Estimating Video Streaming QoE in the 5G Architecture Using Machine Learning

Susanna Schwarzmann
TU Berlin
susanna@inet.tu-berlin.de

Clarissa Cassales Marquezan
Huawei Technologies Germany
clarissa.marquezan@huawei.com

Marcin Bosk
TU Berlin
marcin@inet.tu-berlin.de

Huiran Liu
TU Berlin
huiran@inet.tu-berlin.de

Riccardo Trivisonno
Huawei Technologies Germany
riccardo.trivisonno@huawei.com

Thomas Zinner
TU Berlin
zinner@inet.tu-berlin.de

Abstract

Compared to earlier mobile network generations, the 5G system architecture has been significantly enhanced by the introduction of network analytics functionalities and extended capabilities of interacting with third party Application Functions (AFs). Combining these capabilities, new features for Quality of Experience (QoE) estimation can be designed and introduced in next generation networks. It is, however, unclear how 5G networks can collect monitoring data and application metrics, how they correlate to each other, and which techniques can be used in 5G systems for QoE estimation. This paper studies the feasibility of Machine Learning (ML) techniques for QoE estimation and evaluates their performance for a mobile video streaming use-case. A simulator has been implemented with *OMNeT++* for generating traces to (i) examine the relevance of features generated from 5G monitoring data and (ii) to study the QoE estimation accuracy (iii) for a variable number of used features.

Keywords

QoE; 5G; Mobile Networks; Machine Learning; HAS; Adaptive Streaming

ACM Reference Format:
Susanna Schwarzmann, Clarissa Cassales Marquezan, Marcin Bosk, Huiran Liu, Riccardo Trivisonno, and Thomas Zinner. 2019. Estimating Video Streaming QoE in the 5G Architecture Using Machine Learning. In *4th Internet-QoE Workshop: QoE-based Analysis and Management of Data Communication Networks (Internet-QoE'19), October 21, 2019, Los Cabos, Mexico.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3349611.3355547

Internet-QoE'19, October 21, 2019, Los Cabos, Mexico

ACM ISBN 978-1-4503-6927-5/19/10…$15.00
https://doi.org/10.1145/3349611.3355547

1 Introduction

Online video streaming has become the prevalent way of video consumption and contributes a large fraction of today's global IP traffic. Cisco forecasts that by 2022, mobile video will make up nearly 80% of the overall mobile data traffic [8]. Driven by business incentives, providing a good Quality of Experience (QoE) is of up-most importance for Mobile Network Operators (MNOs), Internet Service Providers (ISPs), and content providers or Virtual Content Providers (VCP). While the QoE can only be estimated reliably by the content provider based on application data, it is the MNO/ISP that has the capability to perform application-aware resource control. When it comes to mobile networks, the 5G system architecture has been significantly enhanced by the introduction of network analytics functionalities and extended capabilities of interacting with third party Application Functions (AFs). 5G Network Functions (NFs) and AFs have standardized interfaces to allow the communication of application-specific information, such as QoE information sent from VCP to MNO. It is desirable for the MNO to estimate the QoE in the system also in the absence of application-specific data or if such data is aggregated by the AF. To achieve this, MNOs of 5G networks can apply Machine Learning (ML) techniques to derive the QoE from network-level monitored statistics, as the newly introduced Network Data Analytics Function (NWDAF) specifies. The current NWDAF specification includes the capability of generating analytics information about service experience based on AF data, i.e., the QoE of applications. In addition, the monitoring data to be used for generating such analytics has been fixed.

To the best of our knowledge, this paper is the first attempt to investigate issues on how NWDAF correlates network statistics with application metrics and the corresponding QoE in 5G systems. In our approach, a third party tenant initially communicates video streaming performance data to NWDAF via the AF. This information is, from the NWDAF perspective, the ground truth QoE. Statistical processing of network-level monitoring data generates a vast number of *features*, whose relation to the ground truth QoE, i.e. relevance, is studied. Then, ML-based models are trained to estimate the QoE from those network-level features that are available to the NWDAF. We study the feasibility of such an approach with traces generated within an *OMNeT++* simulation, that allow for correlating network statistics and QoE. We show that a small set of network-related features and low-complex regression methods already allow an accurate video QoE estimation with a mean squared error below 0.1 on Mean Opinion Score (MOS) scale.

