A Model for Mapping between the Quality of Service and Experience for Wireless Multimedia Applications

Daniel Robalo and Fernando J. Velez
Instituto de Telecomunicações-DEM
Universidade da Beira Interior
6201-001 Covilhã, Portugal
drobalo@lx.it.pt, fjv@ubi.pt

***Abstract*— With nowadays evolution of wireless network technologies and the growth of novel wireless multimedia services it is critical for network and service providers to be able to assess user's perceptual multimedia quality, for commercial and technical reasons. As so, we present wireless network services and applications characterization requirements, updated to current specifications. We considered recommendations for some major regulations entities, such as the ITU, 3GPP, WiMAX Forum and the European project EU-MESH. In the case of audio and video applications, the latest and most used encoding schemes were addressed by considering the WiMAX Forum recommendations. In turn, as not only the Quality of Service (QoS) needs to be addressed, we propose a model for the mapping between QoS parameters and user's Quality of Experience (QoE) for multimedia applications. Measurable QoS parameters can be gathered throughout the network and mapped into an effective QoE estimation. The model has been obtained by means of mathematical regression and the fittings are validated by analysing the coefficient of determination, R^2, between the results from our model and mean opinion scores experimental values. The achieved results are R^2 = 0.94, 0.84 and 0.98 for gaming, video and audio, respectively. The corresponding mean square error is approximately zero in all cases.**

Keywords- Applications; Services; Requirements; Characterization; Quality of Experience, Quality of Service, User's satisfaction, Mean Opinion Score

I. INTRODUCTION

The permanent evolution of wireless network technologies allows for improved data rates and coverage areas while facilitating new multimedia and mobile services. Considering this evolution of services and applications, operator's success does not only depend on their Quality of Service (QoS), but also if it meets the end user's expectations. With the increasing competition, improving the quality of the offered services, as perceived by the users (Quality of Experience, QoE), becomes important as well as a significant challenge to service providers with a goal to minimize the customer churn yet maintaining their competitive edge [1]. Thus, it is critical to be able to assess, predict and possibly control the end-to-end perceptual multimedia (e.g., voice and video) quality for commercial and technical reasons [2].

Our goal is to address the requirements and constraints that must be considered for specific and well detailed services and applications, considering the trends of the evolution of nowadays mobile services. The requirements for applications and services specified by the ITU [3] and 3GPP [4] have been considered. Besides, the characterization of video and voice encoding proposed by the WiMAX Forum [5] has also been taken into account, as well as the requirements specified by EU-MESH [6], an FP7 Working Group European project.

The key QoS parameters that affect users have been addressed for the proposed applications characterization. Assuming that next generation wireless networks should provide superior data rates (and considering updated video and voice encoding rates) the typical downlink and uplink transmission rates have, whenever possible, been chosen by considering the highest values provided by the regulations entities mentioned above.

QoS is generally defined in terms of network delivery capacity and resource availability but not in terms of the satisfaction to the end-user. QoE is very subjective in nature, therefore the most accurate approach to evaluate it is the subjective quality assessment, since there is no better indicator of personal quality then the one given by a human being. This assessment is essentially a subjective measurement of the network performance at the service level. It is important that QoE is expressed as a function of the network and equipment that influence user behaviour and result in a certain level of QoE. Therefore, QoE data should succeed whenever possible in combining both user experience and technical measures; for example, to provide an equation for the user experience when using a particular service with known levels of QoS [7]. As such, QoS metrics gathered from various parts of the network must be mapped onto QoE targets, facilitating the inclusion of the end-user perception into the QoE model.

In this work, we propose a function that characterizes the relation between a set of QoS parameters and the corresponding QoE, providing network and service providers a framework to evaluate user's satisfaction. The novelty of our approach is to introduce a unified model for multimedia applications. Four types of applications are considered: gaming, video, web-browsing and audio. As so, the final model is formed by considering fittings obtained for every individual application. Some are based on Mean Opinion Score (MOS) measurements available in the literature (or provided by other researchers). The remaining ones, for which we could not obtain MOS values, are achievable based on some well-known ITU-T G.1030 and G.107, which, in turn, are also based on

MOS experiments. We are performing regression analysis by means of computing the mean square error and coefficient of determination. These parameters characterize how well the model fits to the experimental results and ITU-T recommendations.

The remainder of this paper is organized as follows. Section II describes the service level parameters for multimedia applications. Section III presents the proposed requirements for applications and services. Section IV addresses the mapping between QoS parameters and the QoE Section V addresses the pioneering QoE model for all the considered applications. Finally, Section VI presents some conclusions and discusses some opportunities to improve this research.

