

# A Survey on Parametric QoE Estimation for Popular Services

Dimitris Tsolkas, Eirini Liotou\*, Nikos Passas and Lazaros Merakos

*National and Kapodistrian University of Athens*

*Department of Informatics & Telecommunications*

*16122, Athens, Greece*

Tel: +302107275158

Fax: +302107275601

E-mail: {dtsolkas, eliotou, passas, merakos}@di.uoa.gr

**Abstract** – As we are moving forward to the 5G era, we are witnessing a transformation in the way networks are designed and behave, with the end-user placed at the epicenter of any decision. One of the most promising contributors towards this direction is the shift from Quality of Service (QoS) to Quality of Experience (QoE) service provisioning paradigms. QoE, i.e., the degree of delight or annoyance of a service as this is perceived by the end-user, paves the way for flexible service management and personalized quality monitoring. This is enabled by exploiting parametric QoE assessment models, namely specific formula-based QoE estimation methods. In this paper, recognizing a gap in the literature between the lack of a proper manual regarding the objective QoE estimation and the ever increasing interest from network stakeholders for QoE intelligence, we provide a comprehensive guide to standardized and state-of-the-art quality assessment models. More specifically, we identify and describe parametric QoE formulas for the most popular service types (i.e., VoIP, online video, video streaming, web browsing, Skype, IPTV and file download services), indicating the key performance indicators (KPIs) and major configuration parameters (MCPs) per type. Throughout the paper, it is revealed that KPIs and MCPs are highly variant per service type, and that, even for the same service, different factors contribute with a different weight on the perceived QoE. This finding can strongly enable a more meaningful resource provisioning across different applications compared to QoE-agnostic schemes. Overall, this paper is a stand-alone, self-contained repository of QoE assessment models for the most common applications, becoming a handy tutorial to parties interested in delving more into QoE network management topics.

**Keywords** – *Quality of experience, objective quality estimation, parametric models, key performance indicators, E-model.*

## 1. Introduction

Quality of Service (QoS) is currently considered inadequate for the thorough characterization of a provided service. The reason is that QoS handles purely technical aspects regarding a service, so it does not incorporate any kind of human-related quality-affecting factors nor does it reflect the actual user experience. This means that the same QoS level might not guarantee the same perceived quality level for two different users. In addition, apart from the system's technical characteristics, other factors, such as the context of use (e.g., surrounding environment, type of device used, etc.), human-specific characteristics (e.g., demographics or current state), the delivered content and the pricing of a service, have a significant impact on the finally perceived user's experience, aspects not captured by QoS. Furthermore, the effect of QoS technical factors on the user's quality perception is, in general, governed by a non-linear relation. Hence, towards the next generation of mobile communications, new terms, such as the Quality of Experience (QoE), have been coined to better depict the end-users' perception of a provided service.

QoE represents the overall quality of a provided service, as it is perceived by the end-users, and as such it is a very appealing alternative for evaluating the quality of a provided service. Similarly to QoS, QoE may be incorporated in network mechanisms and specifically in network decision

\* Corresponding author

processes. “QoE-driven” or “QoE-aware” algorithms can help the network function in a more efficient and effective way. For instance, QoE may be both the criterion and trigger factor of Mobility Management (e.g., [1]) and Radio Resource Management (e.g., [2]) mechanisms, replacing presently used criteria, such as Signal-to-Interference plus Noise Ratio (SINR) measurements. QoE has been even identified as a driver in Software-Defined Networking (SDN) architectures, e.g., in [3], where by monitoring video QoE metrics at the client side, and by dynamically selecting delivery nodes and routing paths, the QoE of video streaming applications is improved.

QoE-awareness may also drive a more resource-efficient network operation, by helping recognize when the provision of extra resources would not improve the QoE that is eventually perceived by the end-user (e.g., [4]). In other words, “over-engineering” could be avoided through QoE monitoring and control. Additionally, QoE awareness may be exploited in many other aspects by network designers and network operators. For instance, network protocol designers may evaluate their implementations based on QoE metrics. Moreover, network problems such as bottlenecks or local failures, may be identified by insufficient QoE values, leading to QoE-improving actions (proactively or reactively). QoE can also drive virtual machine placement, which is an important challenge for modern data centers, aiming to maximize the resource usage but also the final impact on the user experience. Therefore, optimization-based algorithms for virtual machines placement like the promising approach of [5] may be enhanced with QoE objectives, by slightly modifying the already proposed utility functions. However, without the existence of proper estimation methods, such QoE-based solutions are impossible, driving stakeholders to still rely on QoS metrics, despite the recent boost in QoE literature.

Recognizing the importance of QoE awareness, and motivated by the lack of a proper “repository” for QoE estimation methods, we study some approaches that can be used to efficiently acquire the QoE level of a provided service. We focus on *parametric* QoE models, a specific category of QoE estimators, which are appealing for live estimation of QoE based on monitored network performance indicators, namely during actual network operation. More specifically, we provide a thorough description of the dominant QoE estimation models in the literature for the most important service types and applications. Our main intention is to provide a detailed reference map towards efficiently and objectively acquiring the users’ perception of a provided service. To this end, various numerical results for different parametric QoE estimation models are presented on a per service type basis. Therefore, the contribution of this paper is three-fold: a) the identification and description of QoE models and metrics that can be used for real-time QoE monitoring and QoE management inside an operating network, b) the quantification of the contribution and impact of key quality influence factors on the achieved QoE per service type, and c) the proposal of potential research and exploitation directions based on this collected knowledge.

The structure of this paper starts with a taxonomy of QoE estimation methods, dividing them into two main categories: namely subjective and objective. Emphasis is given on objective QoE estimation where various evaluation methods are discussed. The parametric category of objective QoE estimation is identified as the most appealing one for online network control, and therefore, the paper then focuses on this specific category only. Our study is done per-service, since QoE is highly service dependent. First, standardized QoE methods are described (referring to VoIP and real-time video services), followed by the non-standardized literature-based methods (referring to the popular services of FTP, IPTV, Skype and video streaming). Finally, a concise summary of this study is provided together with a discussion on possible research and exploitation directions, before the final conclusions.

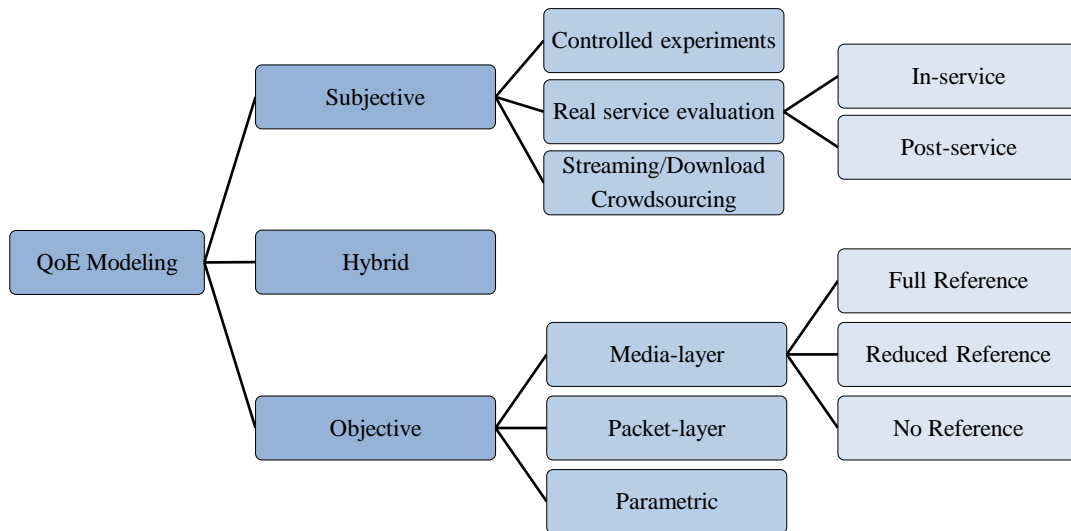
Based on the previous, the remainder of this paper is organized as follows. In section 2, we classify the QoE estimation approaches, highlighting the importance of parametric QoE estimation. In section 3, we present and evaluate standardized QoE estimation models, while, in section 4, we describe literature-based dominant parametric QoE estimation models for different types of services. In section 5, we summarize guidelines and study directions towards designing and efficiently exploiting parametric QoE estimation models. Our conclusions are included in section 6.

## 2. QoE estimation taxonomy

Current literature on the topic of QoE modeling mainly offers classifications of existing standards and focuses on one specific service at a time. For instance, [6] studies speech quality estimation and provides a detailed taxonomy of standardized objective speech quality prediction models. Similarly, [7] conducts a thorough survey of QoE assessment approaches for VoIP services, while [8] focuses on IPTV. However, these considered service types are just a subset of the plethora of services available in current networks. With the availability of 4G and with 5G on the horizon, which allows the co-existence of multiple parallel resource-hungry requests, applications like video streaming and video on demand constitute the prevalent traffic over a network.

What is more, survey papers on QoE estimation usually focus on models that require the originally transmitted signal or part of it to deduce the QoE at the receiver side, not targeting in this way at real-time network management application (e.g., [9]). On the contrary, parametric QoE models are appropriate for this type of scenarios; however, a handy collection of these models for different types of services is currently missing from the literature.

There are various different approaches for quantifying the QoE level of a provided service. A primary classification of the available approaches is based on whether the QoE is evaluated directly by humans or automatically through technical factors. In the first case, specific assessment processes are used, referred to as *subjective tests*, while in the second case mathematical formulas or algorithms are exploited, referred to as *objective models*. The main classification of QoE models is presented graphically in Figure 1, and is further discussed in the next subsections.



**Figure 1:** Classification of QoE modeling approaches

### 2.1 Subjective QoE estimation

Subjective tests are usually based on controlled real life experiments with human participants who directly evaluate their experience of an application or service. These users may be involved in the experiment in a passive way (just viewing/listening) or in an active/interactive way (participating in a conversation) and they judge the quality regarding some stimulus' presentation. For instance, the participants may be called to evaluate the listening or conversational quality of a phone service, the quality of a video, etc. These tests need to be thoroughly designed in advance and the user group needs to be properly selected based on guidelines and recommendations by standardization bodies. Perhaps the most important recommendation towards that direction is the ITU-T P.800 [10]. Various techniques may be used for subjective evaluation. For instance, users may score the quality using an absolute rating scale or they may compare sequential images/videos/sounds stating which one is better. The results are based on user opinions, past experiences, expectations, user perception, judgement and description capabilities, etc. and primarily quantify the effectiveness, efficiency and overall satisfaction of using a service.

These kinds of subjective tests are considered as the most reliable ones, since they incorporate any conscious and unconscious aspects of human quality evaluation, aspects that can otherwise not be captured. Only perceptual quality tests can validly and reliably express the internal state of the human factor. Nevertheless, such subjective techniques are considered reliable, if and only if they are designed carefully and users are unbiased and objective.

One drawback of the above method is that the results of such experiments are valuable only for the laboratory testing of some service, and not for real-time QoE support. One way to overcome this issue is to conduct “real-service” QoE evaluation, where users rate their experience on the run (in-service) or after a service has ended (post-service). Such an example is the “OneClick” paradigm, which may be used for real-time QoE monitoring and feedback, and consequently for QoE control. This framework only requires a subject to click a dedicated key whenever he/she feels dissatisfied with the quality of the application in use [11]. Furthermore, an example of post-service test is that of Skype, where users rate their experience once a session is terminated, using the Mean Opinion Score (MOS) scale.

