression, as will be also discussed later. The classes are shown in Table 2.

Since we have imbalanced datasets, we use random permutation cross-validation with stratification that enforces similar levels of class distribution in each split of the dataset, which provides more consistent results. For the interested reader, we also mention that we have applied oversampling techniques that did not improve results significantly since the number of measurements is not enough and the generated data could not capture realistically the real measurements' dynamics.

We begin with a multinomial logistic regression since we have a multi-class classification problem. We got an MAE score of 1.12, which means that on average the logistic regression estimates one class larger or smaller than the real class. Since the data can be also considered as ordinal (categorical data with a set order and

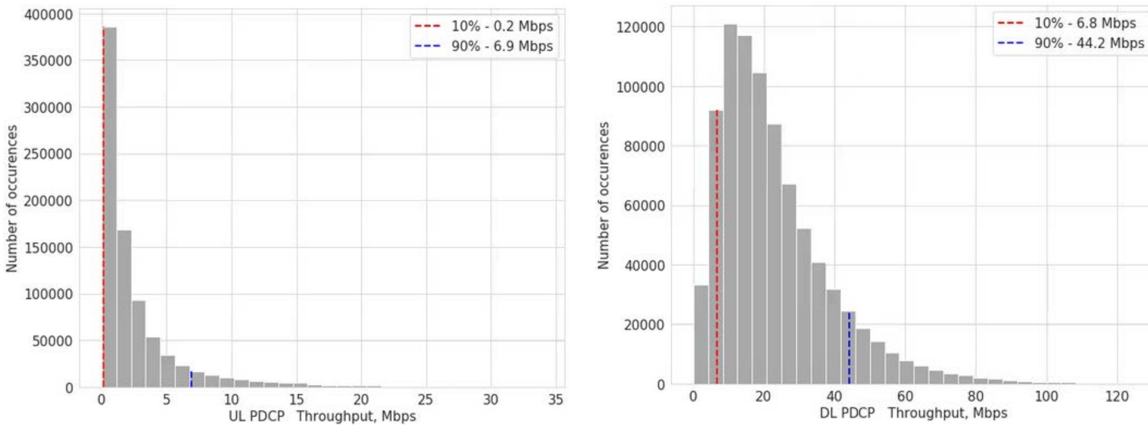


Figure 4: Left: UL PCDP Throughput, Right: DL PDCP Throughput.

distance between the categories), we apply a logistic ordinal regression [9], which is as an extension to the logistic regression. Before applying it we converted the last class to be 30-35 so it is of similar width compared to the other ones. The ordinal regression assumes an input $X \in \mathbb{R}^{n \times p}$ and $y \in \mathbb{N}^n$ as the target values. Then the cumulative probability is modeled as a logistic function $P(y \leq j|X_i) = \phi(\theta_j - w^T X_i) = \frac{1}{1+\exp(w^T X_i - \theta_j)}$ where w, θ are the vectors that are estimated from the data and ϕ is the logistic function defined as $\phi(t) = 1/(1+\exp(-t))$. Compared to the logistic regression there is an extra constraint added so that the hyperplanes that separate the different classes are parallel for all classes. That is the vector w that is common for all possible classes and θ the vector of thresholds. The minimization loss is defined as: $\mathcal{L}(w,\theta) = -\sum_{i=1}^{n} \log(\phi(\theta_{y_i} - w^T X_i) - \phi(\theta_{y_i - 1} - w^T X_i))$ and the gradient of the loss function is calculated

$$\nabla_w \mathcal{L}(w,\theta) = \sum_{i=1}^{n} X_i(1 - \phi(\theta_{y_i} - w^T X_i) - \phi(\theta_{y_i-1} - w^T X_i))$$

as $$\nabla_\theta \mathcal{L}(w,\theta) = \sum_{i=1}^{n} e_{y_i} \left(1 - \phi(\theta_{y_i} - w^T X_i) - \frac{1}{1 - \exp(\theta_{y_i - 1} - \theta_{y_i})}\right) + e_{y_i - 1}\left(1 - \phi(\theta_{y_i-1} - w^T X_i) - \frac{1}{1 - \exp(-(\theta_{y_i-1} - \theta_{y_i}))}\right)$$

The MAE scores for the multinomial ordinal regression is 0.92 which is an improvement compared to the logistic regression with the same complexity. Finally, we include one more algorithm that has comparably low complexity, the linear regression (by converting classification task to a regression one). To avoid overfitting, we apply regularization, where the ridge regressor gives MAE of 0.79 and elastic net performing similarly with MAE of 0.8, performing better than the multinomial ordinal regression. This can be explained probably from the systematic noise radio measurements typically have that makes their distinction to separate classes relatively hard. Various radio phenomena like fast fading can explain such disparities, since measurement reports include time-averaged measurements over a certain time window and cannot indicate the full dynamics and variations of the radio channel. Table 3 refers to the achieved performance from all the models used.