II. Service level parameters for different applications

QoS mechanisms are applied for monitoring the achieved level of performance of an application. QoS parameters are the most accessible and measurable metrics to assess, in real time, the overall performance of the network. We consider the key parameters impacting the users experience as they are defined by ITU-T [3]: the delay, the delay variation and the information loss. Additionally, application specific service level metrics are included in our model. For on-line gaming, a demanding interactive application, we consider the ping or round-trip time. In the case of video applications, the encoding bitrate is also assumed. Applications are generally classified by four different categories, audio, video, data and background, which, in turn, can be divided into various services, whose levels of quality are defined by the ITU [8].

A. Audio

The minimum quality for audio services is similar to ordinary telephony, i.e., the quality that allows a reasonable ease for understanding speech. However, for most audio services, the quality objectives are generally much higher, ranging from a good comfort for listening to speech to the highest quality in reproducing music or other sounds.

B. Video

Video quality is a measure of the ability of a video transmission system to accurately reproduce moving scenes. Video services quality objectives are expressed in terms of spatial and temporal resolution. However, other parameters may also be relevant (e.g., distortion, signal loss or errors).

C. Data

From the user point of view, a prime requirement for any data transfer is to guarantee essentially zero loss of information. At the same time, delay variation is not generally noticeable to the user, although there needs to be a limit on synchronisation between media streams in a multimedia session. Services tend to distinguish themselves on the basis of the delay which can be tolerated by the end-user from the time the source content is requested until it is presented to the user.

D. Background

The only requirement for this type of application is that essentially error free information should be delivered to the user. However, there still is a delay constraint since, for any practical purpose, data is useless if it is received too late.

E. Context-based information

Besides the previous applications defined by ITU, another can be considered from EU-MESH [6], the context-based information application, which belongs to the location dependent services category. This category includes services which are expected to play a significant role in emerging ubiquitous broadband access networks. This application includes services such as one-way location-based broadcasting of local news, travel information, advertisement, and the transmission of multimedia content/information that also involves user feedback (positioning, travel directions, and so on).

III. Proposed requirements for services and applications

In this work, applications requirements from the ITU [3], 3GPP [4], WiMAX Forum [5] and the EU-MESH [6] have been gathered. The goal is to merge and update all recommendations/documents. The trend of wireless technologies has been to provide higher throughputs and QoS. Considering that next generation wireless networks will continue to provide superior data rates, the downlink and uplink speeds were chosen, whenever possible by considering the highest value from the gathered applications requirements. Additionally, the WiMAX Forum [5] provides in-depth information regarding requirements for audio and video applications as a function of their encoding bit rate and other parameters. Our proposal is presented in Table I.

IV. Mapping between the Quality of Service and Experience

A. Gaming

The unified model is divided into four sub-models, one for each application. Our proposal for the sub-model for gaming is based on the MOS measurements from [9]. The performed campaign emulates the delay (in the form ping) and jitter (the variance in packet arrival times at the destination) during subjective experiments. Technical details on the network emulator (Netem) can be found in [9], as well as the range of variation for the QoS parameters considered during the trials.

During two test sessions, six gamers simultaneously played Quake IV in a Free-For-All (FFA) setting. All players had prior experience with First Person Shooter (FPS) games. The FFA setting is one of the most popular FPS game modes, in which players constantly engage in battles to collect the most frags (kills) during a certain time limit. The simulation of each Netem scenario lasted five minutes. After this time period, the gamers were asked to give their opinion on the gaming quality. They were able to select one of the following five values for the opinion score, motivated by the ITU-T ACR scale [10], 5 for excellent gaming quality; 4 for good gaming quality; 3 for fair gaming quality; 2 for poor gaming quality and 1 for bad gaming quality. During the experiment, the gaming server kept track of all kills and deaths of the participating players. The purpose was to identify a relation between these objective performance measures and the chosen impairment factors.

