UE-Based Estimation of Available Uplink Data Rates in Cellular Networks

Georgi Nikolov
Nimbus Centre
Cork Institute of Technology
Email: georgi.nikolov@cit.ie

Michael Kuhn
University of Applied Sciences
Darmstadt
Email: michael.kuhn@h-da.de

Bernd-Ludwig Wenning
Nimbus Centre
Cork Institute of Technology
Email: berndludwig.wenning@cit.ie

***Abstract*—In this paper we propose a passive cellular data rate estimation method, which works with parameters accessible from most Commercial Off-The-Shelf (COTS) modems today. We show that by adjusting the optimization objective we can tweak the estimator to achieve higher accuracy at lower data rates. This is particularly useful in cases where a high data rate cellular link is backed up by a low data rate failover link (e.g. a satellite connection). The way the problem is approached makes this solution easily implementable in a product, for example as an input to a Quality-of-Service (QoS) router. The estimation is performed using both, supervised machine learning algorithms and a simple linear regression. The latter can be considered as a lightweight implementation on systems with low processing power or tight energy restrictions.**

I. Introduction

Intelligent transportation systems (ITS) have become a major area of research in the last few years. Services that make use of a vehicle's connection to infrastructure and backend/cloud systems are part of this research, be it for remote monitoring and maintenance, navigation, or value-added passenger services in the transportation sector. Some of these services not only require a good downlink to the vehicle, but also a good uplink to provide, e.g., critical sensor data to the backend systems. If the vehicle has multiple communication interfaces, such as 2G/3G/4G/5G cellular interfaces, WiFi, or satellite, it needs to make an informed decision on which interface to use, and this decision should take into account the available resources such as the uplink capacity on each interface.

In this context, we look at the estimation of available uplink capacity in cellular communication. Given that actively measuring the uplink capacity continuously would require test data transmission, which might incur cost and put additional load to the network, we instead prefer to estimate the capacity based on the set of channel quality information that is available at the User Equipment (UE).

This paper is structured as follows: Section II describes our use case, and Section III gives an overview of related work. Section IV presents the problem description, and we describe our measurement methodology in Section V. An overview of the prediction algorithms we have considered is given in Section VI. Our results are presented and discussed in Section VII. We conclude the paper and give an outlook to future work in Section VIII.

II. Use Case

Our research presented in this paper is based on a remote monitoring and maintenance scenario where sensors on board a vehicle, e.g. a train or a car, are sending data to a cloud backend to allow for remote monitoring and diagnostics. A mobile network link, e.g. UMTS or LTE, is used as the main data link for uploading this data from the vehicle. This main link is backed up by a satellite link as failover connection in case the mobile network cannot provide the required data rate. The usage of the failover link, however, should be reduced to a minimum due to the higher costs per bit of a satellite connection and should only be used when the available data rate on the mobile link drops below a threshold. Also, under normal conditions, a mobile network offers data rates up to several orders of magnitude higher than a satellite link. For this reason, estimation errors in the high data rate regime are less critical compared to estimation errors at lower data rates, i.e. at data rates close to the threshold where switching is done between the mobile network and the satellite link. Consequently, our goal is to provide an uplink data rate estimation method that is reliable enough in the low data rate regime to make informed decisions on whether the connection should use the main link or the failover link.

III. Related Work

In recent years, a large amount of research has been undertaken in the field of cellular data rate estimation. As a result, numerous models have been developed that can estimate the rate of a mobile link. Most of them use a packet sequence and curve fitting algorithms to approximate the measured queuing delay and account for its dynamic behavior caused by the eNodeB scheduler. The data rate is then calculated from the turning point of the queuing delay function. For example PathQuick3 [1] uses a macroscopic approach along with a nonlinear least squares algorithm to extract the approximation of the queuing delay. The data rate is then estimated from the extracted information. A similar method is also used in [2]. In [3] a different design of the packet sequence together with parabolic curves are used to perform the characterization of the queuing delay. The authors of [4] present a machine learning approach to the curve fitting problem, achieving much better results compared to previous works. Although these methods could be used for downlink as well as uplink data rate estimation and theoretically also be used for standards other than LTE, they all rely on sendprobing mechanisms. The transmitter has to send a sequence of packets to the receiver, resulting in unwanted congestion in the network. This becomes problematic, especially in a mobility scenario where the coherence time of the channel is very small and the channel often exhibits bad conditions that can result in packet loss and retransmissions.