Finally, we look at the ensemble type of algorithms, that are known to deal effectively with outliers but are much more complex . We showcase the performance of Random Forests, which has higher complexity compared to all previous models. The Random Forest algorithm achieved MAE of 0.59 showing the best performance. This achieved performance might be already acceptable for many use cases, specifically given the fact that the measurements are coming from the low-cost transceivers and environment, that is known to be very dynamic, drastically changing over time and space.

Moreover, we would like to stress that applying standard ML metrics can be misleading, especially, when specific use cases are taken into consideration. Metrics need to be linked to the specific use cases and their requirements. For example, predicting less throughput of what is actually achievable might still be acceptable for some applications, since it might satisfy the minimum requirements of the specific use case. On the other side predicting more throughput of what is achievable might cause unsatisfactory QoE levels and abrupt stop of applications.

Even though this might sound trivial, it has not been stressed enough in the literature that often adopts an approach that all estimation errors are equally important. We provide the distribution of MAE in Fig. 5 of all discussed algorithms. Random Forest is performing best with predicting most of the time the correct class, also with more confined prediction errors. As said underestimating the class slightly can still allow specific use cases to run successfully and without interruptions. We contrast that with logistic regression where the error variance is the largest of all models and can completely underestimate or overestimate leading to very poor results. The linear regression algorithms have a better MAE score than the ordinal regression but a much different error distribution. Ordinal regression, for example, seems to have a tendency to not overestimate (although the specific data might have influenced to some extent this). The performance metrics of ML algorithms are more meaningful when presented in conjunction with use cases.

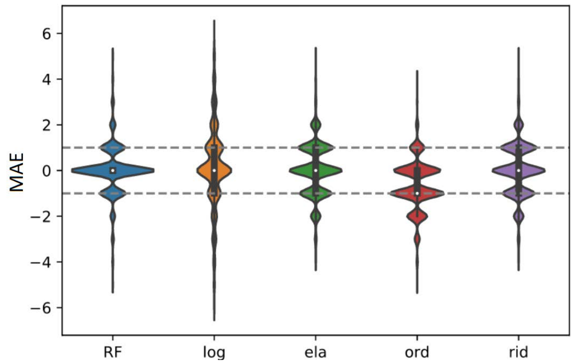


Figure 5: Violin plot of all the estimators used: Random Forest, Logistic Regression, Elastic Net, Ordinal regression, Ridge regression. The dashed lines refer to +1 and -1 points.

Table 4: Feature Importance.

Feature	UE- or Network- feature	RandomForest Importance	Ridge Importance
Cell Bitrate in Uplink (max)	Network	1	4
RSRP	UE	2	1
Cell Bitrate in Uplink (mean)	Network	3	3

4.1 Feature Importance

As was mentioned earlier, most of the literature is based on crowd-sourced data where network-side information is lacking[5],[20],[16],[6],[8]. Here we would like to show some examples of how useful network-side features can be. In Table 4 the most important features are given for the ridge regression and Random Forest models.

We have also repeated the same prediction by using UE-side features only. As shown in Table 3 when no network information is available, prediction performance drops drastically. Even though radio conditions at the UE-side are very important since they define the maximum achievable throughput, at the same time resource availability and load at the base station can limit the throughput due to the fact that resources are shared between all the users in the cell. These results show that the prediction tasks without network information can be mostly good either when the radio conditions at the UE are poor and thus UEs are throughput limited, or when the UEs have good radio conditions and the base stations are low utilized, and thus the base station can provide higher throughput than a UE requests.

4.2 Sensitivity of the results

In this section, we look at the sensitivity of the results in conjunction with the duration that the measurements took place. We try to quantify and give some insights how much time (hours/days) data collection procedures might need to take place. Data collection is an important parameter that is quite often overlooked. Data collection procedures might cause unnecessary high load to the network nodes if collecting reports for very long periods.

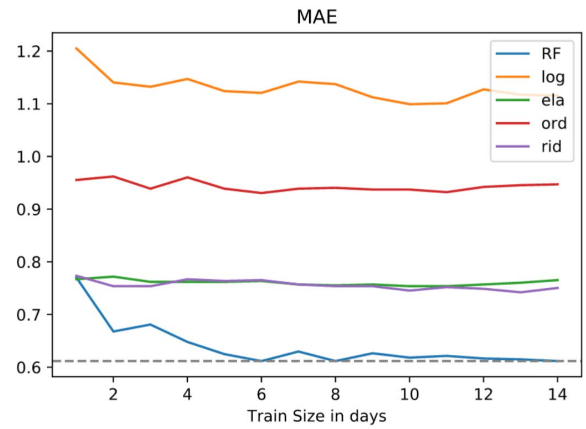


Figure 6: MAE of different algorithms with train size from the first 1 to 14 days, test size always consists of the last 3 days. The dashed line corresponds to the maximum achieved MAE.