TABLE I. APPLICATIONS AND SERVICES REQUIREMENTS

Service	Typical data rates [kbps] DL	Typical data rates [kbps] UL	Interactive [6]	Key performance parameters: One-way delay	Key performance parameters: Delay variation	Key performance parameters: Information loss	Transmission mode [6]
Audio							
VoIP, tele-conferencing	5-64[1] [5]		Yes	< 150ms (limit <400ms) [2 - 4]	<1ms [2 - 4]	<3% [3] [4]	Unicast/ multicast
Internet radio	16-64 [6]	N.A.	No	<10s [6]	<2ms [4]	<1% [6]	Broadcast
Voice mail	30-50 [6]	4-25	Yes	<2s [6]	N.A.	<3% [6]	Unicast
Video							
Video-telephony/conferencing	32 to 1024[1] (Typical = 384) [5]		Yes	< 150ms (limit <400ms) [2 - 4]	<1ms [2 - 4]	<1% [3] [6]	Unicast/ multicast
IPTV	1024-3072 [6]	N.A.	No	<10s [6]	<2ms [4]	<1% [6]	Broadcast
Mobile TV	28-512 [6]						
VoD (SD)	1024-3072 [6]	<8 [6]	Yes				Unicast
VoD (HD)	7168-10240[2] [5]						
On-demand streaming media	32-384 [6]						
Video surveillance				<2s [6]			
Data							
Real-time gaming	32-64 [6]		Yes	<75ms [4] [6]	N.A.	0% [6]	Unicast
Interactive gaming	4-25 [6]			<150ms [5]			
Collaborative working	>500 [6]	4-25 [6]		<2s/transact. [6]			
ASP services							
E-commerce							
Control of remote devices	<28 [6]			<10s [6]			
Peer-to-peer file sharing	>500 [6]			<15s (acceptable <60s) [6]			
User-created content sharing	32-64 [6]	>500 [6]		<2s/upload [6]			
Web browsing	10 to 2048 [6]	4-25 [6]		<2s/page (acceptable<4s/page) [3] [6]			
Email (server access)	>500 [6]			<2s (acceptable <4s) [3] [6]			
File transfer	>2048 [6]	4-25 [6]		<15s (acceptable <60s)			
Telnet	4-25 [6]			<200 ms [3]		0% [4] [6]	
Instant messaging (IM)				250 ms [6]		0% [6]	
Background							
Online chatting	4-25 [6]		Yes	250 ms [6]	N.A.	0% [6]	Unicast
Fax ("real-time")	N.A.			< 30 s/page [3]		$<10^{6}$ [3]	
Usenet	N.A.			Can be several minutes		0% [3]	Broadcast
Context-based Information [6]							
Location-based multimedia broadcast	32-384	N.A.	No	<10s	N.A.	<1%	Broadcast
Location-based interactive multimedia		<8	Yes			N.A.	Unicast
Location-based on demand						<1%	
Alert/notification, advertisement	<28	N.A.	No			0%	Multicast
Presence-based applications		<28	Yes				Multicast
Personalized content	32-64	<8		N.A.			Unicast

[1] Depends of encoding schemes
[2] H.264 video compression

The gaming test setup consisted of 6 clients PC's which were all connected to the gaming server through a Gigabit switch. The Netem network emulator was placed between the switch and the server. Both the gaming clients and the dedicated server were running the latest version of Quake IV at the time (v1.1).

Each player has to assess a total of 33 experimental scenarios. The scenarios included both network impairment caused by a single factor, as well as impairment consisting of combinations of the ping, p, and jitter, j, in ms, and the percentage of loss, ρ.

The results for these experiments were provided by Robert Kooij, one of the authors of the Muse G-Model [11]. We have computed the equation that better fits to the MOS experimental measurements. This regression leads to a sixth degree polynomial equation:

$$\begin{aligned} \mathrm{MOS} = {} & 4.16031 + 0.0349952 \times p - 0.104905 \times j - 0.267622 \times \rho - 0.00315486 \times p^2 + 0.011092 \times j^2 + \\ & 0.119224 \times \rho^2 + 7.84 \times 10^{-5} \times p^3 - 0.00048113 \times j^3 - 0.0135847 \times \rho^3 - 8.48647 \times 10^{-7} \times p^4 + 9.07 \times 10^{-6} \times \\ & j^4 + 0.000412947 \times \rho^4 + 3.92 \times 10^{-9} \times p^5 - 7.51346 \times 10^{-8} \times j^5 + 3.90407 \times 10^{-6} \times \rho^5 - 6.0582 \times 10^{-12} \times p^6 + \\ & 2.16813 \times 10^{-10} \times j^6 - 1.87 \times 10^{-7} \times \rho^6 \end{aligned} \quad (1)$$

The goodness of fit can be verified by the values of the coefficient of determination, R^2, and the mean square error, MSE. In this first case, the achieved results are R^2 = 0.938727189 and MSE = 0.066920089. It can be observed that R^2 is very close to 1 while the MSE is near 0.

B. *Video applications*

In the determination of video perceptual quality, approximating the users' perception of picture quality in real-time via image processing algorithms is a complex procedure, requiring significant processing power [12]. As a consequence, it is not feasible to consider parameters like colour of shape distortion or even I-P-B frame drops for MPEG compressed videos. Network-level connection quality indicators, such as packet loss, jitter or delay can be easily measured by the user terminal and monitored by the wireless network provider [12]. Consequently, it would be desirable to associate the QoS indicators to the actual QoE level, so as to have instantaneously an estimation of the viewed picture [12].