The remainder of the work is structured as follows. Related work is presented in Section 2. Section 3 describes a possible integration of our solution in the 5G architecture for a video streaming use-case. The methodology is described in Section 4, followed by the evaluation in Section 5. Section 6 concludes the paper.

2 Background and Related Work

The fundament for estimating the QoE from network QoS metrics is to understand how they relate to each other. For that reason, [10] studies the causal relations between video QoE and network and application QoS. Making use of those relations to classify video QoE from encrypted network traffic is recently widely discussed and numerous works have been published in the past few years addressing this topic. The works presented in [7, 11, 13, 16, 22] study the capabilities of different ML-based algorithms to classify values for QoE influence factors (QoE-IFs), such as stallings or video quality. Compared to previous works, we do not classify objective QoE-IFs, but estimate actual QoE values. We use the QoE on MOS scale as computed by the standardized ITU-T P.1203 model [20], which uses a set of objective quality assessment modules to measure the subjective application quality perceived by a user. It is used in the context of classifying QoE from encrypted network traffic in [17] and [11].

QoE estimation with a focus on mobile networks is addressed in [12], where the performance of different classifiers and the influence of the used network- and application-specific features is studied. A solution for estimating the QoE solely based on network-related performance indicators for LTE networks is presented in [6]. The authors of [3] study different methods for predicting the QoE of FTP file transfers and propose to apply their methods for detecting anomalies in self-organized networks. Our proposal differs from existing solutions as we estimate the QoE solely based on network-level metrics, including all metrics available to an MNO that have not been considered so far, e.g. channel quality indicator (CQI). We furthermore present an embedding of ML-based QoE estimation into the 5G network architecture and strengthen the applicability of our approach

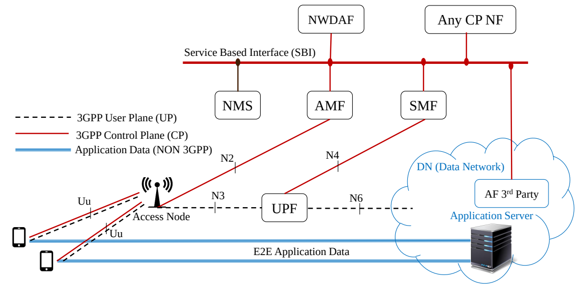


Figure 1: 5G data analytics framework

by taking advantage of 5G-specific NFs, which enable obtaining ground truth data from third parties and retrieving network-level statistics.

Applying machine learning in 5G gains an ever-increasing interest, e.g. to predict the number of active users [19], for traffic forecasting [2], or cognitive networking in [5]. A first QoE-centric approach provides a service quality prediction model for UHD real time video streaming [15]. A network resource allocation system for autonomous QoE-aware 5G network management is discussed in [14].

3 Flexible Approach for Analytics Generation in 5G Systems

This section introduces the data analytics support in 5G first in a general manner and afterwards with a focus on QoE. Then, we describe our approach for integrating ML-based QoE estimation into 5G systems.

3.1 Current Data Analytics Support for 5G

The latest 3GPP release of the 5G Network Architecture, i.e., release 16, includes a new, dedicated specification for network data analytics support in 5G systems.[1] It defines the framework for integrating analytics functionality in 5G and the types of analytics that can be generated by the NWDAF. The 5G analytics framework is based on the following principles: a) The NWDAF as shown in Figure 1 is connected to the Service Based Interface (SBI), allowing the NWDAF to collect data from AFs and from any other 5G control plane NFs, e.g., Access and Mobility Management Function (AMF) or Session Management Function (SMF). The NWDAF is also able to collect data from the 5G management plane (NMS - Network Management System), where user plane data sourcing from Access Node (AN) and User Plane Function (UPF) can be obtained; b) The NWDAF is capable to generate analytics based on ML techniques that can be consumed by 5G NFs, AFs, or 5G management plane entities; c) The internals of NWDAF are vendor specific. Additionally, a set of analytics that can be generated by NWDAF have been standardized. This gives guarantees to MNOs that at least a minimum set