Passive data rate estimation has been considered in [5] and [6]. Although technically passive, the method presented in [5] considers a monitor between the client and the server. The monitor gathers data from the exchanged packets and calculates an estimation of the data rate from it. This makes the approach dependent on an active packet transmission which reduces its reliability in cases where no or only a small amount of data is being exchanged between the server and the UE for an extensive period of time. The authors of [6], however, use a Neural Network (NN) that is trained with lower level channel information, in particular the Reference Signal Received Power (RSRP) and Reference Signal Received Quality (RSRQ) indicators. The method does not require any active data transmission, nor does it need a traffic monitoring tool. Instead, it monitors the cellular link parameters of the target UE as well as the UEs nearby using a Software Defined Radio (SDR) sniffer. This results in a reasonable quality of the prediction, but the need to implement a sniffer (either on the UE directly or as a separate component) makes this approach almost unusable in practical scenarios. Another potential issue of this method is the scalability. Its functionality is yet to be proven in cases where more than one cell is present in the area where the sniffer is acquiring data about the network.

IV. Problem statement

In order to develop a rationale for the estimation approach we have taken, we take a look at the theoretical background of uplink data rate estimation in the considered mobile communication systems, and we start by describing the physical layer. All mobile communication systems (2G/3G/4G/5G) have in common that the data rate in both the uplink and the downlink is chosen by the network and will be assigned to the User Equipment (UE). The network and the mobile device are performing periodic measurements of the mobile communication channel. The measurements performed by the eNodeB are relevant for the characterization of the uplink channel whereas the measurements conducted by the UE describe the downlink channel and are reported back to the network.

For LTE, the values measured by the UE are power and quality of cell-specific reference signals [7]. These values, i.e. RSRQ and RSRP, are transmitted to the eNodeB. Additional values, such as Received Signal Strength Indicator (RSSI) or Signal-to-Noise Ratio (SNR) may be available and accessible at the UE but are not reported. To facilitate the uplink measurements, the UE transmits Sounding Reference Symbols (SRS) which allow the network to perform the channel measurements. Based on these measurements, a Modulation and Coding Scheme (MCS) will be assigned to the mobile device. The MCS will be transmitted by means of Downlink Control Information (DCI) which is transmitted using the Downlink Physical Control Channel (DPCCH). The DCI carries information about downlink, uplink, and sidelink [9]. For the uplink, a five bit field is reserved in order to define the MCS that has to be used by the UE to transmit a message using the PUSCH (Physical Uplink Shared Channel). The mapping from the 5 bit value to the MCS is defined by tables in [10]. Three different modulation schemes are foreseen for the uplink, ranging from QPSK to 64-QAM. Depending on the assigned number of resource blocks N_{RB} and the number of symbols per timeslot N_{SPS}, the number of resource elements N_{RE} can be calculated as

$$N_{RE} = N_{RB} \cdot N_{SC} \cdot N_{SPS} \tag{1}$$

where N_{SC} is the number of subcarriers per resource block which is defined as $N_{SC} = 12$. The number of resource blocks ranges from 6 (for 1.4 MHz bandwidth) to 100 (for 20 MHz bandwidth) and the number of symbols per timeslot can be 6 (for extended cyclic prefix) or 7 (for normal cyclic prefix). The achieved data rate can thus be calculated as

$$R = \frac{N_{RE}}{T_{Slot}} \cdot N_{bps} \tag{2}$$

where N_{bps} is the number of bits per symbol which ranges from 2 (QPSK modulation) to 6 (64-QAM). However, it has to be considered that this data rate cannot be achieved in practice, since not all resource blocks (RBs) are available. Depending on the bandwidth and the network configuration, some RBs may contain the PUCCH (Physical Uplink Control Channel), the PRACH (Physical Random Access Channel), and/or Sounding Reference Symbols (SRS). Depending on the bandwidth configuration, this can significantly decrease the data rate. Even if we assume the parameters mentioned above to be known (or estimated by using a network sniffer), it is still not possible to reliably estimate the user data rate due to the following reasons:

1) The channel quality measurements of the uplink are performed by the eNodeB and not transmitted

to the UE. Therefore, the UE has no information about the channel quality in the uplink. Taking the channel quality of the downlink could be an approximation, however, since frequency division duplex schemes are typically applied, the channel reciprocity cannot be assumed and exploited.

