# Improved VoLTE QoE Estimation Procedure using Network Performance Metrics

Manuela Vaser
*Dpt of Electronic Engineering*
*University of Rome "Tor Vergata"*
Rome, Italy
manuela.vaser@uniroma2.it

Giulio Maggiore
*Core Transport and Network Service Exposure*
*TIM Italy, Via Oriolo Romano*
Rome, Italy
giulio.maggiore@telecomitalia.it

***Abstract* – In this paper a new method to improve the mathematical estimation of Quality of Experience (QoE) perceived by the end user of a Voice call over LTE (VoLTE) is presented.**

**The new estimation leads to Mean Opinion Score (MOS) values more realistic than those available on network monitoring processes.**

**The optimization of the QoE is calculated as a function of several QoS parameters, in particular packet loss, network jitter delay and jitter buffer dimension.**

**The improvement has been realized by identifying the QoE and QoS parameters, named respectively Key Quality Indicators (KQI) and Key Performance Indicators (KPI), for VoLTE services, and then applying a QoE/QoS mathematical relationship in order to adjust the expected QoE level on the basis of QoS.**

***Keywords* – QoE, QoS, user experience, LTE, VoLTE, MOS.**

## I. INTRODUCTION

Telecommunications are today more and more user-centric and for this reason, the network operators must be able to measure the user experience in order to prevent quality issues on the network, before the user perceives them.

In order to do this, it is essential to establish a relation between Quality of Experience (QoE) perceived by the customer and Quality of Service (QoS) offered by the network.

Quality of Service (QoS) [1,2] represents "the ability of the network to provide a service at a guaranteed performance level", and it is measured by quantitative parameters named *Key Performance Indicators (KPI)*, such as packet loss, delay and jitter [3,4].

On the other hand, Quality of Experience (QoE) [5,6,7] defines the quality subjectively perceived by the end-user, and gives information on how well the network meets the user's needs. QoE is measured by qualitative parameters, named *Key Quality Indicators* (*KQI)* [8], such as "very good", "good", "poor".

For what concern of the quality of a real time service like Voice over LTE, Mean Opinion Score (MOS) [9] is the most important parameter to take into account.

In order to obtain the QoE perceived by the customer, MOS monitoring should be carried out at the end device, as close as possible to the user.

This operation today is made in active test scenarios, where ad-hoc commercial test devices are equipped with specific tools, in order to predict Perceptual Objective Listening Quality Assessment (POLQA), a Full Reference model established in ITU-T Rec. P.863 [10].

This estimation can be done only with few handset, so the value is not statistically relevant, as the number of samples is very small.

From a network point of view, the quality is evaluated with G.107 MOS [9], which is easily obtained by monitoring operations of several network interfaces, allowing to obtain both MOS, both QoS parameters.

However, MOS value monitored in this case may not be as realistic as the end user one, due to the fact that measurement is done on the network, and not on user device. Anyhow, it is statically relevant, as there are hundred thousand of available samples resulting from monitoring probes.

With the procedure presented in this paper, using parameters obtained by the network probes, it is possible to obtain a realistic estimation of QoE, together with the statistical relevance of the network approach. This procedure is seamless for the user.

Furthermore, it can be useful, in particular conditions, to express the mathematical dependence on QoE from more than one QoS parameter.

In this work, QoE is expressed as a function of more than one QoS parameter, by following an optimization procedure relying on monitoring operation in specific network sections.

The paper is organized as follows. Section II contains the related works found in literature. Section III shows the network scenario in which the procedure has been applied. Section IV illustrates the core of the procedure. Section V shows the obtained results. Finally, in Section VI some conclusions and remarks are described.

## II. RELATED WORKS

The User Experience assessment in IP real time services has been object of several studies in literature.

- In [11], QoE of Video over LTE (ViLTE) has been evaluated by means of a test bed integrating LTE network with a user equipment under test. Differential quality score (DMOS) derived from G.107 MOS has been calculated in different network conditions and in relation with different values of QoS parameters such as packet loss and packet delay.