In this work, the MOS measurements performed in [13] are considered. The setup for this experiment is shown in Figure 1. The topology consists of a video server that initiates a push-based stream onto the network. The server can encode a video stream at various bitrates and frame rates. The server is connected to an IneoQuest (IQ) Sigulus G1-T box configured in the stimulus mode. In this mode the IQ can generate impairments to the flow by inducing loss, delay, or jitter. The flow is streamed to another IneoQuest G1-T configured in the *analysis and playback* mode. The stream is then passed onto a destination (Terminal-II) for visualization. Additional information on the experimental setup can be found in [13].

From these MOS results, it is also possible to compute the equation that better fits them. The result for this operation is also modelled by a sixth degree polynomial equation:

$$\begin{aligned}\mathrm{MOS} = {} & 3.2147 - 0.00266916 \times b_{rate} - 10.4811 \times d - 20.9894 \times \rho - 5.8875 \times 10^{-6} \times {b_{rate}}^{2} + 40.3305 \times d^{2} + \\ & 166.121 \times \rho^{2} + 1.449 \times 10^{-8} \times {b_{rate}}^{3} - 42.493 \times d^{3} - 730.016 \times \rho^{3} - 4.2939 \times 10^{-12} \times {b_{rate}}^{4} + 18.3884 \times d^{4} + \\ & 1764.47 \times \rho^{4} - 2.29851 \times 10^{-15} \times {b_{rate}}^{5} - 3.48213 \times d^{5} - 2069.09 \times \rho^{5} + 8.08679 \times 10^{-19} \times {b_{rate}}^{6} + \\ & 0.237418 \times d^{6} + 903.102 \times \rho^{6}\end{aligned} \tag{2}$$

where d is the delay in ms, ρ is the percentage of loss and b_{rate} is the bitrate, in kbps. The goodness of fit is confirmed by R^2 = 0.837862868 and the MSE = 0.196665978. As in the previous case, R^2 is very close to 1 while the MSE is near 0.

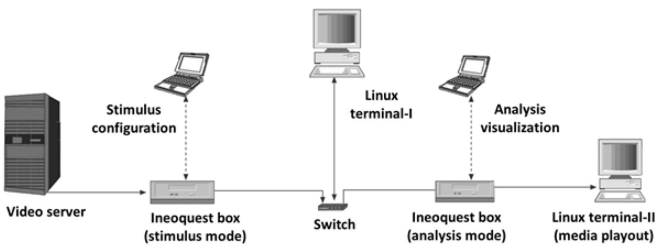

Figure 1. Sequences for the video the metric tests (extracted from [13]).

C. *Web-browsing applications*

We have not been able to obtain MOS experiments results for web-browsing. Instead we used the data available in ITU-T G.1030 [14]. The G.1030 model describes the relation between the different response and download times within web-browsing sessions and the corresponding perceived quality, for a given maximum session time, within a certain network and system configuration. The model is applicable to a wide range of configurations, as well as to web-browsing services for a wide variety of users.

Three subjective experiments mimicked real-life web-browsing experience with time-scales of around 6, 15, and 60 seconds, representing fast, moderate and slow network contexts, respectively. During each session, a subject initially requests a search page and finally retrieves the requested web page. The session time-line is shown in Figure 2.

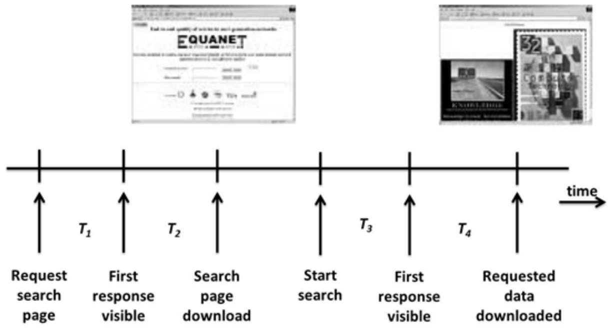

Figure 2. Time-line for the web-browsing application (extracted from [14]).

The first two time intervals, T_1 and T_2, represent the non-interactive response and download times of the search page. The second two time intervals, T_3 and T_4, represent the interactive response and download times of the result page.