2) The mapping from channel measurements to MCS is not defined in the standard and is instead defined by the manufacturer of the eNodeB and/or the network operator. Hence, these values are not accessible.
3) The channel is a shared resource and therefore the scheduler has to assign the resources depending on the load of the cell, the number of users, the requested data rate, and the data plan of the users.

Although some research has been done to map the uplink capacity of LTE to SNR values [11] and the available UMTS uplink data rate to an energy per chip to power spectral density ratio (E_c/I_0) [12], the mappings presented in these works fail to provide realistic estimates for the data rate and are unusable for our purposes. Due to the listed reasons, estimating the achievable uplink data rate based on calculations seems not to be feasible without detailed information from the network operator or exploitation of DCI information, as shown in [6]. However, with COTS chipsets and devices this information is generally not available. The approach chosen in this paper is therefore to use machine-learning algorithms which are trained from measurements performed in a real network only considering measurement data which is generally available on UEs.

V. Measurement Methodology

For the acquisition of data, three separate measurement runs have been conducted. The first and second measurement runs took place in a car, driving around Cork city center on two different days. The third measurement run was done in a train, traveling from Cork to Mallow through a mostly rural environment. The UE we used was a device with two Sierra Wireless MC7455 modules, one with a UMTS Subscriber Identity Module (SIM) and the other with an LTE SIM. Using traffic generated by Iperf [8], both mobile interfaces of our device were fully utilized and data rate as well as channel parameters were recorded. For the LTE link, we recorded SNR, RSRP, RSRQ and the allocated bandwidth. For UMTS, RSSI, E_c/I_0 and RSCP were acquired. The volume of the acquired measurements is as follows: First car measurement: 400 samples, second car measurement: 350 samples, and for the train measurement: 200 samples. Due to hardware constraints the sample time was chosen to be 5 seconds. The UE was on board a moving vehicle, passing through different cells. For each measurement a different route was chosen. In this way, we ensured that the different data sets were taken under different conditions, in order to test how robust the estimation works. The three data sets are used as follows: The first measurement, conducted with a car moving through an urban environment, is used as a training set, the second measurement (also performed in a car moving through an urban environment) is used as development set, and the third measurement (conducted in a train travelling in a rural environment) is used as a test set.

VI. Prediction Algorithms

We consider a Neural Network (NN) for the prediction of the achievable uplink data rate using two different optimization objectives. The first approach is to minimize the absolute error. For the second approach we aim to minimize the relative error. Both neural networks use one hidden layer with 100 neurons and a rectified linear activation function. The activation function was selected since it provided best performance among the following list of activation functions which have been compared using the validation data set: competitive transfer function, hard limit transfer function, symmetric hard limit transfer function, log sigmoid transfer function, rectified linear transfer function, linear transfer function, radial basis transfer function, saturating linear transfer function, symmetric saturating linear transfer function, softmax transfer function, hyperbolic tangent sigmoid transfer function, and triangular basis transfer function. Also the number of neurons was selected best on simulation results varying the hidden layer size. 100 neurons appeared to be a reasonable compromise between complexity/calculation time and performance. Further increase did not significantly increase the performance. Adding additional hidden layers to the network increased the computation time without any noticeable improvement in the network performance.

Training of the network is performed using Bayesian regularization which means that the weight and bias values are updated according to Levenberg-Marquardt optimization. Again, the training algorithm has been selected based on evaluations on the validation data set by comparing about ten different training algorithms in order to select best one. By reducing the relative error we expect to minimize the errors at lower data rates, close to the switching threshold between mobile network and failover link. The authors of [13] present a broad research, which shows that the RSRP and the RSRQ are very good indicators for the quality of the uplink. Using feature selection and optimization we found out that for our purpose using SNR, RSRP, RSRQ and the allocated bandwidth as input features for the neural network provides the best performance. For UMTS, best results have been achieved using RSCP and E_c/I_0.

We further used a linear regression for the prediction of data rates using only E_c/I_0 as a parameter in case of UMTS, and only RSRP in case of LTE, respectively. The

TABLE I
COMPARISON OF DIFFERENT ESTIMATION ALGORITHMS(LTE).