- In [12], a modified version of E-Model has been used to evaluate MOS for Voice over IP (VoIP) services.
- In [13], by using a simulated model, delay and jitter has been observed in a VoIP application.
- In [14], VoIP MOS in 3G mobile network has been evaluated, for Over The Top (OTT) applications like Skype.
- In [15], a closed-loop power control algorithm has been applied to downlink VoLTE radio, and simulation results show the improvements that can be obtained in the VoLTE MOS of an indoor small cell scenario.
- In [16], a dynamic adaptation algorithm of joint source-channel code rate is used to improve the VoLTE QoE. The wideband E-model is used to assess the voice quality.

All above mentioned research works use G.107 MOS as KQI for QoE assessment. In fact, G.107 MOS is easy to obtain, by monitoring operations on network, seamless for the end user, but at the same time it can be poorly representative of real QoE, as the evaluation is made "far" from user device.

On the other hand, test session on for QoE assessment through POLQA tools, can give significant results, but they cannot be implemented in real traffic scenarios.

In this context, a mathematical relation evaluating QoE with QoS parameters can be useful for network operators, as it could allow to obtain QoE as realistic as possible, by using parameters easy to collect.

For this reason, a series of studies can be found in literature [17, 18, 19], deriving QoE from QoS by use of a mathematical relationship holding between the two.

## III. NETWORK SCENARIO

The QoE optimization procedure has been applied on a real scenario, by monitoring network during voice calls.

Voice over LTE (VoLTE) is the voice service delivered in LTE all-IP Packet-Switched domain, in which the network operator provides the service by means of its own network, so that he can manage all phases of the service.

VoLTE calls provisioning is made possible by specific architectural elements composing the Internet Multimedia Subsystem (IMS) section, standardized by 3GPP in [20, 21], responsible for signaling operations of a VoLTE call.

In Figure 1 it is possible to observe the network architecture, in which three different points are marked as PCO, standing for Point of Control and Observation. They indicate the sections of the network in which KPIs are collected by monitoring probes, at application level. PCO#1, PCO#2 and PCO#3 are related to probes for $S_{1u}$, $S_{11}$ and $M_w$ network interfaces monitoring, respectively.

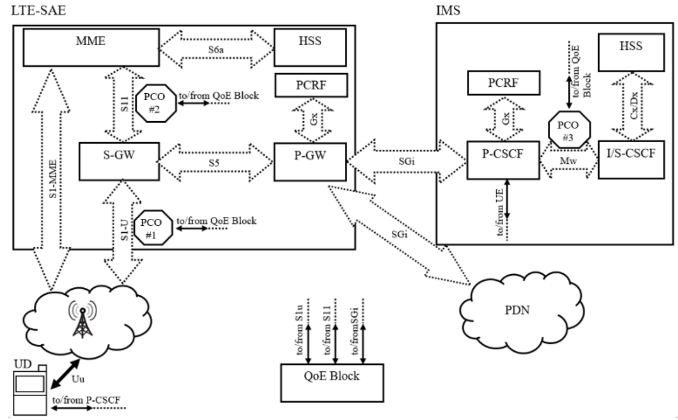

Fig. 1. Network scenario

LTE is an all-IP environment, in which IP packets of the same RTP flow can follow different paths, and for this reason can be affected by different delays, and could be received in a different order. For this reason, different parameters have been collected in the network scenario:

- Packet loss, expressed in number of lost RTP packets;
- Network jitter delay, which is the difference between the timestamps, contained in the RTP protocol header, of consecutively received packets, and it is expressed in ms;
- Jitter buffer dimension, expressed in ms, which is the additional delay of each packet with the aim to re-establish the right order of consecutive packets
- MOS [9], adimensional subjective parameter for the evaluation of speech quality, with values in the range of 1-5.

In addition to the network elements of Figure 1, specific test devices have been used, in which ITU-T P.863 POLQA [10] value is obtained, by means of ad-hoc calculation tools. In this context, POLQA has been used as a comparison entity, as it is the most precise and realistic expression of the QoE of a voice call. It can be useful to remark that the optimization procedure presented here is different from POLQA one, and that QoE is obtained by means of network performance indicators.