[1] 3GPP TS 23.288 V16.0.0, Architecture enhancements for 5G System (5GS) to support network data analytics services, 2019-06

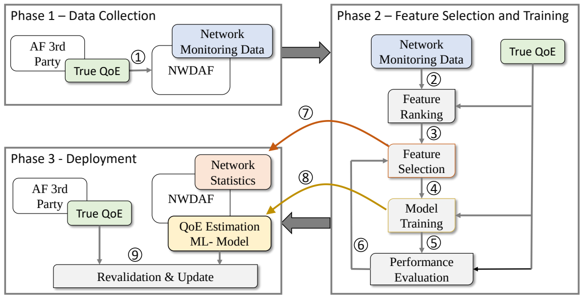


Figure 2: Workflow of proposed ML-based QoE estimation integration into the 5G architecture

of analytics will be offered by any vendor of NWDAF, but does not restrict NWDAF to generate other types of analytics information.

3.2 QoE Analytics in 5G Systems

Given the importance of QoE, 3GPP defined the "Service Experience" analytics information generated by NWDAF[1]. These analytics indicate how QoS parameters satisfy the MOS agreed between the mobile operator and an external party such as an Application Service Provider (ASP) or VCP. Such external party is associated with the AF from where the NWDAF can collect data or provide analytics to. In terms of specific measurements, the set of data to be collected by NWDAF includes: the MOS from the AF (which is used to train the QoE model); and network monitoring data related UL and DL traffic, number of transmitted and retransmitted packets, all per flow of data in 5G, as well as radio quality conditions per User Equipment (UE). The input data defined for generating such analytics information ultimately defines the set of statistics, i.e. features, that can be offered to an ML model for estimating the QoE. However, as there is currently no deployment of 5G networks, there is no evidence of how the approach followed in the standard of fixing the set of features actually performs.

3.3 Proposed ML-based QoE Estimation Integration for Mobile Video Streaming

In this paper, we do not assume an upfront fixed set of parameters provided by the NWDAF, but propose the approach illustrated in Figure 2, which reveals necessary parameters for reliable QoE estimation. Our proposed integration of ML-based QoE estimation considers three phases. During the first phase the third party AF communicates application performance data on a per-session basis to the NWDAF. Hence, the NWDAF's database is enriched by ground truth application QoE ①. In the second phase, a vast number of network features is generated by statistical processing of the network monitoring data. The resulting features are ranked according to their significance in terms of influencing the QoE ②. Then, a subset of significant features ③ and the ground truth QoE values are used to train ML-based models for QoE estimation ④. In the next step, the estimation performance is evaluated ⑤. This process can be repeated for different feature sets and ML-based models ⑥, until a desired estimation accuracy is obtained. The identified feature set dictates the necessary network statistics the NWDAF has to provide ⑦ in order to reliably estimate the QoE during deployment phase based on the trained model ⑧. Although the MNO can now estimate the QoE without the need of application-metrics provided by the AF, the VCP can still communicate such information to facilitate updating, verifying, and improving the trained model ⑨.

4 Methodology

The following section details on the simulation environment and describes the applied QoE estimation techniques.

4.1 Omnet++ Simulation Environment

For simulating the mobile video streaming clients and the underlying network, we use *OMNeT++* [23] together with the frameworks *INET* and *SimuLTE*.[2] Although *SimuLTE* simulates LTE networks, the same type of monitored information in 4G entities, such as PGW (PDN, Packet Data Network, Gateway)[3] and eNB (i.e., access node) [4], is available to be collected from, respectively, UPF and gNB (i.e., access node) in 5G[5,6]. At this stage of our research, the main point of using *SimuLTE* is to obtain monitored information from both, access and core network. We assume the monitoring information to be available at NWDAF and we are not considering any signaling exchange for data collection yet. Therefore, *SimuLTE* can be used for generating user plane traffic in a mobile network. From the perspective of the radio technology, 5G has a much higher performance than 4G. Hence, absolute throughput values in an experiment will be much higher with a 5G gNB. However, the principles of system load (number of UEs in a cell) and radio quality will still play a role in 5G systems.