	Performance Indicator		Estimation Accuracy		
	MARE	MSE	underestimated	good estimation	overestimated
Lin. Reg.	0.96	107.81	7.58%	78.28%	14.14%
NN (AE opt.)	1.12	92.04	0.51%	87.37%	12.12%
NN (RE opt.)	0.89	139.76	8.08%	83.84%	8.08%

objective is to investigate the performance of a scheme that could easily be implemented on a low-power/low-cost UE without the restraint of a complex and power-consuming machine-learning algorithm.

For all algorithms we used the first measurement set as training data. The development set was used to test the performance with respect to different parameter settings, e.g. activation functions, number of neurons, number of hidden layers, network training function, etc. The settings that have provided best results on the development set were finally taken to evaluate the performance on the test set.

VII. RESULTS AND DISCUSSION

Table I presents a summary of the results obtained with our algorithms. It is shown that the Absolute-Error (AE) optimized NN has the best accuracy with respect to the Mean Square Error (MSE) with a value of 92.04. However, its Mean Absolute Relative Error (MARE) is 1.12. The algorithm achieves good estimation (which we define as an estimate within the range of $\pm 50\%$ of the real value) in 87.37% of the cases, underestimating 0.51% of the values, and overestimating 12.12%. The linear regression performs second best when it comes to MSE with a value of 107.81. Regarding the MARE, this algorithm is in between the two NNs with a value of 0.96. It achieves good estimation in 78.28% of all cases, overestimates 14.14% and underestimates 7.58%. The Relative-Error (RE) optimized NN on the other hand has the best MARE of 0.89, but the worst MSE of 139.76. It manages to produce a good estimation in 83.84% of the cases, both underestimating and overestimating 8.08% of the data rates. The results presented in Table I are not surprising. Setting the Mean-Squared-Error as an optimization objective will force the optimizer to adjust the weights of the Neural Network in such way that the absolute error of the estimation stays as small as possible. In a perfect world the algorithm will achieve an error equal to zero. This of course will result in the relative error also being equal to zero. In the real world however, there is always some error between estimation and actual value. Training the estimator to minimize the MSE does not take into consideration whether the absolute values are small or large. Because of that the overall relative error resulting from the imperfection of the optimizer becomes larger, dominated by the relative error at smaller values. In contrast, when the optimizer adjusts the network parameters to minimize the relative error, it will allow for larger errors for large values and small errors for small values respectively, thus resulting in larger absolute error.

Fig. 1 shows the output of the tested algorithms as a function of the real data rate. The closer the red points (prediction) are to the black line (real rate), the more accurate is the prediction. We can see that the linear regression and the NN optimized for minimum Absolute Error (AE) provide more accurate estimates at higher data rates compared to the algorithm which aims at minimizing the Relative Error (RE). On the other hand the RE-optimized network performs better at lower data rates. Fig. 2 illustrates the data rate prediction of the algorithms compared to the real data rate. In this figure, one can also observe that the AE-optimized NN is able to approximate the real data rate best, resulting in the smallest absolute error. The linear regression also performs decently when it comes to absolute error, very close to the AE-optimized NN. In contrast the RE-optimized NN estimates the higher rates with lower accuracy, in exchange for an improvement at the lower rates. As one can see, the linear regression is able to estimate the data rate much more accurately than initially expected, indicating a linear relationship between the RSRP and the data rate.

Fig. 3 illustrates the cumulative distribution functions of the relative error of the three algorithms of interest. While this figure confirms the numbers of Table I we can observe that within the area of good estimations all algorithms tend to slightly underestimate the data rate. While linear regression and AE-optimized NN show approximately the same behaviour, the RE-optimized NN exhibits a slightly stronger tendency to underestimate.

Fig. 4 shows the same analysis as Fig. 1, but for the UMTS-link. Similar to the results of the LTE-link, we can observe that the RE-optimized NN will have its prediction located closer to the real data rate at lower values, whereas the linear regression and the AE-optimized neural network tend to have higher accuracy at higher rates.