A general mapping from session time to web-browsing quality is built by defining minimum, *Min*, and maximum, *Max*, session times whilst using a logarithmic interpolation between these extreme session times. It has been also found that, for shorter duration sessions, the last download time, T_4, has a more severe impact on the final perceived web browsing quality than the other response and download times. T_1, T_2, T_3 and T_4 have to be weighted (as shown in Table II) in order to get a weighted session time that has the highest correlation with the subjectively determined MOS, as follows:

$$\mathrm{WeightedST} = WT_1 \times T_1 + WT_2 \times T_2 + WT_3 \times T_3 + WT_4 \times T_4 \tag{3}$$

The logarithmic interpolation equation is given by:

$$\mathrm{MOS} = \frac{4}{\ln(\frac{Min}{Max})} \times [\ln(WeightedST) - \ln(Min)] + 5 \tag{4}$$

TABLE II. OPTIMAL MODEL WEIGHTING FOR T_1, T_2, T_3 AND T_4 WITH CORRELATIONS BETWEEN OBJECTIVE TIMING AND SUBJECTIVE MOS RESULTS (EXTRACTED FROM [14])

Type of usage	WT_1	WT_2	WT_3	WT_4	*Min*	*Max*	Correlation
6 s expert	0.56	0.84	0.80	1.80			0.97
6 s naïve	0.37	0.40	0.60	2.63			0.93
6 s overall	**0.47**	**0.60**	**0.71**	**2.22**	**0.62**	**13.5**	**0.95**
15 s expert	0.63	0.77	1.11	1.49			0.98
15 s naive	0.48	0.70	0.88	1.95			0.96
15 s overall	**0.54**	**0.72**	**0.98**	**1.76**	**0.81**	**39**	**0.97**
60 s expert	0.84	0.77	1.22	1.18			0.99
60 s naive	0.64	1.01	1.12	1.24			0.98
60 s overall	**0.73**	**0.90**	**1.16**	**1.22**	**2.22**	**151**	**0.98**

The ITU-T G.1030 recommendation further establishes three advanced models by using the best overall weights from Table II in combination with the following mappings from weighted session time to perceived browse quality, in terms of the MOS, clipped between 1 and 5:

$$\mathrm{MOS} = 4.38 - 1.30 \times \ln(WeightedST) \tag{5}$$

for short duration sessions;

$$\text{MOS} = 4.79 - 1.03 \times \ln(WeightedST) \quad (6)$$

for medium duration sessions;

$$\text{MOS} = 5.76 - 0.948 \times \ln(WeightedST) \quad (7)$$

for long duration sessions.

In the context of this research, from a network or operator point of view, it is not feasible to analyse user's web-browsing traffic load to obtain T_1, T_2, T_3 and T_4. Besides, when the session times are shorter, the last download time, T_4, has a higher weight on the perceived quality than the other times T_1, T_2 and T_3. Hence, we are able to conclude that from the user point of view, the main performance factor is how quickly a page appears after it has been requested. Besides, by analysing Figure 2, we can observe that the time required to open the user's desired web page is T_4. Finally, with these considerations in mind, we can also conclude that T_4 is the ratio between the requested amount of *data* (page), in kb, and the available network bitrate, b_{rate}, in kbps:

$$T_4 = {data}/{b_{rate}} \quad (8)$$

We then assume two simplifications, one the one hand we only consider the session time with the highest weight, i.e., time T_4. On the second hand, only medium duration sessions are assumed, i.e., 15 seconds. The MOS is then given by:

$$\text{MOS} = 4.79 - 1.03 \times \ln\left(WT_4 \times {data}/{b_{rate}}\right) \quad (9)$$

By replacing WT_4 by the value extracted from Table II, one obtains:

$$\text{MOS} = 4.79 - 1.03 \times \ln\left(1.76 \times {data}/{b_{rate}}\right) \quad (10)$$

We finally conclude that equation (10) provides a relation between QoS and QoE. By referring the ITU-T G.1030 recommendation we can estimate that the correlation value, R, is approximately 0.9.

D. Audio applications

For audio applications, we considered the well-known ITU-T E-Model from the G.107 recommendation [15]. The E-model assesses the combined effects of varying transmission parameters that affect the conversation quality of narrow band telephony. The E-model is based on the assumptions that transmission impairments can be transformed into psychological factors and psychological factors on the psychological scale are additive.