4.1.1 Network Topology. The network topology used in *SimuLTE* reflects the same user plane entities as illustrated in Figure 1. It consists of a single access node (AN) connected to a gateway (i.e., PGW). This maps to the AN entity connected via N3 interface to the UPF in Figure 1. Such gateway is directly connected to the video server. This is equivalent to UPF connected via N6 interface to the Application Server in Figure 1. The UEs are randomly placed within a square of 500x500 meters around the AN. For these experiments, we consider a single cell without any external interference.

[2] We use Omnet++ 5.1, INET 3.5, and SimuLTE 0.9.1. The simulation environment is made available on Github. https://github.com/fg-inet/vagrant-omnet-simulation
[3] 3GPP TS 32.426 V15.2.0 (2018-12). Performance measurements EPC network
[4] 3GPP TS 32.425 V16.3.0 (2019-06), Performance measurements E-UTRAN
[5] 3GPP TS 28.552 V16.2.0 (2019-06), 5G Performance Measurements
[6] 3GPP TS 28.554 V16.1.0 (2019-06), 5G end to end Key Performance Indicators

Table 1: Configurations in the *OMNeT++* simulator.

CHANNEL	Carrier frequency	*carrierFrequency*	2.1 GHz
	Maximum sending power	*pMax*	20 W
	Uplink (UL) bandwidth	BW_{UL}	20 MHz
	Downlink (DL) bandwidth	BW_{DL}	20 MHz
RLC	Size of fragments	*fragmentSize*	30 B
	Timeout for RX buffer	*timeout*	1 s
MAC	MAC buffers queue size	*queueSize*	5 MiB
	Schedulable Bytes	*maxBytesPerTti*	3 MB
	DL scheduling discpline	*schedulingDDl*	MAXCI
	UL scheduling discpline	*schedulingDUl*	MAXCI
AMC	# resource blocks in DL	*numRbDl*	100
	# resource blocks (*Rb*) in UL	*numRbUl*	100
	# subcarriers per *Rb* in DL	*rbyDl*	12
	# resource blocks in DL	*rbyUl*	12
	# logical bands	*numBands*	100

Table 2: Monitored network data and applied statistical metrics for feature generation.

	Notation	Description
MONITORED DATA	*AN_Tp_Dl*	Access Node downlink (Dl) throughput
	AN_Tp_Ul	Access Node uplink (Ul) throughput
	UE_Tp_Dl	UE Dl throughput
	UE_Tp_Ul	UE Ul throughput
	CQI_Dl	Dl channel quality indicator (CQI)
	CQI_Ul	Ul CQI measured at UE
	M_RTT	Measured RTT at the UE
	S_RTT	Smoothed RTT using a moving average
STATISTICS	mean, min, max, 25th percentile, 75th percentile, median, standard deviation (std), variance, coefficient of variation (cvar), kurtosis, skewness, unbiased standard error of the mean (sem)	

4.1.2 Simulation Parameters. We summarize the most important *OMNeT++* simulation parameters in Table 1. Thereby, we differentiate channel settings, radio link control (RLC) settings, configurations on MAC layer, and adaptive modulation and coding (AMC) configurations. All remaining parameters are kept to the *OMNeT++* default setting.[2]

4.1.3 Simulation Scenarios. We vary the system load by considering different numbers of active video clients, i.e., we vary from 20 to 200 UEs in steps of 20 UEs per configuration ($N_{UE} = 20 : 20 : 200$). Each configuration is run using four different seeds, which determine the UE placement around the access node and the video clients' start times.