The quality parameters collected in PCOs are the input of the QoE block, in which optimization procedure is implemented.

## IV. VOLTE QOE BLOCK

Before the description of the optimization procedure, it is necessary to describe how QoE and QoS are correlated, and why they have to be studied with common understanding.

QoE parameters, the KQIs, are influenced by the total end-to-end network effects, and by the user expectations, degree of delight or annoyance during the fruition of a service. For this reason, it is difficult to express QoE in an objective and mathematical way, as it can depend both from the context in which the user operates and both from the quality negotiated with the network operator for a specific service.

For this reason, in an operational context, a possible solution is to establish a relation between user expectations and the QoS, as it is assured by the differentiated management of the various traffic classes, and it is a technical, objective and network oriented entity. It can be mathematically expressed by KPIs defined by standard organizations, obtained by monitoring operations on the network, that do not depend from the service under monitoring.

The standardized LTE/SAE network focuses on the definition of the network performance in terms of the QoS, but does not define the relationship between QoE and QoS. A complete quality analysis needs to express this relationship in order to maximize the QoE, during the service delivery.

As stated in [22, 23] concerning VoLTE QoE, the most relevant KQI is the Mean Opinion Score (MOS), an adimensional subjective parameter for the evaluation of voice-call quality, with values in the range between 1 and 5. MOS is estimated with the E-Model algorithm, defined in [9] [24-26], whose output is the R-Factor. Table 1 below illustrates the matching between R-factor and MOS values.

According to [9], the R-factor can be expressed as follows:

$$R = R_0 - I_s - I_d - I_{e,eff} + A \tag{1}$$

MOS calculation method through (1) indicates different factors impairing the quality of a voice call, described in detail in [9]. In this context, an element in particular has to be highlighted, Ie,eff. Ie,eff stands for Equipment Factor, and represents the impairment caused by low bit rate CODEC, and packet loss. In ITU-T G.107 recommendation, it has been expressed by:

$$I_{e,eff} = I_e + (95 - I_e)\frac{P_{pl}}{P_{pl} - B_{pl}} \tag{2}$$

where $I_e$ and $B_{pl}$ are parameters that can assume values as indicated in [9], and $P_{pl}$ is the packet loss probability, in a range of values between 0 and 1, calculated as:

$$P_{pl} = 1 - \frac{m}{n} \tag{3}$$

where m is the number of RTP packets received, and n is the number of RTP packets sent, with uncorrelated losses. Ppl is one of the most important KPIs involved in the QoE estimation, and it has been considered in the optimization procedure described here.

TABLE I
R-FACTOR MOS MATCHING [18]

| User Satisfaction Level | R-Factor | MOS |
|---|---|---|
| Maximum using G.711 | 93 | 4.4 |
| Very satisfied | 90-100 | 4.3-5 |
| Satisfied | 80-90 | 4-4.3 |
| Some users satisfied | 70-80 | 3.6-4 |
| Many users dissatisfied | 60-70 | 3.1-3.6 |
| Nearly all users dissatisfied | 50-60 | 2.6-3.1 |
| Not recommended | Less than 50 | 1-2.6 |

Moreover, in packet switched domain, the packets of the same RTP flow related to a VoLTE call can be delivered to destination by following different network paths. For this reason, different delays can affect the packets, and this may cause the packet delay variation, also known as network jitter [27], calculated as the floating average of differences between RTP timestamps [27], of consecutively received packets.

Network jitter is an important impairment to deal with, and at receiver side it is essential to re-order the packets in the right way. To this aim, jitter buffers can compensate this effect by forcing an additional delay to packets, in order to re-establish the original order. In this process, if the end-to-end delay of a packet is greater than jitter buffer dimension, the packet is discarded, and the voice call quality get worse.

In literature, for what concern voice services in IP context, a QoE/QoS mathematical relationship is implemented, known as IQX hypothesis [18], allowing the calculation of MOS (as QoE parameter), as function of one KPI (as QoS parameter):

$$QoE = \alpha e^{-\beta QoS} + \gamma \tag{4}$$

α, β, γ are the regression factors, calculated by means of non-linear regression operations. IQX has been validated in a simulated IP context, and considering packet loss or jitter as KPIs.