The result of any calculation with the E-model, in a first step, is a transmission rating factor R_f, which combines all transmission parameters relevant for the considered connection. This rating factor, R_f, is computed by [15]:

$$R_f = R_0 - I_s - I_d - I_e + A \quad (11)$$

R_o represents the basic signal-to-noise ratio, including noise sources such as circuit noise and room noise. The factor I_s is a combination of all impairments which occur more or less simultaneously with the voice signal. I_d represents the impairments caused by delay. The effective equipment impairment factor I_e represents impairments caused by low bitrate codecs, it also includes impairment due to randomly distributed packet losses. The advantage factor A allows for compensation of impairment factors when the user benefits from other types of access to the user. The term R_o as well as the I_s and I_d, are subdivided into further specific impairment values. By considering the default values for these parameters, R_f is given by [15]:

$$R_f = 93.2 - I_d - I_e \quad (12)$$

For $0<R_f<100$ the dependence between R_f and MOS, defined by the ITU, is the following one:

$$\text{MOS} = 1 + 0.035 \times R_f + R_f \times (R_f - 60) \times (100 - R_f) \times 7 \times 10^{-6} \quad (13)$$

The delay impairment factor, I_d, represents all impairments due to delay of voice signals, and includes impairments due to listener echo, talker echo and absolute delay, defined by [15]:

$$I_d = I_{dte} - I_{dle} - I_{dd} \quad (14)$$

The factor I_{dte} gives an estimate for the impairments due to talker echo. The I_{dle} represents impairments due to listener echo and I_{dd} characterizes the impairment caused by too-long absolute one-way mouth-to-ear delay, T_a. The delay associated with listener echo is T_r, the round trip delay in the four-wire loop and the delay related with Talker Echo is T. Assuming IP-based voice applications, the delay can be defined as [16]:

$$d = T_a = T = {T_r}/{2} \quad (15)$$

I_d can be re-written as a function of the one-way delay in ms, d, by means of the curve fitting provided by [16]:

$$I_d = 0.024\text{R}d + 0.114\text{R}d - 177.3) \times H(x) \times (d - 177.3) \quad (16)$$

where $H(x)$ is the Heaviside or step function.

In order to reach a more accurate fit to the curve from G.107, a sixth degree polynomial fit function is provided [17]:

$$I_d = 1.618 \times 10^{-13} \times d^6 - 1.765 \times 10^{-10} \times d^5 + 6.447 \times 10^{-8} \times d^4 - 8.221 \times 10^{-6} \times d^3 + 0.0002315 \times d^2 + 0.0352 \times d - 0.02434 \quad (17)$$

Given a set of measured I_e values for a codec, we can then derive its model by using regression techniques, without the need for subjective tests. For the AMR codec, the MOS (PESQ) vs. packet loss rate can be obtained by averaging over 30 different packet loss locations, in order to remove the influence of these locations.

Additionally, the MOS score for one packet loss rate is obtained by averaging overall speech samples (a total of 16 samples, consisting of 8 males and 8 females), so that the influence of gender is removed [17]. The reference speech database was taken from the ITU-T data set [18]. The relation between the MOS and packet loss rate can be converted to the equipment impairment, I_e, vs. packet loss rate by considering the following third degree polynomial fitting:

$$R_f = 3.026 \times \text{MOS}^3 - 25.314 \times \text{MOS}^2 + 87.06 \times \text{MOS} - 55.336 \quad (18)$$

By considering only the equipment impairment, R_f can be converted to I_e as follows:

$$I_e = R_0 - R_f \quad (19)$$

Finally, a logarithm fitting function, similar to the one from [16], can be derived through curve fitting:

$$I_e = 16.68 \times \ln(1 + 0.3011 \times \rho) + 14.96 \quad (20)$$

where ρ is the packet loss rate (in percentage).

The MOS vs. packet loss and delay may now be obtained by replacing equations (17) and (20) in (12) and computing the MOS by considering equation (13). To simplify the resulting equation we have also computed the equation that best fits the obtained data and achieved a simpler third degree equation:

$$\mathrm{MOS} = 3.657 - 0.1345 \times \rho - 0.003303 \times d - 0.004012 \times \rho^2 - 3.275 \times 10^{-5} \times d^2 - 4.564 \times 10^{-5} \times \rho^3 + 4.414 \times 10^{-8} \times d^3 \quad (21)$$

The goodness of this fit is guaranteed by $R^2 = 0.984169659$ and MSE = 0.000963288.

V. Unified Model

The unified model is obtained by mathematically integrating the fitness obtained for each application, equation (22). Table III presents the variables to be replaced in the final equation, according to the selected application to compute the expected QoE. A_1, A_2, A_3, A_4 and A_5 are Boolean variables used to select the QoS parameters impacting the QoE. The performed regression analysis validates the final model by means of the goodness of fit obtained for all multimedia applications, i.e., R^2 and MSE. The accuracy of the model can still be improved by considering further MOS experimental results, i.e., with other QoS parameters and additional MOS values for web-browsing and audio applications.