Subjective experiments in controlled laboratory environments need thorough design that strictly follows guidelines provided by standardization bodies. These guidelines describe all aspects such as room conditions (e.g., isolated room, without any noise), audio headset or generally the dedicated equipment used for hearing/viewing/talking, test methodologies, guidelines for the selection of the panel, etc. Regarding the latter, there are guidelines regarding the number of participants, their age, their background (experts or non-experts), their past involvement in similar experiments, the randomness in their selection, etc. However, lately there is also a trend to evaluate the quality of an application in a more relaxed way, i.e., at one’s own and familiar environment, using one’s own equipment and so forth. In this kind of experiments, a service is evaluated using “streaming” or “download” approaches. These methods are considered as more realistic and are open to a much broader public comparing to laboratory experiments, thus allowing for better management. Indeed, a large number of participants may reveal very reliable and realistic QoE scores. Approaches that follow this paradigm are called “Crowdsourcing” techniques [12], because they outsource the task of quality evaluation to arbitrary anonymous online users. One such example is the Google Microworkers platform as well as the Amazon Mechanical Trunk, where an Internet user may conduct QoE experiments designed by other parties (such as researchers), who require a general public for an evaluation task.

Finally, an important issue in subjective test methodologies is the discrimination between “instantaneous” and “overall” quality evaluations. The former method implies a continuous evaluation of the perceived quality by the end-user during one experiment (see ITU-T P.880), whereas the latter simply requires that the user gives one cumulative score for his/her own experience at the end of each experiment. The first method gives a better insight to the system designers, since they can correlate the instantaneous quality with momentary technical parameters in the network; however the latter better describes the overall user experience.

## 2.2 Objective QoE estimation

Subjective tests are costly, time-consuming and not reproducible on demand. Moreover, they are usually not real-time and hence cannot be used for in-service quality monitoring. These constraints have raised the need for the development of objective models that try to measure or predict the quality perceived by end-users, without their intervention. The objective models may be classified using various criteria [13],[14]:

- **Reference signal utilization:** Regarding whether the source signal or part of it is required or not in the QoE estimation process, we distinguish the *Full Reference (FR)* or *reference-based* or *double-ended* models, the *Reduced Reference (RR)* models and the *No Reference (NR)* or *single-ended* models, where “reference” refers to the original signal. [9] conducts a survey on the evolution of video quality assessment methods using this classification.
- **Evaluation method:** Regarding the kind of input information that is used for QoE measurement, we distinguish the: *Media-layer (signal-based)*, *Packet-layer / Bitstream*, and *Parametric* models.

Media-layer models make use of signals and may be Full-, Reduced or No-Reference; Packet-layer models extract information from packet headers, while bitstream models may use both packet headers and payload data. Parametric (or parametric planning) models use network planning parameters and measure the values of specific network metrics. Packet-layer and bitstream models are also referred in literature as “protocol-information-based” models, because they base their estimations on parameters collected at run time from network processes and control protocols. Various surveys in the literature review media-layer models (e.g., [15] thoroughly discusses media-layer models for video quality assessment), while others focus on packet-layer/bitstream models (e.g., [16]). Finally, [17] conducts a study of the correlation models mapping QoS to QoE for multimedia services, providing in this way generic formulas that parametric models usually follow.

- **Model mode:** Regarding the injection of a test signal to the system under test or not, we have: *Intrusive (active)* and *Non-intrusive (passive)* modes respectively.
- **Model timeframe:** *Offline* models refer to pre-service or post-service evaluation, while *online* models require in-service evaluation.
- **Usage purpose:** The usage purpose of a QoE estimation model may vary from network planning, to real-time service monitoring, optimization, benchmarking, etc. The QoE models should be carefully used only within their scope.

Parametric QoE models are basically derived by conducting subjective experiments (lab or crowdsourcing) and then by performing statistical analysis (e.g., regression analysis) on the acquired evaluation results. The derived objective models may be then well-described by providing formulas for the direct computation of QoE based on specific input parameters. On the contrary, signal-based models are based on one-to-one comparison between the original source signal and the degraded destination signal, by exploiting knowledge from the area of Psychophysics.

Also worth mentioning is a third category of QoE modeling, which lies between subjective and objective ones. It operates in a hybrid fashion, namely it works as an automatic and objective quality estimator, relying however on prior available subjective scores. These hybrid methods are based on Machine Learning tools, and they are using subjective test scores as input to train a QoE model. This model then maps network parameters (e.g., codec used, packet loss rate, mean loss burst size, packetization interval, one-way delay, jitter, etc.) to Mean Opinion Score (MOS) values and it can be further used for real-time quality prediction. Characteristic examples of this approach are the Pseudo-Subjective Quality Assessment (PSQA) method [18] and the MLQoE, a modular algorithm for user-centric QoE prediction [19].

What is more, some research works propose methodologies for the construction of objective QoE models. For instance, [20] describes basic principles for building a QoE model, which is based on the egress of parameterized mappings among three layers: the transport-layer, the service-layer and the end-user layer (bottom up). Similarly, in [21], the authors build a QoE estimation function based on a general regression model and prove its applicability to web-browsing and file upload/download scenarios.

At the moment, most objective models account for the human factor in terms of their inherent characteristics, but the context and content of the service are considered only at a limited extent. Under this observation, more work is needed for designing more accurate objective estimation models. More specifically, extra forces should be allocated toward the designing of new parametric QoE estimation models, since they are currently the most appealing candidates for quantifying QoE levels in an indirect and user-transparent way in mobile networks [22]. Taking this into account, the dominant parametric QoE estimation models are studied in this paper.

Having described the taxonomy of QoE methods, we now focus on objective QoE estimation and specifically on parametric QoE evaluation methods. Our goal is to study parametric QoE estimation approaches for popular services, since this category is particularly useful for real-time network management and control. We, therefore, discuss the key models for file download (FTP), web browsing, voice over IP (VoIP), online/real-time video (e.g., IPTV or video-calls), video streaming

services (e.g., YouTube), and Skype, explaining in detail the factors and the parameters that each model requires. Our structure is as follows: we first study standardized QoE approaches and then literature-based ones. The former study handles VoIP and online/real-time video QoE estimation, while the latter refers to the remaining popular services.

### 3. Standardized parametric QoE estimation

In this section, we present the two standardized parametric models that are available in the literature, namely G.107 for VoIP and G.1070 for online video. Our main goal is to study the impact of various factors on the user's QoE; however, for completion, we also present the most substantial parts of these standards, the full versions of which may found in the ITU-T portal. In this way, this document can comprise a stand-alone guide of QoE estimation models.

#### 3.1 Parametric QoE estimation for VoIP services

In VoIP applications, similarly to conventional telephony, the QoE is expressed in terms of how clearly the user can listen and understand his or her interlocutor's speech, and how easy or not the communication is due to potential arrival delays of speech internet packets. Because of this, the models for this service are divided into the following categories: listening-only and conversational. Subjective assessment methods, in VoIP services, are based on four testing axes [23]: Comprehensibility, Multi-dimensional, Listening Quality and Conversational Quality tests. The MOS is the most extensively used measurement scale for observations of this kind. Concerning parametric objective methods, the ITU-T Rec. G.107, a.k.a. the "E-model" [24],[25] is the most reliable and representative approach.

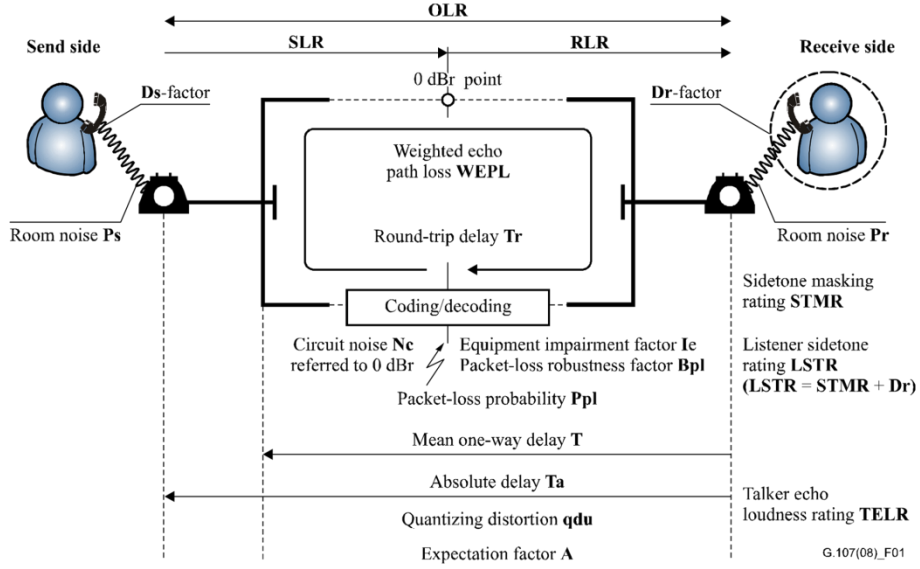
##### 3.1.1 The basic rating factor

The E-model provides a formula that can be used for the computation of the transmission quality of voice communications by estimating the mouth-to-ear conversational quality as perceived by the user at the receive side, both as listener and talker (Figure 2). It is a parametric model that takes into account a variety of transmission impairments producing the so-called Transmission Rating factor ( $R$  factor). The conversational quality is estimated by means of this rating factor  $R$ , scaling from 0 (worst) to 100 (best):

$$R = R_0 - I_s - I_d - I_{e-eff} + A$$

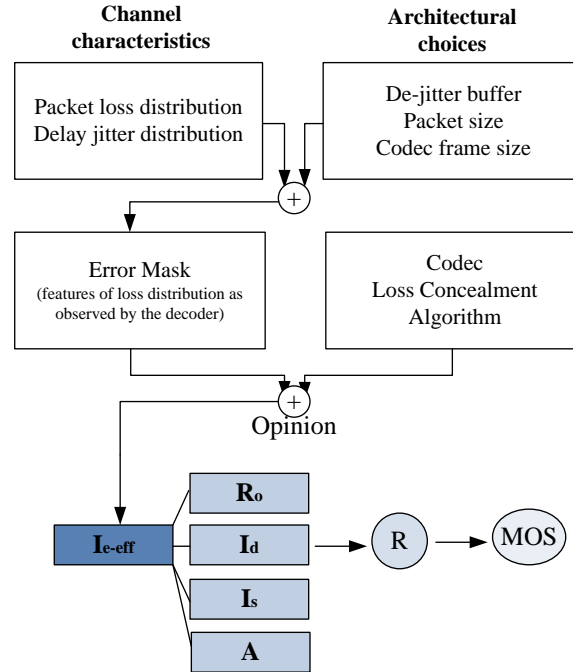
where:

- $R_0$  represents in principle the basic signal-to-noise ratio, including noise sources such as circuit noise and room noise.
- $I_s$  is a combination of all impairments which occur more or less simultaneously with the voice signal.
- $I_d$  represents the impairments caused by delay.
- $I_{e-eff}$  represents impairments caused by low bit-rate codecs (effective equipment impairment factor). It also includes impairments 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.



**Figure 2:** Reference connection for the E-model [24]

Worth mentioning at this point is the methodology proposed in [26] and presented in Figure 3. According to this, the E-model is reduced to transport-level parameters only, which can be easily measured within the network. The main idea is to combine transport-level measurements such as delay and packet loss with architectural-specific parameters such as the de-jitter buffer at the receiver side to get an estimation of the effective equipment impairment factor  $I_{e-eff}$  presented above. This methodology can be therefore directly used for VoIP quality measurement and monitoring.



**Figure 3:** Methodology for VoIP quality measurement (based on [26])

The detailed calculation of each one of the factors used in G.107 is provided in the following subsections, while summarized default values can be found in Appendix A.