4.1.4 Client Parameters and Video Properties. The video server provides a video of 300 seconds length, split into segments of 5 seconds duration. We consider three different quality levels having bitrates of 500 *kbps*, 1500 *kbps*, and 3000 *kbps*. The client is equipped with a simple buffer-based heuristic, has a maximum buffer length of 30 seconds, and quality switching thresholds of 10 and 20 seconds for switching to the second and third quality level.

4.2 Estimating QoE from Network QoS

In the following, we first detail on the generated features and how their relevance can be studied. Afterwards we give a short summary on the applied ML-techniques.

4.2.1 Feature Generation. During the simulations, we monitor different network metrics, which can be available to *NWDAF* in 5G, as previously discussed. In order to train ML models to estimate the user QoE from the available network data, we generate numerous features from these data. Table 2 summarizes the monitored network data and lists the applied statistics for feature generation. In total, we consider 8 network monitoring metrics and 12 statistical sizes, summing up to 96 features. In the case of *UE_Tp_Dl* and *UE_Tp_Ul*, we use a subset of the monitored network traffic. For instance, packet headers and acknowledgments are omitted, so to obtain the application-layer goodput.

4.2.2 Feature Relevance. In order to study the relevance of certain features, we apply three different state-of-the-art methods: *Univariate F-test Statistics (F-test)*, *Mutual Information Regression (MIR)* [18], and *Principal Component Analysis (PCA)* [9]. When MNOs apply the F-Test or MIR, the features are ranked according to their relevance without modifying them. Therefore, MNOs can explicitly identify the set of monitoring information most relevant. In contrast, PCA applies an orthogonal transformation on the input features to convert them into a new set of features, referred to as *principal components*. Hence, in the case of *PCA*, revealing the most relevant monitoring points is more complex.

4.2.3 Applied Regression Techniques. The ultimate goal is to estimate the QoE solely from data available at NWDAF, i.e. network statistics only. We study the feasibility of such an approach for two typical regression methods, *Linear Regression (LR)* [21] and *Support Vector Regression (SVR)* [4]. Both, LR and SVR, are supervised learning models, i.e., they need to be trained on a ground truth data set. They furthermore output actual values, i.e., they estimate a value rather than their affinity to predefined classes.

We study the estimation accuracy of these models for an increasing number of used features according to the obtained feature ranking. The training set consists of 70% of the ground truth data points. The remaining 30% are used for testing the estimation performance. For quantifying the estimation accuracy, we apply the following typical measures: Median absolute error (MedAE), mean squared error (MSE), and the coefficient of determination, referred to as R^2.

5 Evaluation

In order to retrieve the ground truth QoE values for training and evaluating the ML models, we apply the ITU-T P.1203 model on the relevant video metrics of each session. Figure 3 illustrates the CDF of the obtained QoE values on MOS scale.

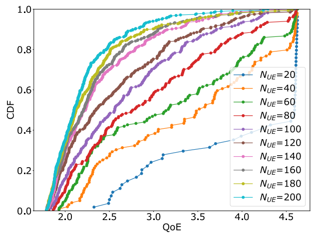


Figure 3: True QoE scores obtained for varying numbers of active UEs

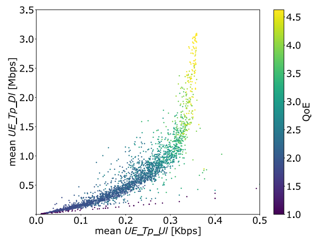


Figure 4: Relationship between filtered UL and DL throughput and QoE

Table 3: Feature ranking obtained by MIR and F-test

Rank	MIR	F-test
1	mean *UE_Tp_Ul*	mean *UE_Tp_Dl*
2	mean *UE_Tp_Dl*	skewness *UE_Tp_Dl*
3	cvar *UE_Tp_Dl*	variance *UE_Tp_Ul*
4	max *CQI_Ul*	75th percentile *CQI_Ul*
5	kurtosis *M_RTT*	std *AN_Tp_Dl*
6	min *M_RTT*	max *AN_Tp_Dl*
7	sem *M_RTT*	skewness *M_RTT*
8	25th percentile *S_RTT*	skewness *AN_Tp_Ul*
9	kurtosis *S_RTT*	75th percentile *CQI_Dl*
10	max *UE_Tp_Ul*	mean *M_RTT*

The QoE tends to decrease with an increasing number of active clients, which is attributed to the fact of increasing overall system load. The set of ground-truth data is used in the following to retrieve the feature ranking and to study the estimation accuracy of the applied regression techniques.