TABLE III. Variables to be replaced in the final equation, according to the selected application to compute the expected QoE

	Gaming	Video	Web	Audio
a_0	4.160	3.215	4.790	3.657
A_1	1	0	0	0
A_2	0	1	0	0
A_3	0	0	1	0
A_4	1	1	0	1
A_5	0	1	0	1
a_1	-0.268	-20.989	0	-0.135
a_2	0.119	166.121	0	-4.012×10^{-3}
a_3	-0.014	-730.016	0	-4.564×10^{-5}
a_4	41.295×10^{-5}	1764.470	0	0
a_5	3.900×10^{-6}	-2069.090	0	0
a_6	-1.870×10^{-7}	903.102	0	0
a_7	0	-10.481	0	-0.003303
a_8	0	40.331	0	-3.275×10^{-5}
a_9	0	-42.493	0	4.414×10^{-8}
a_{10}	0	18.388	0	0
a_{11}	0	-3.482	0	0
a_{12}	0	0.237	0	0

This single equation for multiple multimedia application is provided to build an effective QoE control mechanism onto measurable QoS parameters for multimedia networks, i.e., for improving packet scheduling in mesh and or cognitive radio (CR) networks. In particular, when considering the additional delay and possible packet loss induced by spectrum mobility in CR networks, the model facilitates to verify their impact on the achieved user QoE, explicitly for delay-sensitive applications, such as interactive gaming.

VI. Conclusions and Future work

Next generation wireless networks are bound to provide a multitude of multimedia services. The QoE concept is introduced in these networks to describe the satisfaction of subscriber's quality requirements, as it is becoming critical to operators, to guarantee their clients degree of satisfaction in nowadays competitive market. A provider needs to be able to observe and react quickly on quality problems. Otherwise, he risks losing market share to other providers. Such challenge motivates the creation of strategies to assess and map the user perception, experience, and expectations with measurable technical parameters such as application and network-level QoS. This will give operators some sense of the contribution of the network performance to the overall level of user’s satisfaction.

As such, in this work, applications requirements from the ITU [3], 3GPP [4], WiMAX Forum [5] and the EU-MESH European project [6] have been gathered. The goal is to merge all recommendations/documents, to provide and updated services and applications characterization. The trend of wireless technologies has been to provide higher throughputs and QoS. Considering that next generation wireless networks will continue to provide superior data rates, the downlink and uplink speeds were chosen, whenever possible, by considering the highest value from the gathered applications requirements.

Furthermore, the aim of this work is also to propose a pioneering model that maps the relation between the service quality parameters and the user’s QoE. Four main types of multimedia applications have been addressed and four individual fittings are obtained to estimate the QoE as a function of service level parameters. MOS experimental results have been considered for gaming and video applications. These results have been fitted to compute an equation of user’s expectation as a function of QoS parameters. Due to the scarce number of MOS result available in the literature we have considered the ITU-T recommendations for web-browsing and audio applications. These recommendations are based on experimental results. We have validated our fittings by means of regression analysis. We have shown that the values for the coefficient of determination between the model for MOS and the available data (either experimental values or the ones from ITU-T G.107 recommendation) are higher than R^2 = 0.93, 0.98 and 0.84 for gaming, audio and video, respectively. The computed mean square error is approximately zero in all cases.

$$\mathrm{QoE(QoS)} = a_0 + A_1(0.035p - 0.105j - 0.003p^2 + 0.011j^2 + 7.84 \times 10^{-5}p^3 - 48.113 \times 10^{-5}j^3 - 8.486 \times 10^{-7}p^4 + 9.07 \times 10^{-6}j^4 + 3.92 \times 10^{-9}p^5 - 7.513 \times 10^{-8}j^5 - 6.058 \times 10^{-12}p^6 + 2.168 \times 10^{-10}j^6) + A_2\left(-26.6916 \times 10^{-4}b_{rate} - 5.888 \times 10^{-6}{b_{rate}}^2 + 1.449 \times 10^{-8}{b_{rate}}^3 - 4.294 \times 10^{-12}{b_{rate}}^4 - 2.299 \times 10^{-15}{b_{rate}}^5 + 8.087 \times 10^{-19}{b_{rate}}^6\right) + A_3\left(4.79 - 1.03\ \ln\left(1.76 \times \frac{data}{b_{\mathrm{rate}}}\right)\right) + A_4(a_1\rho + a_2\rho^2 + a_3\rho^3 + a_4\rho^4 + a_5\rho^5 + a_6\rho^6) + A_5(a_7 d + a_8 d^2 + a_9 d^3 + a_{10} d^4 + a_{11} d^5 + a_{12} d^6) \quad (22)$$

ACKNOWLEDGMENT

This work has been partially supported and funded by the FCT project PEst-OE/EEI/LA0008/2013, Ubiquimesh, OPPORTUNISTIC-CR, LTE-Advanced Enhancements using Femtocells, CREaTION, HANDCAD, COST IC0905 “TERRA”, COST IC 0902, COST IC 1004, and by the Marie Curie Reintegration Grant PLANOPTI (FP7-PEOPLE -2009-RG). The authors also acknowledge Robert E. Kooij and Jânio M. Monteiro for granting access to their MOS experimental results.