### 3.1.1.1 Basic signal-to-noise ratio, $R_o$

$R_o$  represents in principle the basic signal-to-noise ratio, including noise sources such as circuit noise and room noise and it is calculated as follows:

$$R_o = 15 - 1.5(SLR + N_o)$$

where  $SLR$  (Send Loudness Rating) represents the loudness rate at the sender and  $N_o$  the power addition of different noise sources. More specifically,  $N_o$  is as follows:

$$N_o = 10 \log \left[ 10^{\frac{N_c}{10}} + 10^{\frac{N_{os}}{10}} + 10^{\frac{N_{or}}{10}} + 10^{\frac{N_{fo}}{10}} \right]$$

where  $N_c$  is the sum of all circuit noise powers, all referred to the 0 dBr point, and  $N_{fo}$  is the “noise floor” at the receive side.  $N_{os}$  is the equivalent circuit noise at the 0 dBr point, caused by the room noise  $P_s$  at the send side:

$$N_{os} = P_s - SLR - D_s - 100 + 0.004(P_s - OLR - D_s - 14)^2$$

where,  $D_s$  is the D-Value of telephone at the sender side (normally  $D_s = 3$ ) and the overall loudness rating ( $OLR$ ) is the sum of the loudness rate at the sender and the receiver (denoted by  $SLR$  and  $RLR$ , respectively), i.e.:

$$OLR = SLR + RLR$$

Similarly,  $N_{or}$  is the equivalent circuit noise at the 0 dBr point for the room noise  $P_r$  at the receive side:

$$N_{or} = RLR - 121 + P_{re} + 0.008(P_{re} - 35)^2$$

where:

$$P_{re} = P_r + 10 \log \left[ 1 + 10^{\frac{10-LSTR}{10}} \right]$$

and:

$$N_{fo} = N_{for} + RLR$$

with  $N_{for} = -64 \text{ dBmp}$

### 3.1.1.2 Simultaneous impairment factor, $I_s$

This factor includes all the potential degradations that can arise together with the voice transmission. Three main impairments are considered:

$$I_s = I_{olr} + I_{st} + I_q$$

where  $I_{olr}$  represents the decrease in quality caused by too-low values of OLR,  $I_{st}$  represents the impairment caused by non-optimum sidetone, and  $I_q$  represents the impairment caused by quantizing distortion. More specifically the formula used for the calculation of  $I_{olr}$  is:

$$I_{olr} = 20 \left[ \left\{ 1 + \left( \frac{X_{olr}}{8} \right)^8 \right\}^{\frac{1}{8}} - \frac{X_{olr}}{8} \right]$$

where:

$$X_{olr} = OLR + 0.2(64 + N_o - RLR)$$

For the  $I_{st}$  factor the formula is the following:

$$I_{st} = 12 \left[ 1 + \left( \frac{STMR_o - 13}{6} \right)^8 \right]^{\frac{1}{8}} - 28 \left[ 1 + \left( \frac{STMR_o - 1}{19.4} \right)^{35} \right]^{\frac{1}{35}} - 13 \left[ 1 + \left( \frac{STMR_o - 3}{33} \right)^{13} \right]^{\frac{1}{13}} + 29$$

where:

$$SMTR_o = -10 \log \left[ 10^{\frac{SMTR}{10}} + e^{-\frac{T}{4}} \cdot 10^{\frac{TELR}{10}} \right]$$



and finally, the  $I_q$  factor is calculated as follows:

$$I_q = 15 \log[1 + 10^Y + 10^Z]$$

where:

$$Y = \frac{R_o - 100}{15} + \frac{46}{8.4} - \frac{G}{9}$$

$$Z = \frac{46}{30} - \frac{G}{40}$$

$$G = 1.07 + 0.258Q + 0.0602Q^2$$

$$Q = 37 - 15\log(qdu)$$

and  $qdu$  refers to the number of quantizing distortion units from the sender to the receiver.

### 3.1.1.3 Delay impairment factor, $I_d$

The impairment factor  $I_d$  represents all impairments due to the delay of voice signals and it is further divided into three factors as follows:

$$I_d = I_{dte} + I_{dle} + I_{dd}$$

More specifically, the factor  $I_{dte}$  represents the impairment due to echo at the talker:

$$I_{dte} = \left[ \frac{R_{oe} - R_e}{2} + \sqrt{\frac{(R_{oe} - R_e)^2}{4} + 100} - 1 \right] (1 - e^{-T})$$

$$R_{oe} = -1.5(N_o - RLR)$$

$$R_e = 80 + 2.5(TERV - 14)$$

$$TERV = TELR - 40 \log \frac{1 + \frac{T}{10}}{1 + \frac{T}{150}} + 6e^{-0.3T^2}$$

Note that when  $SMTR < 9dB$ , the term  $TERV$  is replaced by the term  $TERVs$ :

$$TERVs = TERV + \frac{I_{st}}{2}$$

while when  $SMTR > 20dB$ , the term  $I_{dte}$  is replaced by the term:

$$I_{dtes} = \sqrt{I_{dte}^2 + I_{st}^2}$$

The factor  $I_{dle}$  represents impairments due to listener echo calculated by:

$$I_{dle} = \frac{R_o - R_{le}}{2} + \sqrt{\frac{(R_o - R_{le})^2}{4} + 169}$$

where:

$$R_{le} = 10.5(WEPL + 7)(T_r + 1)^{-0.25}$$

Finally, the factor  $I_{dd}$  represents the impairment caused by too-long absolute delay  $T_a$ , which occurs even with perfect echo cancelling:

$$I_{dd} = \begin{cases} 0, & T_a \leq mT \\ 25 \left\{ (1 + X^{6 \cdot sT})^{\frac{1}{6 \cdot sT}} - 3 \left( 1 + \left[ \frac{X}{3} \right]^{6 \cdot sT} \right)^{\frac{1}{6 \cdot sT}} + 2 \right\}, & T_a > mT \end{cases}$$

where:

$$X = \frac{\log\left(\frac{T_a}{mT}\right)}{\log 2}$$

The  $mT$  and  $sT$  reflect different use cases and user groups. Two aspects are addressed by these settings:

- The interactivity of the conversation and the sensitivity of the users to the delay-effect
- The application scenario, that is, whether a given call is being made in a business context or in an everyday situation. Even if users may not notice the delay, it may be very critical for the efficiency or even effectiveness of a given call, for example in a business context.

**Table 1:** Values of  $mT$  and  $sT$  for different use cases

Delay class	$sT$	$mT$	Use case
Default	1	100	Applicable to all types of telephone conversations.
Low	0.55	120	Must be used for: - Carrier-grade fixed or mobile telephony - Enterprise-grade fixed or mobile telephony - When targeted user group and delay requirements are unknown.
Very low	0.4	150	Applicable only in cases where it is known that users have very low sensitivity to delay, e.g., in primarily non-interactive cases, such as to mainly listen to a conversation or a lecture.

As a consequence, in case that it is uncertain what user group or what application scenario is being addressed with the planned service, it is recommended that the default class is used. Any case of non-default value usage should be explicitly mentioned when reporting results. Based on these considerations, Table 1 depicts the settings that are recommended for parameters  $mT$  and  $sT$  for different use cases.

#### 3.1.1.4 Equipment impairment factor, $I_{e-eff}$

The  $I_{e-eff}$  factor is based on metrics close related to packet losses. More specifically:

$$I_{e-eff} = I_e + (95 - I_e) \cdot \frac{P_{pl}}{\frac{P_{pl}}{BurstR} + B_{pl}}$$

where:

- $I_e$  is the codec related factor and its potential values have been defined based on subjective results. Recommended values can be seen in Table 2.
- $B_{pl}$  is the packet-loss robustness factor
- $P_{pl}$  is the packet-loss probability, and
- $BurstR$  is the burst ratio, which is defined as:

$$BurstR = \frac{\text{Average length of observed bursts in an arrival sequence}}{\text{Average length of bursts expected for the network under "random" loss}}$$

When packet loss is random (i.e., independent)  $BurstR = 1$ , and when packet loss is bursty (i.e., dependent)  $BurstR > 1$ .

For example, for packet loss distributions corresponding to a 2-state Markov model with transition probabilities  $p$  between a “found” and a “loss” state, and  $q$  between the “loss” and the “found” state, the burst ratio can be calculated as:

$$BurstR = \frac{1}{p + q} = \frac{P_{pl}/100}{p} = \frac{1 - P_{pl}/100}{q}$$

**Table 2:** Recommended values for  $I_e$  factor when Adaptive Multi-Rate (AMR) mode is used

AMR mode index	Application layer Encoding Rate (Kbps)	$I_e$ value
1	4,75	29
2	5,15	27
3	5,9	23
4	6,7	20
5	7.4	16
6	7.95	15
7	10.2	9
8	12.2	5

### 3.1.1.5 Advantage factor, $A$

The use of factor  $A$  and its selected value in a specific application is up to the planner's decision. However, the values in Table 3 should be considered as absolute upper limits for factor  $A$ .

**Table 3:** Provisional examples for the advantage factor  $A$

Communication system	$A$ (Max value)
Conventional wired communication	0
Mobility by cellular networks in a building	5
Mobility in a geographical area or moving in a vehicle	10
Access to hard-to-reach locations, e.g., via multi-hop satellite connections	20

### 3.1.2 Quality metrics derived by the basic rating factor

The transmission rating factor  $R$  ranges from 0 to 100, where  $R = 0$  represents an extremely bad quality and  $R = 100$  represents a very high quality. Even if the  $R$  factor can be directly used as an assessment value, in most of the times it is preferable to transform it to MOS values to retrieve results comparable with results provided by subjective methods. The transformation formula is as follows:

$$MOS_{VoIP} = \begin{cases} 1, & \text{if } R < 0, \\ 1 + 0.035R + R(R - 60)(100 - R) \cdot 7 \cdot 10^{-6}, & \text{if } 0 \leq R \leq 100, \\ 4.5, & \text{if } R > 100 \end{cases}$$

Other useful quality values that can be derived by the rating factor  $R$  are the percentages for a judgement Good or Better (GoB) and Poor or Worse (PoW). There values are derived with the aid of the Gaussian function:

$$E(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$$

For the GoB value it holds that:

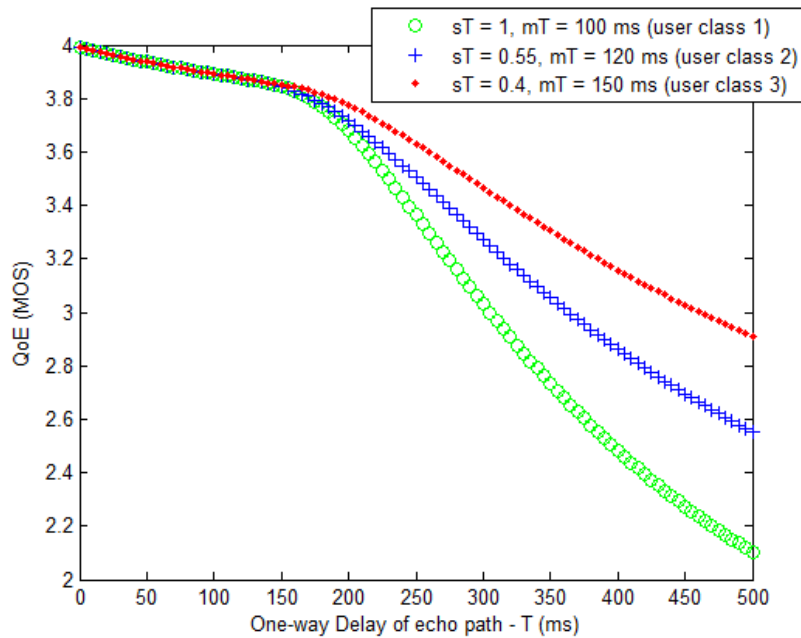
$$GoB = 100 \cdot E \left( \frac{R - 60}{16} \right) \%$$

while for the PoW value it holds that:

$$PoW = 100 \cdot E \left( \frac{45 - R}{16} \right) \%$$

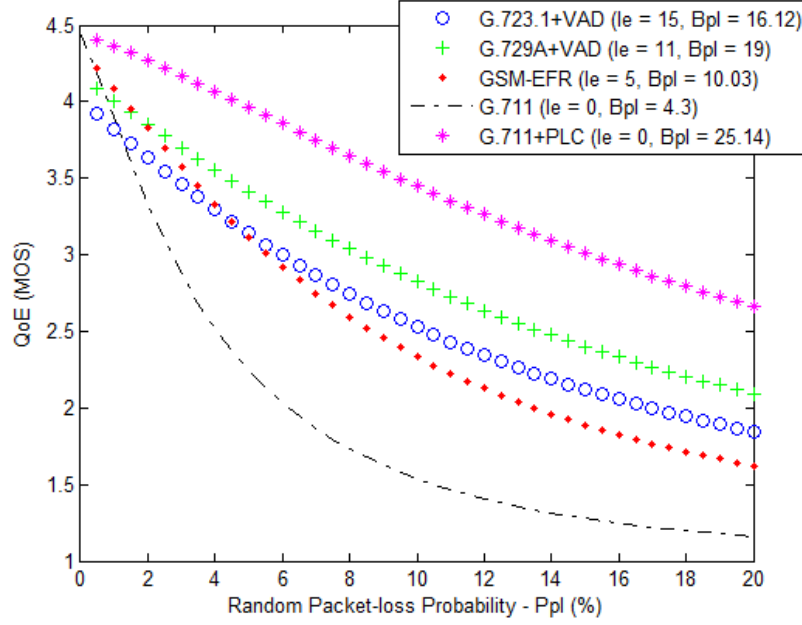
### 3.1.3 Numerical results

Below we quantify the impact of key parameters referred in the previous section on the MOS domain for VoIP services. First of all, Figure 4 depicts the MOS level for an increasing value of the echo path delay. As expected, an increase at the echo path delay leads to QoE downgrade. However, the rate of reduction varies depending on the type of user and the delay value. When the user's profile refers to high delay sensitivity (e.g., user class 1), then an increase of  $T$  causes a steeper decline in QoE, compared to a delay-tolerant user (e.g., user class 3). In Figure 4, we can also observe that for  $T < 150ms$ , all three user categories perceive the same, almost constant quality, whatever their requirements. This suggests that the value  $T = 150ms$  is the maximum acceptable delay for VoIP services, namely the delay value by which no rapid change in perceived QoE of speech is observed. Once the delays become greater than this threshold, an exponential degradation is observed, the slope of which depends on the parameter  $sT$ , i.e., the term that expresses the user's sensitivity to delays.

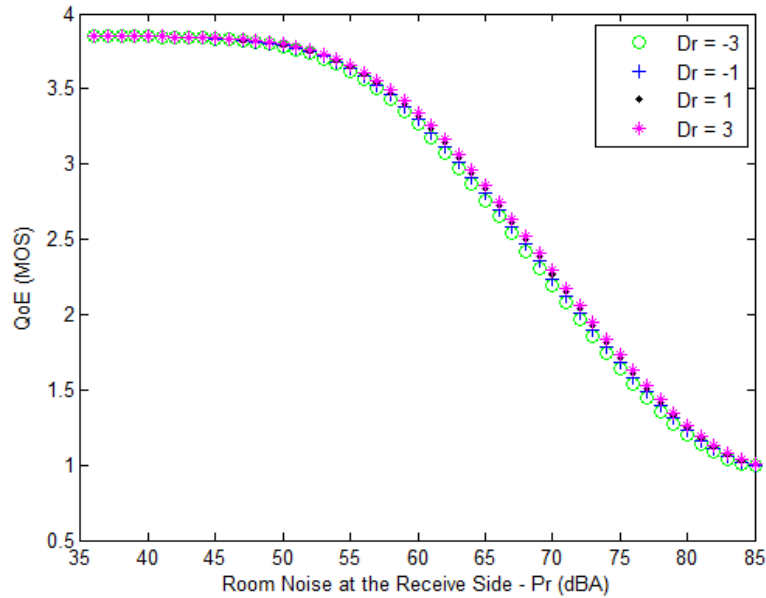


**Figure 4:** MOS values versus delay for three different user classes

Figure 5 studies the relationship between QoE and packet loss probability. It is shown that higher MOS can be offered by a coding system with smaller equipment downgrades and greater robustness (resistance) to packet loss. Specifically, the relationship between QoE and packet loss rate is exponential, depending on the  $I_e$  and  $B_{pl}$  values. As it can be observed, the smaller the value of the parameter  $I_e$ , the higher the starting point of the curve (higher MOS). Additionally, the term  $B_{pl}$  seems to affect the slope of the curves, with more steep curves for small values of  $B_{pl}$ . A typical example is the case of G.711 in Figure 5. For this case, the small deterioration due to equipment impairments ( $I_e = 0$ ) leads to a high QoE for a small percentage of lost packets ( $P_{pl}$ ); however a smaller packet-loss robustness factor ( $B_{pl} = 4.3$ ) leads to a steep slope, i.e., to a sharp QoE degradation as the percentage of lost packets ( $P_{pl}$ ) increases.



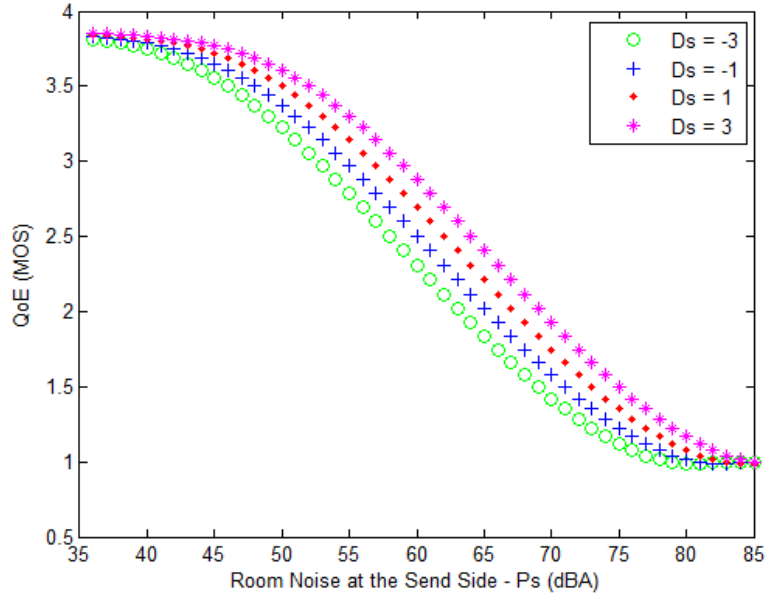
**Figure 5:** Impact of packet-loss on MOS



**Figure 6:** Impact of the noise at receiver's side on MOS

In Figure 6, we can observe that as the intensity of noise at the receiver increases, the feeling of good speech quality decreases. This is a reasonable result, but a quite interesting observation is that the  $D_r$  parameter slightly affects the way that QoE is related with  $P_r$ . More specifically, for  $P_r < 45\text{dBA}$ , and for all evaluated  $D_r$  values, no effect on QoE level is observed. This quantifies the noise range that does not influence the QoE. However, for higher noise values the QoE level reduces exponentially.

Finally, Figure 7 depicts the impact of the environment at the sender side on the QoE level at the receiver, and as it can be observed, as the noise increases the QoE decreases exponentially. The way that this exponential decrement is performed depends on the  $D_s$  value. Practically, since possible errors and faults in the device of the sender travel throughout the network, small noise degradations at the sender device escalate on their way to the receiver causing notable impact to QoE, as depicted by the shift of the curves in Figure 7.



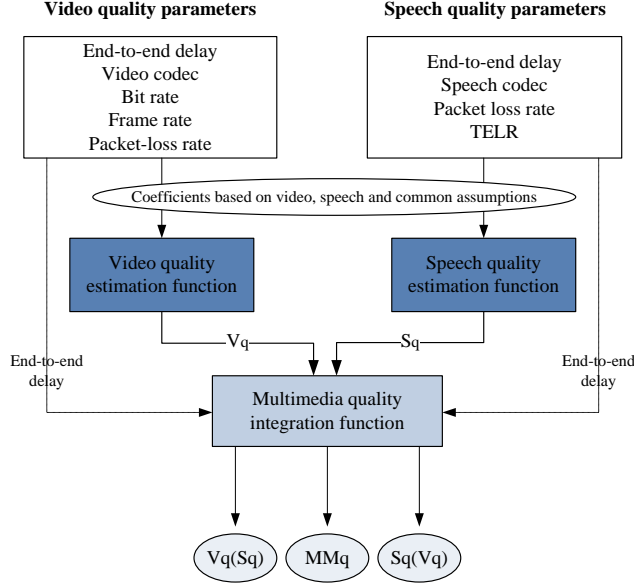
**Figure 7:** Impact of the noise at sender's side on MOS

### 3.2 Parametric QoE estimation for online video

In this section we focus on the online video service and we examine the ITU G.1070 recommendations [27] for QoE estimation. ITU Rec. G.1070 describes a parametric model applicable to online multimedia services over IP, such as a video conference. The model consists of three functions, named the *video quality estimation*, the *speech quality estimation* and the *multimedia quality integration* functions.

The degradation caused by pure delay is considered only in the multimedia quality integration function. The model provides three output quality metrics in MOS scale, named the multimedia quality ( $MM_q$ ), the video quality influenced by speech quality ( $V_q(S_q)$ ), and the speech quality influenced by video quality ( $S_q(V_q)$ ). Note that various implementations can be found for a coding technology (e.g., MPEG-4 codecs) due to variations in coding-parameter settings and decoder characteristics. Therefore, the coefficients of video and speech quality estimation functions in this model were determined by referring to tables prepared in advance for each video and speech codec.

The framework and methodology of G.1070 is presented in Figure 8, where the key influence parameters for each one of the three aforementioned dimensions are presented (multimedia – video – speech). By mapping video, speech and common assumptions into specific coefficients, the impact of e.g., terminal type, monitor characteristics, environmental noise and conversational task on the multimedia quality is quantified.



**Figure 8:** Methodology for multimedia quality assessment based on G.1070 (based on [27])

### 3.2.1 Narrowband speech

For the estimation of the narrowband speech quality an index similar to the  $R$  factor for VoIP communications is defined. This index is denoted by  $Q$  and it is calculated as follows:

$$Q = 93.193 - I_{dte} - I_{e-eff}$$

$I_{dte}$  in the above formula represents the degradation caused by talker echo and  $I_{e-eff}$  represents the degradation caused by speech coding and packet loss. More specifically,  $I_{e-eff}$  is calculated as in the description for VoIP services, while  $I_{dte}$  is as follows:

$$I_{dte} = \left[ \frac{94.769 - R_e}{2} + \sqrt{\frac{(94.769 - R_e)^2}{4} + 100} - 1 \right] (1 - e^{-T_s})$$

where:

$$R_e = 80 + 2.5(TERV - 14)$$

and:

$$TERV = TELR - 40 \log \frac{1 + \frac{T_s}{10}}{1 + \frac{T_s}{150}} + 6e^{-0.3T_s^2}$$

With the aid of the index  $Q$  the speech quality in MOS scale, denoted by  $S_q$ , is given by the following formula:

$$S_q = \begin{cases} 1, & \text{if } Q < 0 \\ 1 + 0.035Q + Q(Q - 60)(100 - Q) \cdot 7 \cdot 10^{-6}, & \text{if } 0 < Q < 100 \\ 4.5, & \text{if } Q > 100 \end{cases}$$