In Fig. 1, we have already seen some notions of similarity between the measurement days. We would like to explore in this section if such dynamics can reduce significantly the time needed for data collection. Data collection is not only needed once for the training but it should be repeated in order to update models and monitor their performance.

In Fig. 6, we observe the behavior of each algorithm in terms of performance as the training set increases. We use the measurements from the last three days as a testing data to have a fair comparison. As the number of days increases, performance seems to increase only for one model - Random Forest. Simpler models reach their maximum performance much faster. This is a useful attribute that shows that if used carefully the overhead for data collection can be significantly reduced. At the same time, the Random Forest model improves its performance quite a lot if more data is available for training. In given settings, 6 or more days of data collection might be needed to achieve the aforementioned levels of performance. Random Forest model can handle imbalanced datasets and outliers more effectively. At the end, it is not only about having access to the data, but the data should have specific qualities, like representatively enough data for all classes. This is not guaranteed to be so especially if the data collection procedures do not last for a series of days. The smaller session sizes and achieved throughput typically dominate the datasets.

5 CONCLUSION

In this paper, we have explained some of the challenges of developing ML algorithms using the data collected from the large-scale commercial network. Provided results reveal some of the future challenges that ML algorithms need to cope with, such as heavily imbalanced datasets, systematic measurement noise, and data with large outliers. Even though this might sound rather trivial, it has not been stressed enough in the current literature. Data collection procedures might need to last few days or even weeks to provide a good representation of all possible cases that ML algorithms need to learn to differentiate with. Our results show that regardless of the challenges some promising results might be achieved, especially with better-optimized sampling processes.

Moreover, some of the network dynamics were compared with different operators from different continents, which boosts the confidence in the presented results. Aforementioned network dynamics (network load, traffic patterns, throughput, etc.) are beneficial as an input for the simulator-based studies as it enables more realistic models.

We also tried a few algorithms of low complexity, and explained why conventional performance metrics should be extended and explained in terms of requirements of a particular use case. More complex ML models provide some performance benefits but this comes with a rather steep increase in complexity and longer data collection procedures. For large-scale networks, this can be a very important factor that needs to be considered.

Finally, we discussed the performance gains by utilizing network-side information and not only solely relying on crowd-based UE-side data. Another interesting fact is that the network dynamics at the aggregated level (overall utilization of the base stations) change relatively slow and statistics seem to have a relatively easy trackable nature, that can considerably simplify sampling procedures enabling a much easier adoption of ML models.

In the future, we plan to further expand this discussion and present on how prediction errors should be handled, so new use cases can be enabled even in the presence of large prediction errors.

ACKNOWLEDGMENT

The authors would like to thank Mats Zachrison for providing valuable insights on traffic patterns of commercial networks.

REFERENCES

[1] 3GPP. 2017. *Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification.* Technical Specification (TS) 36.331. 3rd Generation Partnership Project (3GPP). https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2440 Version 14.2.2.

[2] 3GPP. 2018. *Service requirements for enhanced V2X scenarios.* Technical Specification (TS) 22.186. 3rd Generation Partnership Project (3GPP). Version 15.4.0.

[3] 3GPP. 2019. *Evolved Universal Terrestrial Radio Access (E-UTRA); Services provided by the physical layer.* Technical Specification (TS) 136.302. 3rd Generation Partnership Project (3GPP). Version 14.5.0.

[4] S. Ertekin, L. Bottou, and C. L. Giles. 2011. Nonconvex Online Support Vector Machines. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 2 (2011), 368–381. https://doi.org/10.1109/TPAMI.2010.109

[5] A. Ghasemi. 2018. Predictive Modeling of LTE User Throughput Via Crowd-Sourced Mobile Spectrum Data. In *2018 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. 1–5. https://doi.org/10.1109/DySPAN.2018.8610464

[6] K. Kousias, Ö. Alay, A. Argyriou, A. Lutu, and M. Riegler. 2019. Estimating Downlink Throughput from End-user Measurements in Mobile Broadband Networks. In *2019 IEEE 20th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*. 1–10. https://doi.org/10.1109/WoWMoM.2019.8792968