5.1 Feature Ranking

Table 3 shows the 10 most relevant features in descending order obtained by *MIR* and *F-test*. As both methods are very likely to rank highly correlated features similarly, we exclude those features that have a correlation of at least 0.7 to another, higher ranked feature. Please note that PCA modifies the features in such a way, that they cannot straightforwardly be interpreted. Hence, they are omitted in the table.

As expected *UE_Tp_Dl* is identified as an important indicator for the QoE. MIR recognizes the mean *UE_Tp_Ul* as the most relevant feature, motivating to take a closer look on the relationship of *UE_Tp_Dl*, *UE_Tp_Ul*, and QoE, as depicted in Figure 4. It shows an exponential relationship between the average application uplink and downlink throughput, i.e. the throughput omitting packet headers and acknowledgments.

Besides the filtered UE throughput, features related to the channel quality, i.e. CQI, obtained high rankings. Both, MIR and F-Test, rank CQI-related features on position four. With regard to the applied statistics, we see that skewness and kurtosis have notable relevance for estimating QoE. To the best of our knowledge, their relevance has not yet been examined in detail and only few works [1, 10] consider these statistics for generating features.

5.2 QoE Estimation Accuracy

We detail now on the estimation accuracy that can be achieved by linear regression and support vector regression with a varying number of features.

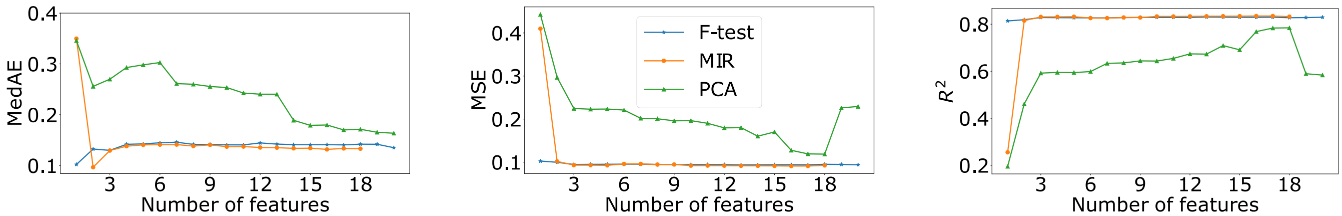
5.2.1 Linear Regression. Figure 5 summarizes the results obtained with linear regression. The number of used features is increased according to the ranking given in Table 3. If we use the features as ranked by the F-test, the lowest median absolute error can be achieved by only using the mean *UE_Tp_Dl*. Using a larger feature set does not involve a better performance in terms of MedAE. Estimating the QoE only with the highest ranked feature from MIR (mean *UE_Tp_Ul*) results in a MedAE of roughly 0.35. This value can be reduced to 0.1 when additionally considering the mean *UE_Tp_Dl*. For PCA, a large feature set is required to obtain comparable results. Similar conclusions can be drawn when examining MSE and R^2 score. While the first two features retrieved from F-test and MIR already provide an MSE below 0.15 and an R^2 score above 0.8, a set of 16 features is required in the case of PCA to obtain a similar performance.

5.2.2 Support Vector Regression. The performance metrics for the support vector regression are given in Figure 6. When estimating the QoE only based on a single feature, the lowest median absolute error can be achieved with the F-test as feature selection method. Here, the QoE can be estimated in most cases only differing less than 0.1 from the true QoE value. To achieve similar accuracy with feature sets revealed from MIR and PCA, two and four features are needed respectively. A similar behavior can be observed for the mean squared error, as shown in Figure 6b. In terms of R^2 (Figure 6c), acceptable scores can be achieved using at least two features obtained from MIR and F-test or at least four PCA features.