### 3.2.2 Wideband speech

For wideband speech the  $Q$  index is refined as follows:

$$Q = 129 - I_{dte, WB} - I_{e-eff, WB}$$

where the corresponding  $I_{dte}$  and  $I_{e-eff}$  for wideband speech, denoted by  $I_{dte, WB}$  and  $I_{e-eff, WB}$  respectively, are used.  $I_{e-eff, WB}$  is equivalent to the  $I_{e-eff}$  term, while  $I_{dte, WB}$  is calculated as follows:

$$I_{dte,WB} = \left[ \frac{129 - R_{e,WB}}{2} + \sqrt{\frac{(129 - R_{e,WB})^2}{4} + 100} - 1 \right] (1 - e^{-T_s})$$

where:

$$R_{e,WB} = 80 + 3(TERV_{WB} - 14)$$

and:

$$TERV_{WB} = TELR + K - 40 \log \frac{1 + \frac{T_s}{10}}{1 + \frac{T_s}{150}} + 6e^{-0.3T_s^2}$$

$$K = \begin{cases} 0.08 \cdot T_s + 10, & \text{if } T_s < 100 \text{ ms} \\ 18, & \text{if } T_s > 100 \text{ ms} \end{cases}$$

Similarly, to the speech quality for narrowband speech, the speech quality for wideband speech is given by the following formula:

$$S_q = \begin{cases} 1, & \text{if } Q_x < 0 \\ 1 + 0.035Q_x + Q_x(Q_x - 60)(100 - Q_x) \cdot 7 \cdot 10^{-6}, & \text{if } 0 < Q_x < 100 \\ 4.5, & \text{if } Q_x > 100 \end{cases}$$

where:

$$Q_x = \frac{Q}{1.29}$$

### 3.2.3 Video quality estimation function

The function that provides the video quality ( $V_q$ ) is:

$$V_q = 1 + I_{coding} \cdot e^{-\frac{P_{pIV}}{D_{pIV}}}$$

where  $I_{coding}$  represents the basic video quality affected by the coding distortion under a combination of video bit rate ( $B_{rV}$  [kbit/s]) and video frame rate ( $F_{rV}$  [fps]), while the packet loss robustness factor  $D_{pIV}$  expresses the degree of video quality robustness due to packet loss, where  $P_{pIV}$  [%] represents the packet-loss rate. More specifically, it holds that:

$$I_{coding} = I_{ofr} \cdot e^{-\frac{(\ln(F_{rV}) - \ln(O_{fr}))^2}{2 \cdot D_{frV}^2}}$$

where:

$$O_{fr} = v_1 + v_2 \cdot B_{rV}, \quad 1 \leq O_{fr} \leq 30, \quad v_1, v_2: \text{constants}$$

Note that when  $F_{rV} = O_{fr}$ , then  $I_{coding} = I_{ofr}$ , and  $I_{ofr}$  represents the maximum video quality at each bit rate:

$$I_{ofr} = v_3 - \frac{v_3}{1 + \left(\frac{B_{rV}}{v_4}\right)^{v_5}}, \quad 0 \leq I_{ofr} \leq 4, \quad v_3, v_4, v_5: \text{constants}$$

Additionally,  $D_{frV}$  represents the degree of video quality robustness due to frame rate and is given by:

$$D_{frV} = v_6 + v_7 \cdot B_{rV}, \quad 0 < D_{frV}, \quad v_6, v_7: \text{constants}$$

Finally,  $D_{pIV}$  parameter is as follows:

$$D_{pIV} = v_{10} + v_{11} \cdot e^{-\frac{F_{rV}}{v_8}} + v_{12} \cdot e^{-\frac{B_{rV}}{v_9}}, \quad v_8, v_9, v_{10}, v_{11}, v_{12}: \text{constants}$$



For the case studies defined in Table 4, we depict in Table 5 the values of the  $v_i, i = 1, 2, \dots, 12$  coefficients.

**Table 4:** Case studies for the derivation of the  $v_i, i = 1, 2, \dots, 12$  coefficients

Factors	#1	#2	#3	#4	#5
Codec type	MPEG-4	MPEG-4	MPEG-2	MPEG-4	ITU H.264
Video format	QVGA	QQVGA	VGA	VGA	VGA
Key frame interval	1	1	1	1	1
Video display size (inch)	4.2	2.1	9.2	9.2	9.2

**Table 5:** Values for  $v_i, i = 1, 2, \dots, 12$  coefficients

Parameter	#1	#2	#3	#4	#5
$v_1$	1.431	7.160	4.78	1.182	5.517
$v_2$	$2.228 \times 10^{-2}$	$2.215 \times 10^{-2}$	$1.22 \times 10^{-2}$	$1.11 \times 10^{-2}$	$1.29 \times 10^{-2}$
$v_3$	3.759	3.461	2.614	4.286	3.459
$v_4$	184.1	111.9	51.68	607.86	178.53
$v_5$	1.161	2.091	1.063	1.184	1.02
$v_6$	1.446	1.382	0.898	2.738	1.15
$v_7$	$3.881 \times 10^{-4}$	$5.881 \times 10^{-4}$	$6.923 \times 10^{-4}$	$-9.98 \times 10^{-4}$	$3.55 \times 10^{-4}$
$v_8$	2.116	0.8401	0.7846	0.896	0.114
$v_9$	467.4	113.9	85.15	187.24	513.77
$v_{10}$	2.736	6.047	1.32	5.212	0.736
$v_{11}$	15.28	46.87	539.48	254.11	-6.451
$v_{12}$	4.17	10.87	356.6	268.24	13.684

### 3.2.4 Multimedia quality integration function

The multimedia quality,  $MM_q$ , is estimated as follows:

$$MM_q = m_1 \cdot MM_{SV} + m_2 \cdot MM_T + m_3 \cdot MM_{SV} \cdot MM_T + m_4, \quad m_1, m_2, m_3, m_4: \text{constants}$$

The first term in the formula above ( $MM_{SV}$ ) is:

$$MM_{SV} = m_5 \cdot S_q + m_6 \cdot V_q + m_7 \cdot S_q \cdot V_q + m_8, \quad m_5, m_6, m_7, m_8: \text{constants}$$

while the term  $MM_T$  is:

$$MM_T = \max\{AD + MS, 1\}$$

where:

$$AD = m_9 \cdot (T_S + T_V) + m_{10}, \quad m_9, m_{10}: \text{constants}$$

and:

$$MS = \begin{cases} \min\{m_{11} \cdot (T_S - T_V) + m_{12}, 0\}, & \text{if } T_S > T_V \\ \min\{m_{13} \cdot (T_V - T_S) + m_{14}, 0\}, & \text{if } T_V > T_S \end{cases}$$

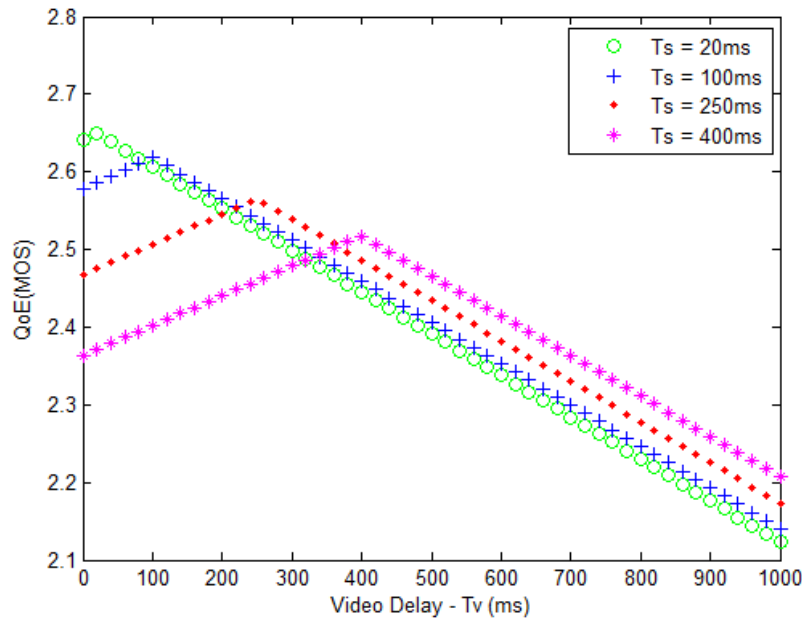
Some values for the  $m_i, i = 1, 2, \dots, 14$  coefficients are given in Table 6.

**Table 6:** Values of  $m_i, i = 1, 2, \dots, 14$  coefficients for 4.2 and 2.1 inch display

Coefficients	4.2 inch	2.1 inch
$m_1$	$-4.457 \times 10^{-1}$	$-6.966 \times 10^{-1}$
$m_2$	$-6.638 \times 10^{-1}$	$-8.127 \times 10^{-1}$
$m_3$	$4.042 \times 10^{-1}$	$4.562 \times 10^{-1}$
$m_4$	2.321	3.003
$m_5$	$-3.255 \times 10^{-1}$	$-1.638 \times 10^{-1}$
$m_6$	$3.309 \times 10^{-1}$	$3.626 \times 10^{-1}$
$m_7$	$1.494 \times 10^{-1}$	$1.291 \times 10^{-1}$
$m_8$	$5.457 \times 10^{-1}$	$5.456 \times 10^{-1}$
$m_9$	$-3.235 \times 10^{-4}$	$-1.251 \times 10^{-4}$
$m_{10}$	3.915	3.763
$m_{11}$	$-1.377 \times 10^{-3}$	$-1.065 \times 10^{-3}$
$m_{12}$	0.000	$1.465 \times 10^{-2}$
$m_{13}$	$-1.095 \times 10^{-3}$	$-1.002 \times 10^{-3}$
$m_{14}$	0.000	0.000

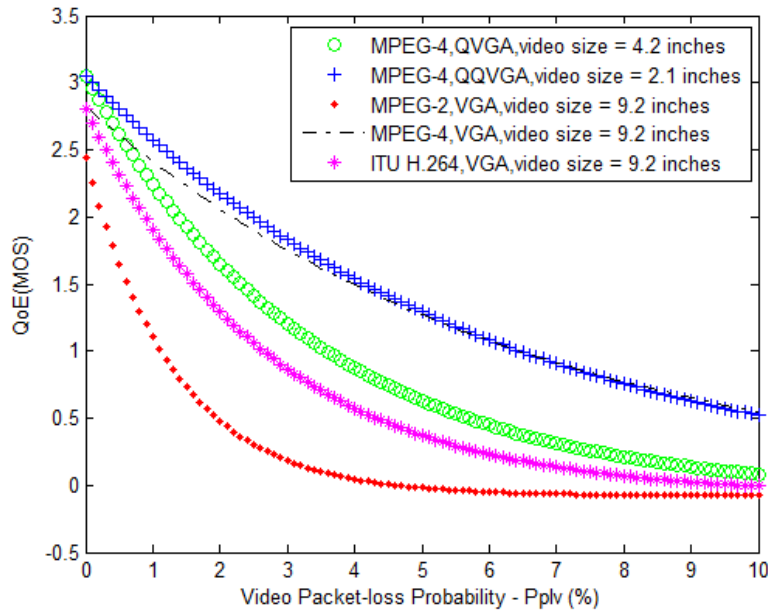
### 3.2.5 Numerical results

Below, we study the influence of key quality parameters on the MOS level of online video services. First, of all, we evaluate the impact of video delay on QoE and present the results in Figure 9. The curves in this figure show a “split-linear” behavior for the QoE level as the video delay increases. As it can be observed, the higher QoE level of a multimedia service, among other conditions, requires similar delay values between speech and video signals. This explains why in each curve of Figure 9, the higher point is observed when  $T_S = T_V$  (i.e., full synchronization of speech and video).