[7] A. Kulkarni, A. Seetharam, A. Ramesh, and J. D. Herath. 2020. DeepChannel: Wireless Channel Quality Prediction Using Deep Learning. *IEEE Transactions on Vehicular Technology* 69, 1 (2020), 443–456. https://doi.org/10.1109/TVT.2019.2949954

[8] Jinsung Lee, Sungyong Lee, Jongyun Lee, Sandesh Dhawaskar Sathyanarayana, Hyoyoung Lim, Jihoon Lee, Xiaoqing Zhu, Sangeeta Ramakrishnan, Dirk Grunwald, Kyunghan Lee, and Sangtae Ha. 2020. PERCEIVE: Deep Learning-Based Cellular Uplink Prediction Using Real-Time Scheduling Patterns. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services* (Toronto, Ontario, Canada) *(MobiSys '20)*. Association for Computing Machinery, New York, NY, USA, 377–390. https://doi.org/10.1145/3386901.3388911

[9] Peter McCullagh. 1980. Regression Models for Ordinal Data. *Journal of the Royal Statistical Society. Series B (Methodological)* 42, 2 (1980), 109–142. http://www.jstor.org/stable/2984952

[10] D. C. Moreira, I. M. Guerreiro, W. Sun, C. C. Cavalcante, and D. A. Sousa. 2020. QoS Predictability in V2X Communication with Machine Learning. In *2020 IEEE 91st Vehicular Technology Conference (2020-Spring)*. 1–5. https://doi.org/10.1109/VTC2020-Spring48590.2020.9129490

[11] A. Palaios, J. Riihijärvi, O. Holland, and P. Mähönen. 2013. A week in London: Spectrum usage in metropolitan London. In *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. 2522–2527. https://doi.org/10.1109/PIMRC.2013.6666571

[12] A. Palaios, J. Riihijärvi, P. Mähönen, V. Atanasovski, L. Gavrilovska, P. Van Wesemael, A. Dejonghe, and P. Scheele. 2012. Two days of spectrum use in Europe. In *2012 7th International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*. 24–29. https://doi.org/10.4108/icst.crowncom.2012.249475

[13] T. Park, J. Han, and E. Hong. 2020. UE throughput guaranteed small cell on/off algorithm with machine learning. *Journal of Communications and Networks* 22, 3 (2020), 223–229. https://doi.org/10.1109/JCN.2020.000020

[14] Caleb Phillips, Douglas Sicker, and Dirk Grunwald. 2013. A survey of wireless path loss prediction and coverage mapping methods. *IEEE Commun. Surveys Tuts.* 15, 1 (1st Quart. 2013), 255–270.

[15] D. Raca, A. H. Zahran, C. J. Sreenan, R. K. Sinha, E. Halepovic, R. Jana, and V. Gopalakrishnan. 2020. On Leveraging Machine and Deep Learning for Throughput Prediction in Cellular Networks: Design, Performance, and Challenges. *IEEE Communications Magazine* 58, 3 (2020), 11–17. https://doi.org/10.1109/MCOM.001.1900394

[16] J. Riihijarvi and P. Mahonen. 2018. Machine Learning for Performance Prediction in Mobile Cellular Networks. *IEEE Computational Intelligence Magazine* 13, 1 (2018), 51–60. https://doi.org/10.1109/MCI.2017.2773824

[17] Janne Riihijärvi, Daisy Maibam, and Petri Mähönen. 2017. Impact of model uncertainties on quantitative evaluation of interference risks. In *Proc. IEEE Int. Symposium on Dynamic Spectrum Access Networks (DySPAN)*. 1–2.

[18] Cort J. Willmott and Kenji Matsuura. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research* 30, 1 (2005), 79–82. https://www.int-res.com/abstracts/cr/v30/n1/p79-82

[19] X. Yang, G. Zhang, J. Lu, and J. Ma. 2011. A Kernel Fuzzy c-Means Clustering-Based Fuzzy Support Vector Machine Algorithm for Classification Problems With Outliers or Noises. *IEEE Transactions on Fuzzy Systems* 19, 1 (2011), 105–115. https://doi.org/10.1109/TFUZZ.2010.2087382

[20] C. Yue, R. Jin, K. Suh, Y. Qin, B. Wang, and W. Wei. 2018. LinkForecast: Cellular Link Bandwidth Prediction in LTE Networks. *IEEE Transactions on Mobile Computing* 17, 7 (2018), 1582–1594. https://doi.org/10.1109/TMC.2017.2756937

[21] R. Zhohov, A. Palaios, H. Ryden, R. Moosavi, and J. Belgrund. 2021. Reducing the Latency: Improving Handover Procedure Using Machine Learning. In *2021 IEEE 93rd Vehicular Technology Conference (2021-Spring)*. 1–6.