Fig. 5 shows a similar CDF to Fig. 3. In contrast to LTE, the tendency to underestimate in the good estimation range is less prominent. Within this range, the AE-optimized NN and the linear regression slightly overestimate the values whereas the RE-optimized NN slightly underestimates. Further, compared to LTE, the

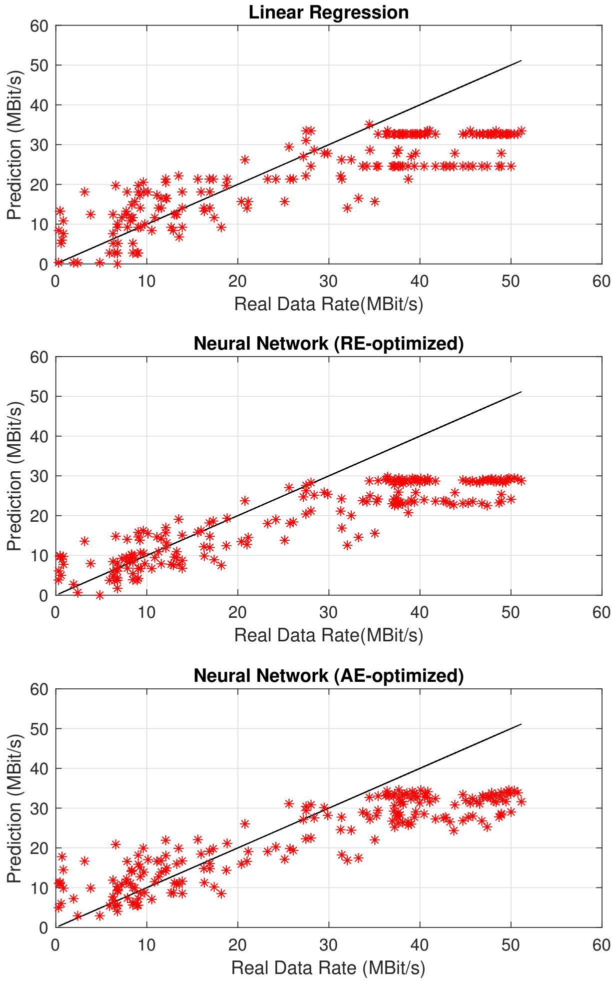

Fig. 1. Output of all estimators as a function of the real data rate (LTE)

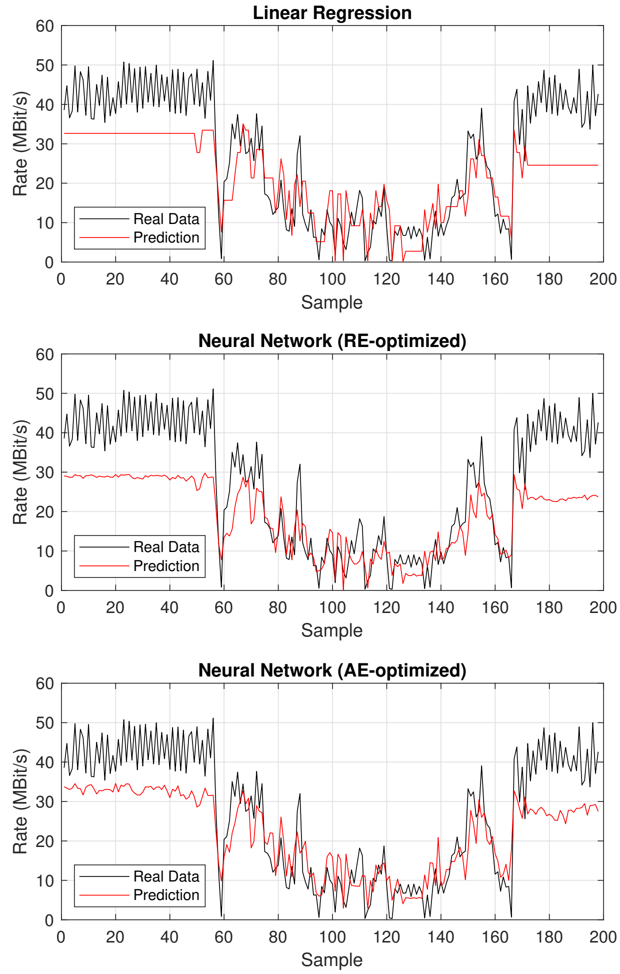

Fig. 2. Prediction compared to real data rate (LTE)

AE-optimized NN and the linear regression have a stronger tendency to overestimate by more than 50%.