**Figure 9:** MOS values for different Video ( $T_V$ ) and Audio/Speech ( $T_S$ ) delays

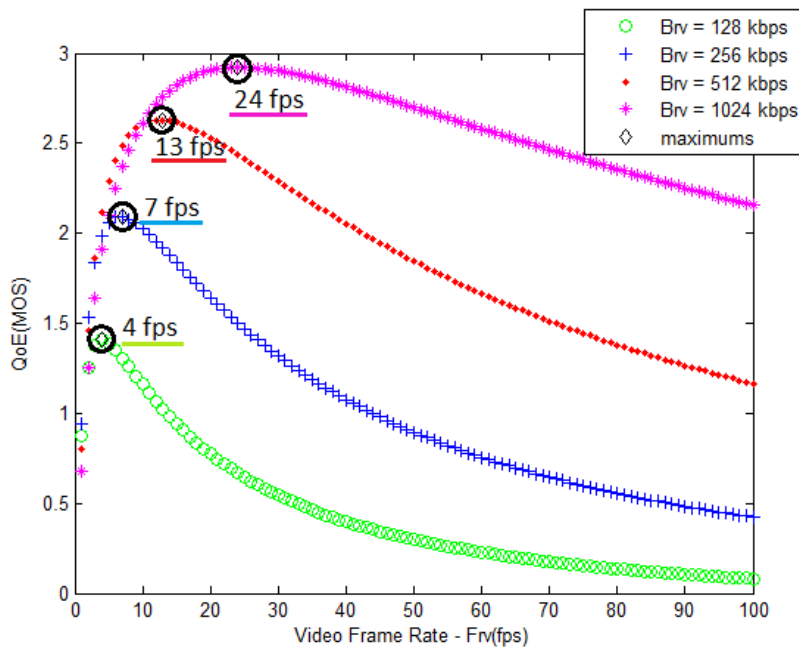
According to Figure 10, an exponential decreasing behavior for the QoE is observed as the percentage of lost packets increases. However the rate of the QoE degradation, i.e., the slope of the curves in this

figure, depends on the characteristics of the video and of the projector. Specifically, in Figure 10 it is shown that with MPEG-4 encoding type, QQVGA type analysis (i.e., 160x120 or 120x160 pixels), and 2.1 inches display size, the packet losses affect the end user's QoE less. In other words, it is revealed that in a network with high packet loss rates the multimedia services projected to small (mobile) screens ("smartphones") offer users a better experience than large computer monitors.



**Figure 10:** MOS for increasing packet loss probability

Figure 11 depicts the relationship between the QoE level and the video frame rate. As it can be observed, for increasing frame rate the QoE increases up to a certain point (maximum) and then decreases exponentially. This behavior is reasonable considering that the frame transmission rate ( $F_{rV}$ ) indicates the number of frames per second entering the network, while the bit rate, denoted by  $B_{rV}$ , represents the serving rate of the network.



**Figure 11:** Impact of Frame rate and Service rate combinations on MOS

Having studied standardized parametric QoE estimation for VoIP and real-time video, next we move on to literature-based estimation methods for other popular services (namely FTP, web browsing, IPTV, video streaming, Skype). Since standardization efforts are still ongoing for these common services, we present and analyze well-cited models from research literature.

## 4. Literature-based parametric QoE estimation

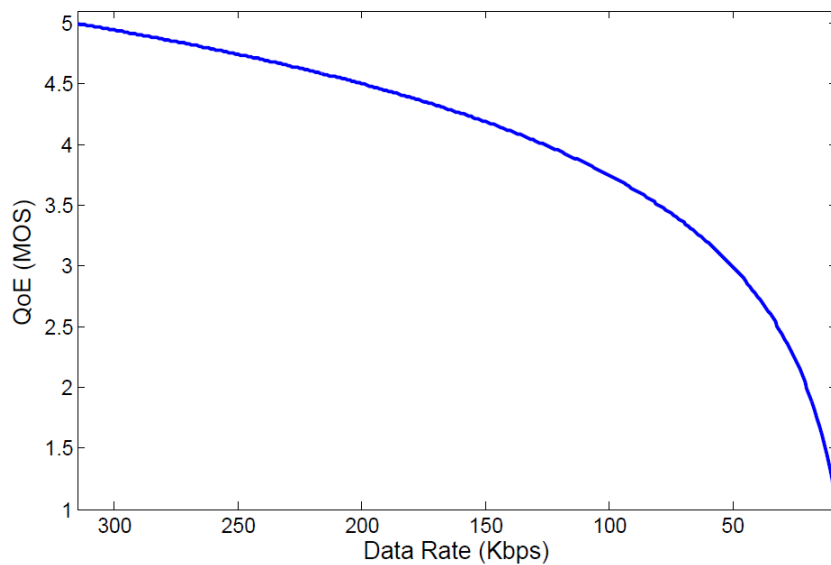
In this section, we describe non-standardized parametric QoE models which can be used for a reliable estimation of QoE for various types of services, namely file download (FTP), web browsing, lossy video streaming (IPTV), lossless video streaming (conventional and adaptive YouTube) and Skype applications.

### 4.1 Parametric QoE estimation for FTP services

The main characteristic of FTP services is that there is no need for a continuous and in-sequence packet arrival. Taking into account that the delay expected by the end-user is proportional to the size of the downloaded file and the fact that the FTP service is not adjusted in the application layer, the data rate is the dominant factor that affects the QoE level. More specifically, the model that provides MOS for an FTP service is as follows [28]:

$$MOS_{FTP} = \begin{cases} 1 & u < u^- \\ b_1 \cdot \log_{10}(b_2 \cdot u) & u^- \leq u < u^+ \\ 5 & u^+ \leq u \end{cases}$$

where  $u$  represents the data rate of the correctly received data, i.e.,  $u = R \cdot (1 - P_{error})$  where  $R$  is the data rate and  $P_{error}$  the error ratio. The values of the  $b_1$  and  $b_2$  coefficients are obtained from the upper ( $u^+$ ) and lower rate ( $u^-$ ) expectations for the service. For instance for  $u^- = 8kbps$  and  $u^+ = 315kbps$  it holds that  $b_1 = 2.5037$  and  $b_2 = 0.3136$ , while the estimated MOS values are depicted in Figure 12.



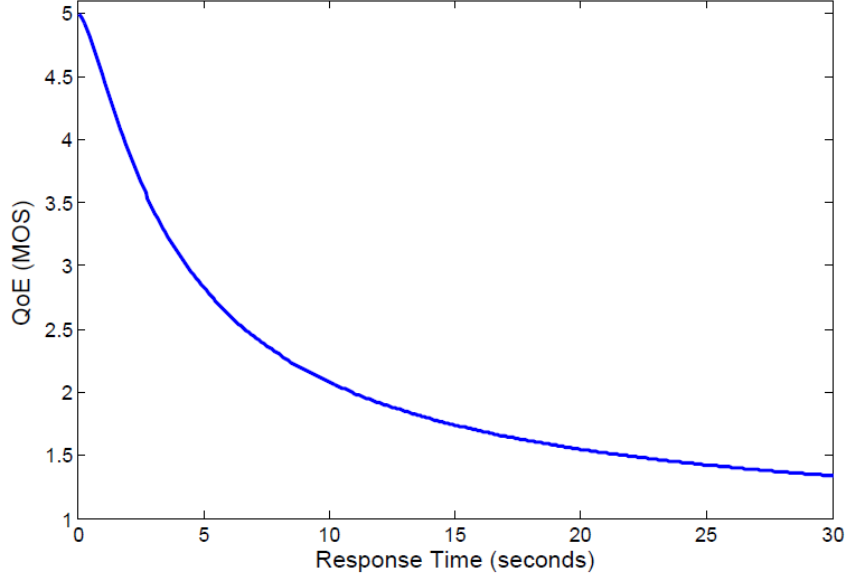
**Figure 12:** Estimated MOS for data rates in the range from 8 to 315 Kbps

### 4.2 Parametric QoE estimation for web browsing

The main observation for web browsing services is that the delay is the key QoE performance indicator. A long waiting time for the response of web will make users lose patience and negatively affect their perception for the provided service. Taking this into account, the model described in [29] is a good candidate for QoE estimation. This model is based on subjective validation tests, and the resulted empirical formula is as follows:

$$MOS_{web} = 5 - \frac{578}{1 + (11.77 + 22.61/\tau)^2}$$

where  $\tau$  is the response time. In Figure 13 the estimated MOS for different response times is depicted.



**Figure 13:** Estimated MOS for various response times

### 4.3 Parametric QoE estimation for video streaming

Due to their increasing popularity, video services cause the majority of traffic over the Internet, while they are characterized by high resource requirements. Therefore, the estimation of QoE for video streaming applications becomes of notable importance. In this section, we study two different types of video streaming: a) IPTV, which is a lossy type of service, and b) Video on Demand, which is a lossless service type. For the latter, we focus on the paradigm of YouTube, and further study it in two versions, namely adaptive (or monolithic) and non-adaptive streaming over HTTP.

#### 4.3.1 IPTV

IPTV is a common video streaming service. It is UDP-based and is therefore prone to packet losses. In [30],[31] a MOS prediction formula is proposed for three video content types, named “Slight movement (SM)”, “Gentle walking (GW)” and “Rapid movement (RM)”. This formula considers the objective parameters Send Bitrate ( $SBR$ ), Frame Rate ( $FR$ ) and Packet Error Rate ( $PER$ ):

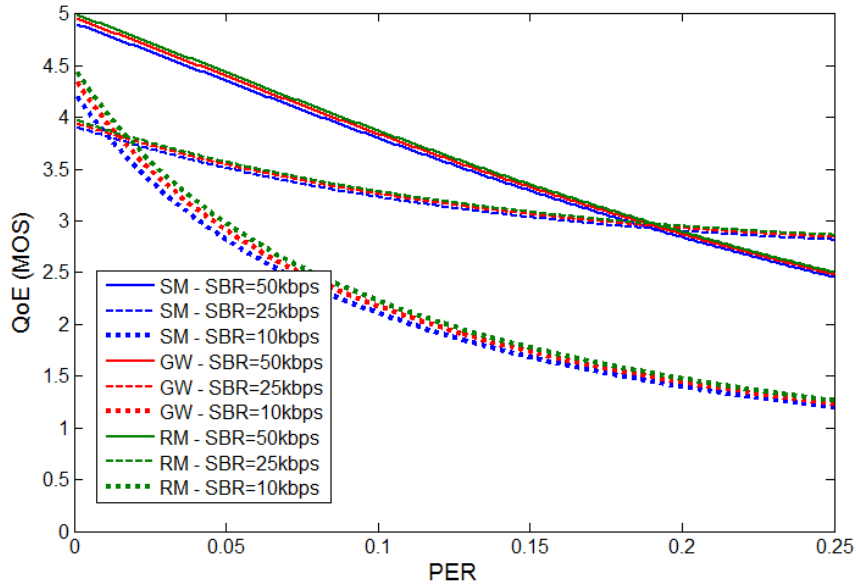
$$MOS_{IPTV} = \frac{a_1 + a_2 \cdot FR + a_3 \cdot \ln(SBR)}{1 + a_4 \cdot PER + a_5 \cdot (PER)^2}$$

where the coefficients:  $a_1, a_2, a_3, a_4, a_5$ , are obtained by linear regression of the proposed model with the training set of video sequences. More specifically, the values depicted in Table 7 have been experimentally calculated.

**Table 7:** Typical values for the coefficients per video content type

Coefficient	SM	GW	RM
$a_1$	4.5796	3.4757	3.0946
$a_2$	-0.0065	0.0022	-0.0065
$a_3$	0.0573	0.0407	0.1464
$a_4$	2.2073	2.4984	10.0437
$a_5$	7.1773	-3.7433	0.6865

In Figure 14 we consider the values included in Table 7 and  $FR = 30$  to estimate MOS for various  $PER$  and  $SBR$  values. As depicted in this figure, there is a decrement of the MOS value for increasing  $PER$  values, while this decrement is strongly correlated with the  $SBR$  value. More specifically, as it can be observed in Figure 14, for the selected  $FR$  value ( $FR = 30$ ) a more stable performance is achieved when  $SBR = 25Kbps$ . For higher  $SBR$  values (e.g.,  $SBR = 50Kbps$ ) a linear degradation of the MOS is observed, leading to low MOS performance for high  $PER$  values. Additionally, for lower  $SBR$  values (e.g.,  $SBR = 10Kbps$ ) the degradation of the MOS as the  $PER$  increases is exponential, leading very fast to low MOS values.