However, under certain real network conditions, in evaluating the relation between MOS and packet loss, it can happen that MOS varies regardless to packet loss, that remains always the same. This situation is illustrated in Figure 2, in which real packet loss and MOS are showed, respectively on x-axis and on y-axis, obtained by probes on the network scenario depicted in Figure 1.

For what concern the quality perceived by the user, by applying IQX in (4), the fitting Matlab code was not able to establish a relationship. So, in this case, it is impossible to mathematically establish the relation between QoE, expressed by MOS, and QoS, expressed by packet loss by using pure IQX. Then, it is necessary to introduce further elements that can allow an estimation of QoE.

In order to express the importance of jitter on MOS

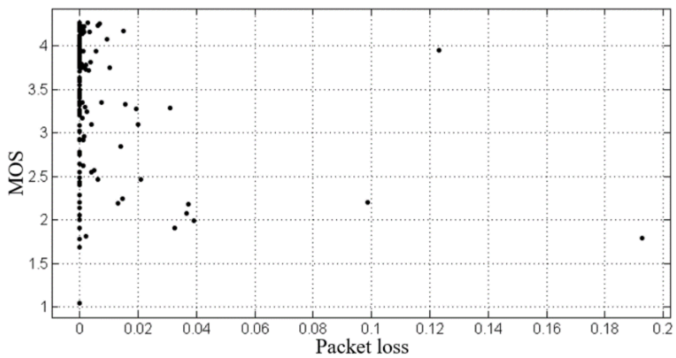


Fig. 2: MOS vs packet loss

calculation, in [19] the combined effect of jitter, packet loss and jitter buffer is represented named Packet Loss effective:

$$P_{pl,eff} = 1 - \left(1 - P_{pl}\right)\left(1 - P_{jitter}\right) \quad (5)$$

where jitter is expressed as a Pareto probability by writing:

$$P_{jitter} = \frac{\left(1 - \frac{0.1x}{\sigma}\right)^{20}}{2} \quad (6)$$

calculated according to jitter buffer dimension and network jitter delay, respectively $x$ and $\sigma$, both expressed in ms.

The diagram in Figure 3 represents the framework of the procedure, whose purpose is the estimation of QoE of a VoLTE call through the calculation of MOS value using performance parameters obtained by monitoring operations on the network.

First of all, Voice Call set up procedure is executed by the network.

After that, VoLTE call begins. From this moment, in the procedure two operations run in parallel, and they are represented by two branches.

In the first branch, probes on PCOs give the measurements of voice data and performance parameters of RTP packets, and collect them in vectors. These vectors are used in input of the "Self Learning" process, in which IQX in (4) is implemented.

The output of the "Self Learning" phase are regression parameters α, β, γ, together with information about jitter buffer dimension [ms], used in the second branch, in the "Correlation Function", which is an exponential function like the IQX one.

However, in this procedure, expression of $P_{pl,eff}$ (5) is used as QoS KPI, in order to obtain a "Refined Correlation Function":

$$VM\_MOS = \alpha e^{-\beta\left[1-\left(1-P_{pl}\right)\left(1-P_{jitter}\right)\right]} + \gamma \quad (7)$$

In (7) the calculation of VM_MOS contains an explicit indication to two network KPI: packet loss and network jitter delay.

The action nodes in the second branch are executed as long as the VoLTE call is in progress.