REFERENCES

[1] J. A. Hassan, S. K. Das, M. Hassan, C. Bisdikian, D. Soldani, "Improving quality of experience for network services [Guest Editorial]," *IEEE Network,* Vol. 24, No. 2, pp. 4-6, Mar.-Apr. 2010 (available online DOI 10.1109/MNET.2010.5430137).

[2] H. Koumaras, F. Liberal, L. Sun, “Quality of experience issues in multimedia provision,” *Telecommunication Systems*, Vol. 49, No 1, pp. 1-3, 2012 (available online DOI 10.1007/s11235-010-9349-4).

[3] ITU-T, *End-user multimedia QoS categories*, Recommendation of the ITU-T, Recommendation G.1010, International Telecommunication Union, Geneva, Switzerland, 2001.

[4] 3GPP, *TS 22.105 (2008-12) Services and service capabilities (Release 9)*, 3GPP Technical Specification Group Services and System Aspects, Services and service capabilities, V9.0.0, 2008.

[5] WiMAX Forum: WiMAX System Evaluation Methodology. Version 2.1, July 2008.

[6] EU-MESH: Enhanced, Ubiquitous, and Dependable Broadband Access using MESH Networks, Public Deliverables, “D2.1 Usage Scenarios and Application Requirements”, April 2008.

[7] P. Brooks, B. Hestnes, "User measures of quality of experience: why being objective and quantitative is important," *IEEE Network*, Vol.24, No.2, pp.8-13, Mar.-Apr. 2010 (available online DOI 10.1109/MNET.2010.5430138).

[8] ITU-T, *Framework Recommendation for multimedia services*, Recommendation of the ITU-T, Recommendation F.700, International Telecommunication Union, Geneva, Switzerland, 2000.

[9] A.F. Wattimena, R.E. Kooij, J.M. van Vugt, O.K. Ahmed, “Predicting the perceived quality of a First Person Shooter: the Quake IV G-model”, in *Proc. of NetGames 2006*, Singapore, 30-31 Oct. 2006, (available online DOI 10.1145/1230040.1230052).

[10] ITU-T, *Methods for Subjective Determination of Transmission Quality*, Recommendation of the ITU-T, Recommendation P.800, International Telecommunication Union, Geneva, Switzerland, 1996.

[11] DSL Forum, Technical Report TR-126: Triple-play Services Quality of Experience (QoE) Requirements, Dec. 2006.

[12] G. Gardikis, G. Xilouris, E. Pallis, A. Kourtis, “Joint assessment of Network and Perceived-QoS in video delivery networks,” *Telecommunication Systems*, Vol. 49, No 1, pp 75-84, Jan., 2012 (available online DOI 10.1007/s11235-010-9349-4).

[13] M. Venkataraman, M. Chatterjee, "Inferring video QoE in real time," *IEEE Network*, Vol. 25, No. 1, pp.4-13, Jan.-Feb. 2011.

[14] ITU-T, *Estimating end-to-end performance in IP networks for data applications*, Recommendation of the ITU-T, Recommendation G.1030, International Telecommunication Union, Geneva, Switzerland, 2005.

[15] ITU-T, *The E-model: a computational model for use in transmission planning*, Recommendation of the ITU-T, Recommendation G.107, International Telecommunication Union, Geneva, Switzerland, 2009.

[16] R. G. Cole and J. Rosenbluth, “Voice over IP Performance Monitoring,” *SIGCOMM Computer Communications Review*, Vol. 31, No. 2, pp. 9-24,Apr. 2001 (available online DOI 10.1145/505666).

[17] Lingfen Sun, *Speech Quality Prediction for Voice Over Internet Protocol Networks*, PhD thesis, University of Plymouth, Plymouth, United Kingdom, Jan. 2004.

[18] ITU-T, *Objective measuring apparatus, Appendix 1: Test signals*, ITU-T Recommendation P.50, International Telecommunication Union, Geneva, Switzerland, 1998.