**Figure 14:** Estimated MOS for various  $SBR$  values and increasing  $PER$

#### 4.3.2 YouTube - conventional streaming

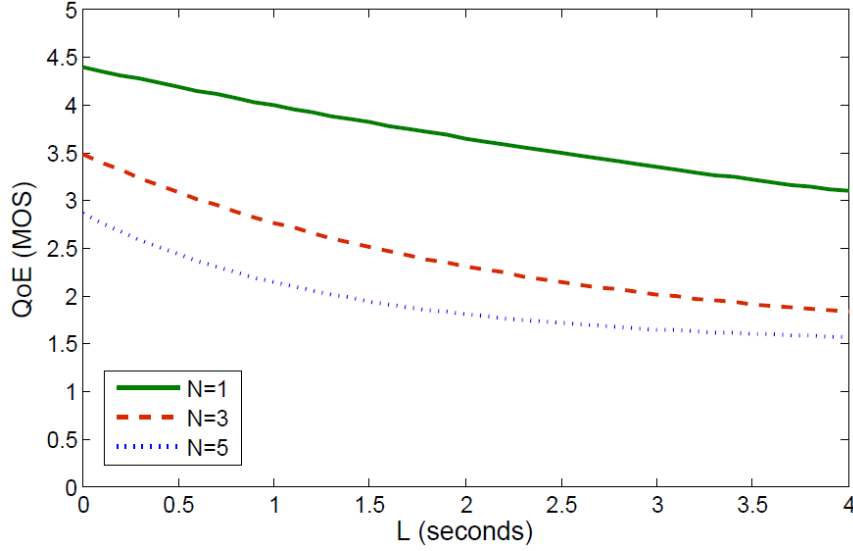
Another type of video content delivery that deserves attention is that of streaming pre-encoded video, known as Video on Demand (VoD). YouTube is the most popular paradigm in this category. YouTube is not subject to packet losses, since the connection is TCP-based. The following description regarding YouTube QoE analysis is based mostly on [32], the authors of which demonstrate an extensive literature on the topic. The most popular models of YouTube QoE are built based on subjective experiments, conducted either in controlled laboratory environments or using crowdsourcing tests and field studies. Through these well-designed experiments, the system-level key influence factors that affect the YouTube video delivery quality may be found, which are:

- Number of stalling events,  $N$
- Duration of stalling events,  $L$
- Total video duration,  $T$  (significant is the total stalling duration over the whole video duration)
- Initial delay (video start-up delay).

Regarding these influence parameters, some important findings, that follow up from these subjective experiments are:

- The number of stalling events together with the stalling length (i.e., the stalling pattern) are clearly dominating the user perceived quality.
- Initial delays have almost no influence on MOS for videos of duration 60s and 30s, namely they are tolerated up to a reasonable level.

- User ratings are statistically independent from video characteristics such as resolution, video motion, content type, encoding scheme and video bit rate. The stalling pattern is what really influences the end-user's experience.



**Figure 15:** Estimated MOS for various number of stalling events and stalling durations

Specifically for YouTube QoE, the mapping function proposed in [32] has the form:

$$MOS_{youtube} = \alpha \cdot e^{-\beta(L) \cdot N} + \gamma$$

where  $\alpha$ ,  $\gamma$  and  $\beta(L)$  are coefficients derived from the experimental process. More specifically,  $\beta(L)$  has a linear relation with the stalling duration  $L$ , which is defined as:  $\beta(L) = 0.15 \cdot L + 0.19$ . Typical values for  $\alpha$  and  $\gamma$  coefficients are  $\alpha = 3.5$  and  $\gamma = 1.5$ . In Figure 15, the MOS is depicted for various numbers of stalling events and stalling durations. As it can be observed from this figure, the number of stalling events is the dominant factor that affects MOS. Also, for a stalling duration of up to a point the degradation of the MOS is fast, while after that point the degradation of MOS is slower. This is a reasonable result, since when the MOS reaches a very low value (of about 1.5 in Figure 15, e.g., for  $N = 2$  and  $N = 3$  this is reached when  $L = 2$  seconds) the effect of a longer stalling duration is hardly considered by the end-users.

QoE metrics tailored to the YouTube application have also been defined. These metrics can be used instead of the MOS scale to get an indication about the quality of the user viewing experience. For instance, the reception ratio,  $\rho$ , is calculated as follows [32]:

$$\rho = \frac{\text{Download throughput or Bottleneck capacity}}{\text{Video encoding rate}}$$

Although the reception ratio cannot be directly related to QoE, it is a good indicator about whether there are problems in the network. If  $\rho > 1$ , the video has good quality, otherwise poor.

Moreover, rate  $\lambda$  gives a good indication of YouTube video delivery quality according to [33], and its value should ideally be zero or close to zero:

$$\lambda = \frac{\text{total stalling time}}{\text{total video elapsed time}}$$

Following a similar logic, [34] proposes the following metric:

$$\rho' = \min_{t \in (0, T)} \frac{b(t) \Delta}{B - t}$$

where:

- $t$  is the instantaneous download time
- $T$  is the total download time
- $\Delta$  is the video duration
- $B$  is the total video size in bytes
- $b(t)$  are the bytes downloaded so far

In this case,  $\rho' > 1$  indicates a stalling-free video, while  $\rho' < 1$  implies a non-seamless video session.

#### 4.3.3 YouTube - adaptive streaming

In adaptive streaming scenarios, a video file is broken into multiple segments, while each segment is available at different quality levels. These levels may differ in video bit rate (bps) or in the video resolution, etc. Then, each user independently requests the next segment in a specific quality level, based on the user's current perception of available bandwidth for this session.

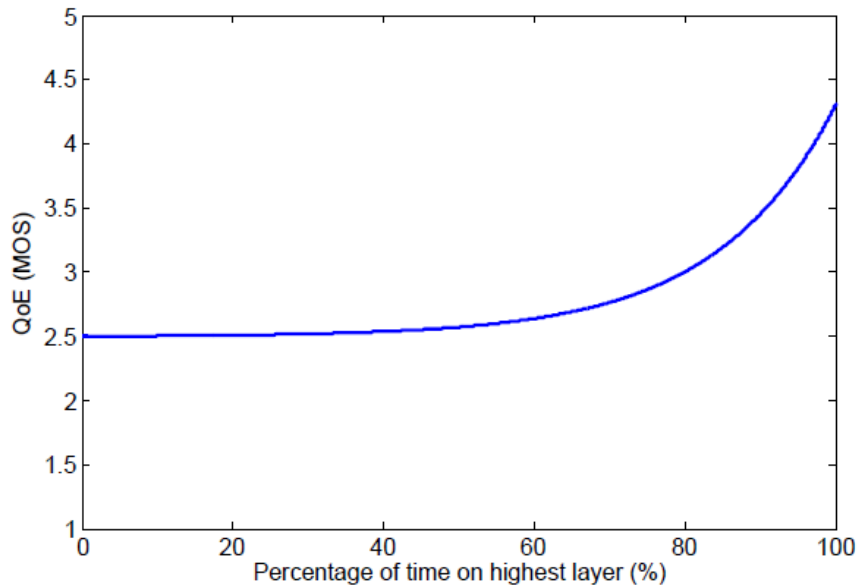
For the case of HTTP Adaptive Streaming (HAS), [35] proposes a simple but highly accurate QoE model, that is:

$$MOS_{HAS} = 0.003 * e^{0.064*t} + 2.498, \text{ where } t = \text{time on highest layer.}$$

Based on this formula, the QoE of HAS applications depends mainly on the fraction of time that the highest layer is being played out over the total viewing time. Moreover, as it can be seen in Figure 16, the MOS is bounded by the quality that can be achieved by the highest and lower layers (4.3 and 2.498 respectively).

Another important influence factor of HAS QoE based on [35] is the “adaptation amplitude”, which refers to the gap between two subsequent quality levels. In the case that the highest and then the lowest quality levels are sequentially selected (or vice versa), the amplitude will be high and the QoE impression will be low; if however such intense switches are refrained, the amplitude will be lower. The higher the amplitude, the worse the perception of the overall quality at the end user.

Moreover, some additional quality influence factors, with lower impact though, are the frequency of switches (i.e., adaptation events) and their direction. Last but not least, it has been shown that the buffer length of the user's application and the size of the segment encoded at the server's side play an important role on the QoE [36].



**Figure 16:** Estimated MOS for different percentages of time on highest HAS layer

For the case of HAS, additionally the “activity factor” metric proposed in [34] applies:



$$a = \frac{\text{total time of actual data download}}{\text{total time elapsed for complete video download}}$$

If this metric is close to 1, it means that the client was “struggling” to download each segment on time; however if this factor is much lower than 1, it means that the client had sufficiently available bandwidth and could even afford higher video resolutions, if those were available. Note that gaps in the video download are occurring because the client is not buffering the full content at once, but is just targeting to maintain an acceptable buffer threshold.

#### 4.4 Parametric QoE estimation for Skype

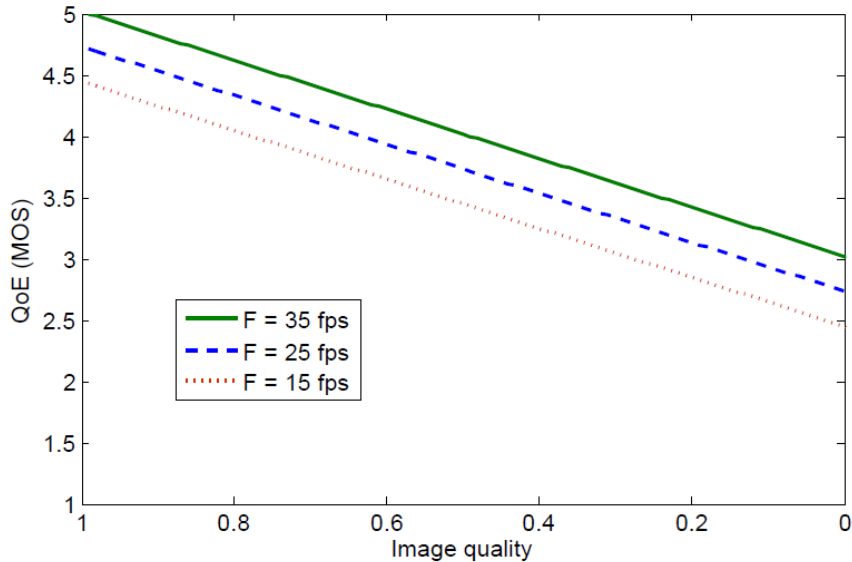
For Skype applications, a practical QoE estimation approach can be found in [37]. The proposed model has been derived by measurements conducted on Skype video calls. It has been found that three types of resolutions are available, namely 160x120, 320x240 and 640x480. Moreover, the maximum frame rate is 35fps. Then, the MOS level for this service type is as follows:

$$MOS_{skype} = \begin{cases} 1 & res = 160 \times 120 \\ 2 & res = 320 \times 240 \\ 3 + \frac{F}{35fps} + (2I - 1) & res = 640 \times 480 \end{cases}$$

where  $I$  is the Image Quality ranging from 0 (worst) to 1 (best) and  $F$  is the Frame Rate.

In Figure 17, we evaluate the relationship between Skype QoE and the image quality. Moreover, we vary the frame rate to study its impact. As expected, MOS degrades linearly while the image quality is reduced, while the frame rate also has a significant influence on the perceived QoE.