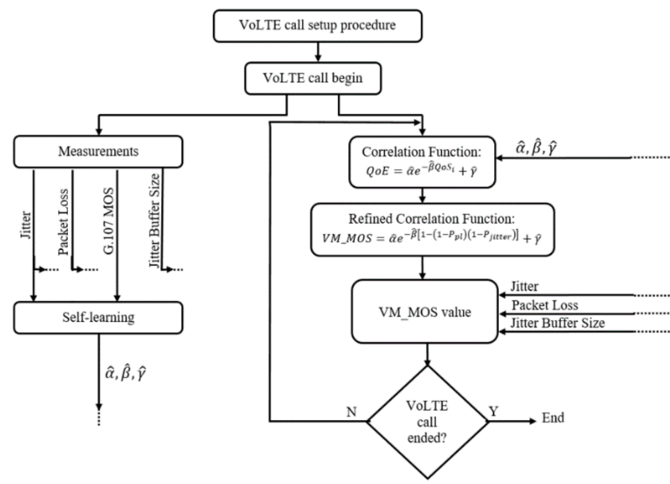


Fig. 3: VM_MOS procedure

## V. Obtained Results

Different sets of VoLTE data have been analyzed, and User Experience has been estimated, by using the optimization framework in Figure 3.

The network has been monitored for seven days, during which probes on PCO#1, PCO#2, PCO#3 indicated in Figure 1 collected the parameters listed below:

- G.107 MOS of the calling;
- G.107 MOS of the called;
- Network jitter delay of the calling [ms];
- Network jitter delay of the called [ms];
- Packets lost of the calling;
- Packets lost of the called.

All parameters have been collected at application level.

At the same time, POLQA values have been collected by means of test devices, and compared to VM_MOS ones calculated in the optimization framework.

In order to evaluate the effectiveness of optimization procedure, VM_MOS values calculated in the process have been compared with POLQA values collected by test devices end-to-end. The comparison is showed in Table 2, in which performances of calling side are indicated.

The monitoring operations lead to thousands of values per day, as the probes give one value per call and the sampling can be every second for each performance parameter. All values are in input to the optimization procedure illustrated in Figure 3.

For the sake of readability, values indicated in Table 2 shows the average values over daily measurements.

The second column of the table reports G.107 MOS value, as monitored from the network probes. VM_MOS calculated with

TABLE II
RESULTS

| | G.107 MOS | VM_MOS | POLQA |
|---|---|---|---|
| Day 1 | 3.21 | 3.39 | 3.46 |
| Day 2 | 3.14 | 3.44 | 3.46 |
| Day 3 | 3.24 | 3.4 | 3.46 |
| Day 4 | 2.89 | 3.4 | 3.47 |
| Day 5 | 3.15 | 3.39 | 3.46 |
| Day 6 | 3.16 | 3.33 | 3.48 |
| Day 7 | 3.17 | 3.41 | 3.48 |

the optimization procedure is indicated in third column. Finally, the last column shows the correspondent values of POLQA obtained from test devices.

As can be seen by comparing the numeric results, the optimization procedure can allow a MOS value that differ from POLQA of 0.02 in the best case (the worst case is 0.15), better than G.107 MOS values.

## VI. CONCLUSIONS

In this paper, a new method to improve the mathematical estimation of Quality of Experience (QoE) perceived by the end user of a Voice call over LTE (VoLTE) systems has been showed.

This method has been compared to those following Mean Opinion Score (MOS) specification.

The new estimation called VM_MOS provides values comparable with POLQA (Perceptual Objective Listening Quality Assessment) established in ITU-T Rec. P.863.

With the VM_MOS procedure, using a network monitoring approach, it is possible to obtain a realistic estimation of QoE, together with the statistical relevance of the network approach, and being seamless for the user.

## REFERENCES


[1] 3GPP Technical Specification TS 23.207 V10.0.0, "*End-to-end Quality of Service (QoS) Concept and Architecture*" March 2011.

[2] 3GPP Technical Specification TS 23.107 V11.0.0, "*Quality of Service (QoS) Concept and Architecture*" June 2012.

[3] 3GPP TS 32.455 version 10.0.0 Release 10 *"LTE; Telecommunication management; Key Performance Indicators (KPI) for the Evolved Packet Core (EPC)"*.

[4] 3GPP TS 32.450 version 8.0.0 Release 8 "*Universal Mobile Telecommunications System (UMTS); LTE; Telecommunication management; Key Performance Indicators (KPI) for Evolved Universal Terrestrial Radio Access Network (E-UTRAN): Definitions*".

[5] ITU-T Recommendation P.10/G.100, "*Vocabulary for performance and quality of service*", 2006.