VIII. Conclusion and Future Work

In this paper we have proposed both, a neural network and a linear regression scheme for the passive estimation of the uplink data rate in UMTS and LTE networks. The neural network learns from parameters which are easily accessible from almost every cellular modem, making this solution widely implementable. It has been shown that by changing the cost function of the neural network optimizer, a significant reduction of the relative error can be achieved. This is particularly useful in cases where low data rates have to be predicted with higher accuracy compared to higher data rates. This can prove beneficial in cases of one high data rate link being backed up by a low data rate link. In this scenario, it becomes essential that the prediction is more accurate at lower rates than it is at higher rates. We also have shown that a linear regression can be a good solution in case complexity

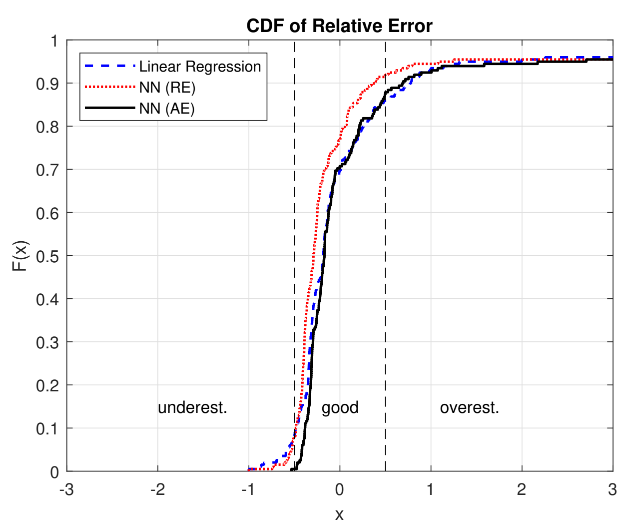

Fig. 3. Cumulative distribution function of the relative error of all algorithms compared (LTE)

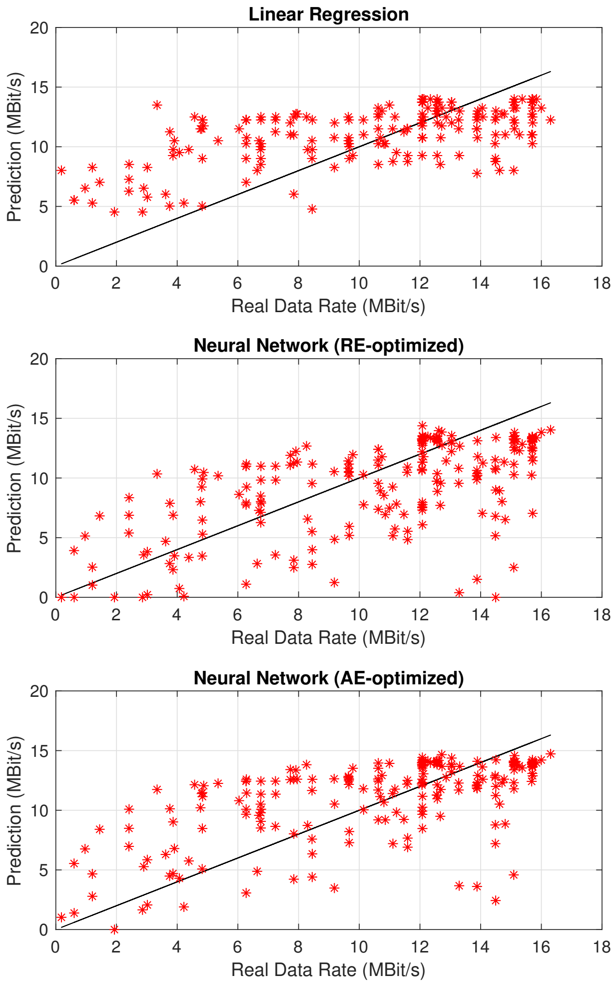

Fig. 4. Output of all estimators as a function of the real data rate (UMTS)