5.3 Discussion

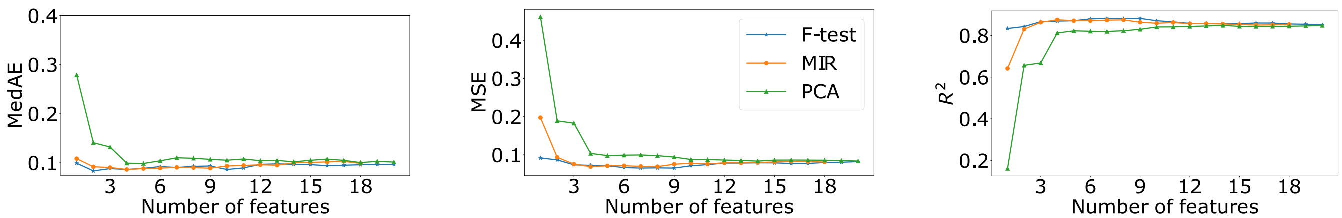
The presented results can provide new insights to MNOs and act as a first step towards understanding the correlation of 5G monitoring data to user QoE and towards designing ML-based QoE estimation approaches for their systems. Low-complex regression techniques and small feature sets already suffice for a reliable QoE estimation. Our results show that MNOs can rely on simple feature selection methods, such as MIR and F-test, which provide a direct mapping between high ranked features and data to be monitored.

Finally, the experiments discussed in this paper are a first step towards understanding the issues of network data correlation with application QoE in 5G systems, specifically on the monitoring data needed for accurate QoE estimation. Comparing our first results with the current specification of service experience analytics, our results confirm some of the monitoring data listed in the standard as relevant, but also



(a) Median absolute error **(b) Mean squared error** **(c) R^2 score**

Figure 5: Estimation performance of linear regression depending on the number of applied features



(a) Median absolute error **(b) Mean squared error** **(c) R^2 score**

Figure 6: Estimation performance of support vector regression depending on the number of applied features

provide initial evidences that not all monitoring data specified in the standard might be actually beneficial. In addition, we also show with our work, that the approach of fixing upfront the set of data to be collected for generating analytics by NWDAF might not be the best strategy, or should be carefully considered during standard work process.

6 Conclusion

The upcoming 5G networks will enhance access to network analytics data and standardized interfaces for information exchange between MNOs and third party applications. In this work, we investigated issues on how NWDAF correlates network statistics with application metrics and the corresponding QoE in 5G systems for an initial scenario. We showed that the subjective QoE score can reliably be estimated with standard regression approaches, solely based on network monitoring data, and that a small set of features allow for obtaining a sensible accuracy. However, given the simulation-based nature of the evaluation, the findings still need to be verified within a real 5G deployment.

Our future work will include more diverse and complex scenarios, i.e., we plan to extend the setup towards a multi-cell deployment, moving clients, and varying video and client characteristics. We will also enlarge our set of ML techniques, among others by including classifiers, and we plan to investigate the requirements for the data provided by NWDAF, e.g. in terms of granularity.

Acknowledgments

This work has been funded by the German BMBF Software Campus Grant “BigQoE” (01IS17052).