[6] ETSI Technical Report 102 643 V1.0.2."*Human Factors (HF); Quality of Experience (QoE) requirements for real-time communication services.*" October 2010.

[7] S. Barakovi, L. Skorin-Kapov, "*Survey and Challenges of QoE Management Issues in Wireless Networks*", Journal of Computer Networks and Communications, 2013.

[8] TeleManagement FORUM, The Open Group, "*SLA Management Handbook – Volume 4: Enterprise Perspective*", October 2004.

[9] ITU-T Recommendation G.107, *"The E-Model: A Computational Model for Use in Transmission Planning"*, December 2011.

[10] ITU-T Recommendation P.863, "*Perceptual Objective Listening Quality Prediction*", March 2018.

[11] C. Callegari et al, *"Experimental Analysis of ViLTE Service"*, IEEE Access, published on April 2nd, 2018.

[12] E.Silva et al, *"Mean Opinion Score Measurements Based on E-Model During a VoIP call"*, The Eleventh Advanced International Conference on Telecommunications, 2015.

[13] M.H. Miraz et al, "*Simulation and Analysis of Quality of Service (QoS) Parameters of Voice over IP (VoIP) Traffic through Heterogeneous Networks"*, International Journal of Advanced Computer Science and Applications, Vol. 8, N. 7, 2017.

[14] P. Wuttidittachotti, T. Daengsi, *"Quality Evaluation of Mobile Networks Using VoIP Applications: A case study with Skype and LINE based-on Stationary Tests in Bangkok"*, I.J. Computer Network and Information Security, 2015, 12, 28-41.

[15] F.B. Mismar, B.L. Evans, *"Q-Learning Algorithm for VoLTE Closed-Loop Power Control in Indoor Small Cells"*, 2018.

[16] D. Nguyen, H. Nguyen, *"A Dynamic Rate Adaptation algorithm using WB E-Model for voice traffic over LTE network"*, 2016.

[17] T. Hossfeld, D. Hock, P. Tran-Gia, K. Tutschku, M. Fielder, "*Testing the IQX Hypothesis for Exponential Interdependency between QoS and QoE of Voice Codecs iLBC and G.711*", 18th Seminar on Quality Experience, Sweden 2008.

[18] Fiedler M., Hossfeld T., Tran-Gia P., Blekinge Inst. of Technol., Karlskrona, Sweden; "*A generic quantitative relationship between quality of experience and quality of service* ", Network, IEEE (Volume:24 , Issue: 2 ), 2010.

[19] A. Kovac, M. Halas, M. Orgon, M. Voznak, '*E-Model MOS Estimate Improvement through Jitter Buffer Packet Loss Modelling*'', Information and Communication Technologies and Services, Vol 9, n. 5, 2011.

[20] 3GPP Technical Specification TS 23.228 V11.8.0, "IP Multimedia Subsystem (IMS)", June 2013.

[21] 3GPP Technical Specification TS 22.228 V12.6.0, "Service Requirements for the Internet Protocol (IP) Multimedia Core Network Subsystem (IMS)" June 2013.

[22] S. Forconi, M. Vaser, "*4G LTE Architectural and Functional Models of Video Streaming and VoLTE Services*", the Seventh International Conference on Ubiquitous and Future Networks (ICUFN 2015). Sapporo, Japan, July 2015.

[23] M. Vaser, *"Analisi Funzionale del Servizio Voce in LTE e Studio Simulativo delle Relazioni tra QoE e QoS"*, PhD Thesis, XXVII Cycle.

[24] ITU-T Recommendation P.862, "*Perceptual Evaluation of Speech Quality (PESQ): An Objective Method for end-to-end Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs*", February 2001.

[25] ITU-T Recommendation P.862.2, "*Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs*", November 2007.

[26] ITU-T G.109 Amendment 1, *"Definition of Categories of Speech Transmission Quality Amendement 1: New Appendix I – The E-Model Based Quality Contours for Predicting Speech Transmission Quality and User Satisfaction from Time-Varying Transmission Impairments"*, January 2007.

[27] RFC 3393 "*IP Packet Delay Variation Metric for IP Performance Metrics (IPPM)*".