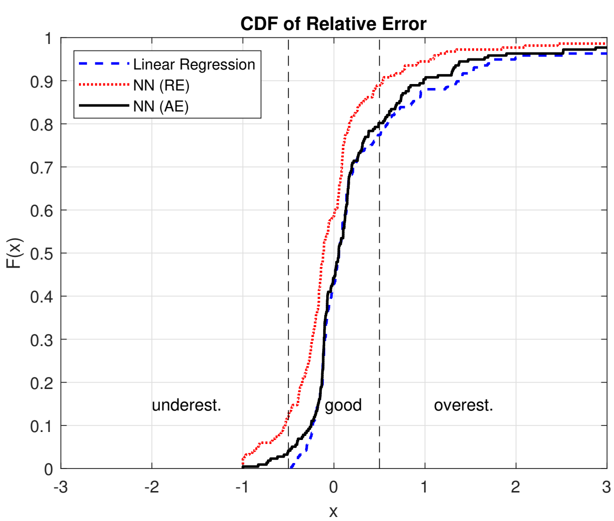

Fig. 5. Cumulative distribution function of the relative error of all algorithms compared (UMTS)

and power consumption of the implementation is key, e.g. for battery powered devices.

Our next step will be to gather more training and test data for the neural network evaluation, particularly in different cell load conditions. Additionally, we will model the switching between the low data rate link and the high data rate link as a classification problem and investigate the performance of machine learning classifiers in this scenario.

ACKNOWLEDGMENT

SCOTT (www.scott-project.eu) has received funding from the Electronic Component Systems for European Leadership Joint Undertaking under grant agreement No 737422. This joint undertaking receives support from the European Union's Horizon 2020 research and innovation program and Austria, Spain, Finland, Ireland, Sweden, Germany, Poland, Portugal, Netherlands, Belgium, Norway.

REFERENCES

[1] T. Oshiba, K. Nogami, K. Nihei and K. Satoda ”Robust Available Bandwidth Estimation Against Dynamic Behavior of Packet Scheduler in Operational LTE Networks” *IEEE Symposium on Computers and Communication (ISCC)*, 2016

[2] A. K. Paul, A. Tachibana and T. Hasegawa ”NEXT-FIT: Available Bandwidth Measurement over 4G/LTE Networks – A Curve-fitting Approach” *IEEE 30th International Conference on Advanced Information Networking and Applications (AINA)*, 2016.

[3] S.Shiobara and T. Okamawari ”A novel Available Bandwidth Estimation Method for Mobile Networks using a Train of Packet Groups” *IMCOM Proceedings of the 11th International Conference on Ubiquitous Information Management and Communication, Article No. 59*, 2017

[4] N. Sato, T.Oshiba, K. Nogami, A. Sawabe and K. Satoda ”Experimental Comparison of Machine Learning-Based Available Bandwidth Estimation Methods over Operational LTE Networks” *IEEE Symposium on Computers and Communications (ISCC)*, 2017

[5] J. Huang, F. Qian, Y. Guo, Y. Zhou, Q. Xu, Z. M. Mao, S. Sen and O. Spatscheck ”An In-depth Study of LTE: Effect of Network Protocol and Application Behavior on Performance” *Proceedings of the ACM SIGCOMM 2013 conference on SIGCOMM, pp. 363-374* , 2013.

[6] R. Falkenberg, K. Heimann and C. Wietfeld ”Discover Your Competition in LTE: Client-Based Passive Data Rate Prediction by Machine Learning” *IEEE Global Communications Conference(GLOBECOM)*, 2017

[7] 3GPP TS 36.214 Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer Measurements (version 10.1.0 Release 10)

[8] https://iperf.fr/ (last accessed 5.09.18)

[9] 3GPP TS 36.212 Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding (version 13.1.0 Release 13)

[10] 3GPP TS 36.213 Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures (version 14.2.0 Release 14)

[11] J. Blumenstein, J. C. Ikuno, J. Prokopec and M. Rupp ”Simulating the Long Term Evolution Uplink Physical Layer” *International Symposium ELMAR*, 2011.

[12] Goncalo Nuno Coelho e Ferraz de Carvalho ”Data Rate Performance Gains in UMTS Evolution at the Cellular Level” *Dissertation submitted for obtaining the degree of Master in Electrical and Computer Engineering, Universidade Tecnica de Lisboa, p. 27*, 2009.

[13] C. Ide, R. Falkenberg, D. Kaulbars and C. Wietfeld ”Empirical Analysis of the Impact of LTE Downlink Channel Indicators on the Uplink Connectivity” *IEEE Vehicular Technology Conference (VTC Spring)*, 2016