References

[1] V. Aggarwal, E. Halepovic, J. Pang, S. Venkataraman, and H. Yan. 2014. Prometheus: toward quality-of-experience estimation for mobile apps from passive network measurements. In *HotMobile 2014*. ACM.
[2] I. Alawe, A. Ksentini, Y. Hadjadj-Aoul, and P. Bertin. 2018. Improving traffic forecasting for 5G core network scalability: A Machine Learning approach. *IEEE Network* 32, 6 (2018), 42–49.
[3] A. Angus, C. Murudkar, and R. Gitlin. 2019. Machine Learning for QoE Prediction and Anomaly Detection in Self-Organizing Mobile Networking Systems. *International Journal of Wireless & Mobile Networks* 11 (2019).
[4] D. Basak, S. Pal, and D. Patranabis. 2007. Support vector regression. *Neural Information Processing-Letters and Reviews* 11, 10 (2007).
[5] D. Bega, M. Gramaglia, M. Fiore, A. Banchs, and X. Costa-Perez. 2019. DeepCog: Cognitive Network Management in Sliced 5G Networks with Deep Learning. (2019).
[6] T. Begluk, J. Baraković Husić, and S. Baraković. 2018. Machine learning-based qoe prediction for video streaming over lte network. In *INFOTEH 2018*. IEEE, 1–5.
[7] G. Dimopoulos, I. Leontiadis, P. Barlet-Ros, and K. Papagiannaki. 2016. Measuring video QoE from encrypted traffic. In *IMC 2016*. ACM, 513–526.
[8] Cisco Visual Networking Index. 2018. Forecast and Trends, 2017–2022. *Cisco Systems* (2018), 1–7.
[9] I. Jolliffe. 2011. *Principal component analysis*. Springer.
[10] M. Katsarakis, R. Teixeira, M. Papadopouli, and V. Christophides. 2016. Towards a causal analysis of video QoE from network and application QoS. In *Internet-QoE 2016*. ACM, 31–36.
[11] M. Khokhar, T. Ehlinger, and C. Barakat. 2019. From Network Traffic Measurements to QoE for Internet Video. In *IFIP Networking Conference*.
[12] Y. Lin, E. Oliveira, S. Jemaa, and S. Elayoubi. 2017. Machine learning for predicting QoE of video streaming in mobile networks. In *ICC 2017*. IEEE.
[13] T. Mangla, E. Halepovic, M. Ammar, and E. Zegura. 2018. eMIMIC: Estimating HTTP-based Video QoE Metrics from Encrypted Network Traffic. In *TMA 2018*. IEEE.
[14] A. Martin, J. Egaña, J. Flórez, J. Montalbán, I. Olaizola, M. Quartulli, R. Viola, and M. Zorrilla. 2018. Network resource allocation system for QoE-aware delivery of media services in 5G networks. *IEEE Transactions on Broadcasting* 64, 2 (2018).
[15] J. Nightingale, P. Salva-Garcia, J. Calero, and Q. Wang. 2018. 5G-QoE: QoE modelling for ultra-HD video streaming in 5G networks. *IEEE Transactions on Broadcasting* 64, 2 (2018), 621–634.
[16] I. Orsolic, D. Pevec, M. Suznjevic, and L. Skorin-Kapov. 2017. A machine learning approach to classifying YouTube QoE based on encrypted network traffic. *Multimedia tools and applications* (2017).
[17] I. Oršolić, P. Rebernjak, M. Sužnjević, and L. Skorin-Kapov. 2018. In-Network QoE and KPI Monitoring of Mobile YouTube Traffic: Insights for Encrypted iOS Flows. In *CNSM 2018*. IEEE.
[18] H. Peng, F. Long, and C. Ding. 2005. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 8 (2005), 1226–1238.
[19] M. Polese, R. Jana, V. Kounev, K. Zhang, S. Deb, and M. Zorzi. 2018. Machine learning at the edge: A data-driven architecture with applications to 5G cellular networks. *arXiv preprint arXiv:1808.07647* (2018).
[20] W. Robitza, S. Göring, A. Raake, D. Lindegren, G. Heikkilä, J. Gustafsson, P. List, B. Feiten, U. Wüstenhagen, M. Garcia, et al. 2018. HTTP adaptive streaming QoE estimation with ITU-T rec. P. 1203: open databases and software. In *MMSys 2018*. ACM, 466–471.
[21] G. Seber and A. Lee. 2012. *Linear regression analysis*. Vol. 329. John Wiley & Sons.
[22] M. Seufert, P. Casas, N. Wehner, L. Gang, and K. Li. 2019. Stream-based Machine Learning for Real-time QoE Analysis of Encrypted Video Streaming Traffic. In *ICIN 2019*. IEEE.
[23] A. Varga. 2010. OMNeT++. In *Modeling and tools for network simulation*. Springer, 35–59.