Based on this model, the authors in [37] also propose an adaptation mechanism of the Skype application to poor network conditions. Assuming a maximum acceptable threshold of the packet delay, if this threshold is reached, the Skype application starts to gradually degrade first the frame rate, then the image quality and finally, if required, the resolution. This adapting behavior helps sustain a viable and meaningful communication between two Skype applications, compromising on the quality though.



**Figure 17:** The impact of Image Quality and Frame Rate on Skype MOS

Having described both standardized and literature-based QoE estimation models, next we move on to summarized results and potential research and exploitation directions.

## 5. Summarized results and exploitation directions

The parametric QoE estimation models described above define a major set of formulas that can be exploited by academia and industry to understand how the end users perceive the quality of a provided service. Summarizing the study in the previous sections, in Table 8 we indicate the key parametric QoE estimation models available in the literature and list the major configuration parameters (MCPs) and key performance indicators (KPIs) that affect the QoE performance per service type.

**Table 8:** Parametric QoE estimation per service type

Service Type	QoE estimation Model	MCPs and KPIs
File transfer	Data rate-based formula [28]	Data rate, expected upper and lower data rate
Web Browsing	Response time-based formula [29]	Response time
Skype	Skype-specific formula [37]	Frame rate, image quality, resolution
VoIP	ITU-T Rec. G107, E-model [24][25][26]	Packet loss ratio, delay, codec, coding rate
Video streaming	IPTV model [30][31]	Data rate, frame rate
	YouTube (conventional) model [32]	Number of stalling events, duration of stalling events, video duration
	YouTube with adaptive streaming model [35][36]	Time on highest layer, amplitude, frequency of quality switches
Online video	ITU-T Rec. G 1070, E-model [27]	Packet loss ratio for audio and video packets, relative delay between video and audio packets, data rate, frame rate, monitor size

From an academic and research perspective, through a clear collection of parametric QoE models, an easy and straightforward “translation” of QoS research works to QoE vocabulary may be applied. To be more specific, a potential research direction that can be aided by our work is the direct quantification of the impact of existing research works on QoE. This may be possible either via the realization of appropriate QoE estimation models (column 2 of Table 8) or by quantifying a potential improvement on specific MCPs and KPIs per service (column 3 of Table 8). Furthermore, the collection of KPIs helps identify the specific influence factors that play the most important role on the end user’s perceived quality, guiding in this way future works towards devising network and application mechanisms that target at improving exactly those factors.

Regarding the impact of explicit QoE parametric models on the industry sector, this is twofold. On the one hand, it can help operators design their networks in a QoE- rather than QoS-meaningful way. That is, the operators are guided to give emphasis on designing and maintaining their networks in such a way that requirements regarding the KPIs per service are met (e.g., through a QoE-meaningful resource provisioning, a proper positioning of network servers and gateways – e.g., close to the end-user, etc.). Furthermore, attention on the per-service KPIs has to be given during the network management process, i.e., during the network’s real-time operation. Mechanisms such as scheduling, mobility management and power control can be tuned so that proper weight is given on the actual QoE impact factors per service. In this way, an indirect QoE improvement will be achieved through the targeted enhancement of carefully selected QoS parameters. What is more, if we consider the recently emerged paradigm of “User Provided Networking” (UPN), like the one proposed in [38], a massive potential is unlocked. According to this paradigm, end-users are actively involved not only in service evaluation tasks by providing feedback about the experienced quality either passively (e.g.,

device capabilities, response times, context of use) or actively (e.g., MOS feedback), but they can also participate in the service provisioning loop by becoming “micro-providers”, given the proper incentives.

Another important, even though less obvious capability exposed by the collection of the different KPIs per service, is the opportunity to achieve a more meaningful cross-service resource provisioning towards a) higher QoE and b) higher resource utilization. All services are currently competing for the same resources on an equal basis; nevertheless it would make more sense to allocate the limited resources on a service-dependent rather than on a service-oblivious (i.e., blindly fair) way. This may be possible a) by performing the scheduling process on a per-flow basis (e.g., prioritize a more delay-critical service with respect to another) or b) by optimizing the sum QoE in a cell by taking appropriate cross-service management decisions. Regarding the latter, a potential enabler is to exploit the adapting behavior of Skype or HAS applications and provoke a deliberate quality degradation at specific Skype or HAS flows, so that resources are moved to other applications, whose QoE/KPI is currently at a critical level. The goal would be to keep all users’ QoE above a critical threshold, or, to achieve a maximum possible summed QoE. Note, that average QoE values per cell should not be considered as appropriate indicators of quality though (e.g., a cellular MOS of 3 may be a result of user1’s MOS=1 and user2’s MOS=5); on the contrary each flow should be treated independently, or, at least, standard deviations should be considered as well (see [39]).

## **6. Conclusions**

As we are moving closer and closer to future network generations, the human factor is becoming the epicenter of attention and the driving force for the network design. Thus, the comprehension and, in extension, the control of the provisioned QoE to the end-users has become a necessity for network operators. Parametric QoE estimation models are a prerequisite for this purpose. They constitute the ideal tools towards live network quality monitoring and, hence, QoE management. Nevertheless, despite the increased interest from academia and industry to push towards a QoE service provisioning model, a clear/comprehensive manual on the available parametric models and the critical QoE performance parameters per service type is currently missing. Identifying this gap, this paper aspires to become a thorough and handy “manual”, currently absent from the literature, that identifies and describes appropriate parametric models for popular services nowadays, such as YouTube, Skype and IPTV, as well as describe and study standardized ones. Therefore, this paper may become a stand-alone, useful tutorial both for researchers and operators, who are interested in moving from the pure technical QoS-domain to a more meaningful QoE-domain, so that they can really understand and influence the impact of their network decisions on the final recipient, the end-user.

## **Acknowledgments**

This work has been partly funded by the European Commission under the auspices of the FP7-PEOPLE MITN-CROSSFIRE project (grant 317126) and the H2020-MSCA-RISE CASPER project (grant 645393).

## Appendix A: Default values for ITU-T G.107 (E-model)

Parameter	Abbr.	Unit	Default value	Permitted range	Remark
Send loudness rating	SLR	dB	+8	0 ... +18	Total values between microphone or receiver and 0 dBr-point
Receive loudness rating	RLR	dB	+2	-5 ... +14	Total values between microphone or receiver and 0 dBr-point
Sidetone masking rating	STMR	dB	15	10 ... 20	Fixed relation: $LSTR = STMR + D$ Equations provide also predictions for $STMR > 20$ dB. However, such values can hardly be measured in a reliable way.
Listener sidetone rating	LSTR	dB	18	13 ... 23	Fixed relation: $LSTR = STMR + D$
D-Value of telephone, send side	Ds	-	3	-3 ... +3	Fixed relation: $LSTR = STMR + D$
D-Value of telephone, receive side	Dr	-	3	-3 ... +3	Fixed relation: $LSTR = STMR + D$
Talker echo loudness rating	TELRL	dB	65	5 ... 65	
Weighted echo path loss	WEPL	dB	110	5 ... 110	
Mean one-way delay of the echo path	T	ms	0	0 ... 500	
Round-trip delay in a 4-wire loop	Tr	ms	0	0 ... 1000	
Absolute delay in echo-free connections	Ta	ms	0	0 ... 500	
Delay sensitivity	sT	—	1	0.4 ... 1	Only predefined settings are allowed for sT, reflecting specific use cases
Minimum perceivable delay	mT	ms	100	20 ... 150	Only predefined settings are allowed for mT, reflecting specific use cases
Number of quantization distortion units	qdu	-	1	1 ... 14	
Equipment impairment factor	Ie	-	0	0 ... 40	If Ppl > 0%, then the Bpl must match the codec, packet size, and PLC assumed
Packet-loss robustness factor	Bpl	-	4.3	4.3 ... 40	Currently under study. If Ppl > 0%, then the Bpl must match the codec, packet size, and PLC assumed
Random packet-loss probability	Ppl	%	0	0 ... 20	Currently under study. If Ppl > 0%, then the Bpl must match the codec, packet size, and PLC assumed
Burst ratio	BurstR	-	1	1 ... 8	Currently under study. E-model predictions for values of BurstR > 2 are only valid if the packet loss percentage is Ppl < 2%
Circuit noise referred to 0 dBr-point	Nc	dBm0p	-70	-80 ... -40	
Noise floor at the receive side	Nfor	dBmp	-64	-	Currently under study.
Room noise at the send side	Ps	dB(A)	35	35 ... 85	
Room noise at the receive side	Pr	dB(A)	35	35 ... 85	
Advantage factor	A	-	0	0 ... 20	

## References

- [1] C. Quadros, E. Cerqueira, A. Neto, A. Pescape, A. Riker, R. Immich, and M. Curado, "A quality of experience handover system for heterogeneous multimedia wireless networks," in *IEEE International Conference on Computing, Networking and Communications (ICNC)*, 2013, pp. 1064–1068.
- [2] V. F. Monteiro, D. A. Sousa, T. F. Maciel, F. R. M. Lima, E. B. Rodrigues, and F. R. P. Cavalcanti, "Radio resource allocation framework for quality of experience optimization in wireless networks," *IEEE Netw.*, vol. 29, no. 6, pp. 33–39, Nov. 2015.
- [3] H. Nam, K.-H. Kim, J. Y. Kim, and H. Schulzrinne, "Towards QoE-aware video streaming using SDN," in *IEEE Global Communications Conference (GLOBECOM)*, 2014, pp. 1317–1322.
- [4] D. Tsolkas, E. Liotou, N. Passas, and L. Merakos, "The need for QoE-driven interference management in femtocell-overlaid cellular networks," *Mobile and Ubiquitous Systems: Computing, Networking, and Services*, Springer International Publishing, Editors: I. Stojmenovic, Z. Cheng, and S. Guo, vol. 131, pp. 588–601.
- [5] F. Song, D. Huang, H. Zhou, H. Zhang, and I. You, "An optimization-based scheme for efficient virtual machine placement," *Int. J. Parallel Program.*, vol. 42, no. 5, pp. 853–872, Oct. 2014.
- [6] S. Möller, W.-Y. Chan, N. Cote, T. Falk, A. Raake, and M. Waltermann, "Speech quality estimation: Models and trends," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 18–28, Nov. 2011.
- [7] S. Jelassi, G. Rubino, H. Melvin, H. Youssef, and G. Pujolle, "Quality of Experience of VoIP service: A survey of assessment approaches and open issues," *IEEE Commun. Surv. Tutorials*, vol. 14, no. 2, pp. 491–513, Jan. 2012.
- [8] A. Takahashi, D. Hands, and V. Barriac, "Standardization activities in the ITU for a QoE assessment of IPTV," *IEEE Communications Magazine*, vol. 46, no. 2, pp. 78–84, Feb. 2008.
- [9] Y. Chen, K. Wu, and Q. Zhang, "From QoS to QoE: A tutorial on video quality assessment," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 2, pp. 1126–1165, Jan. 2015.
- [10] Recommendation ITU-T P.800 (1998), Methods for subjective determination of transmission quality.
- [11] K.-T. Chen, C.-C. Tu, and W.-C. Xiao, "OneClick: A framework for measuring network quality of experience," in *IEEE International Conference on Computer Communications (INFOCOM)*, 2009, pp. 702–710.
- [12] T. Hoßfeld, C. Keimel, M. Hirth, B. Gardlo, J. Habigt, K. Diepold, and P. Tran-Gia, "Best practices for QoE crowdtesting: QoE assessment with crowdsourcing," *IEEE Trans. Multimed.*, vol. 16, no. 2, pp. 541–558, Feb. 2014.
- [13] Recommendation ITU-T G.1011 (2010), Reference guide to quality of experience assessment methodologies.
- [14] R. Schatz, T. Hoßfeld, L. Janowski, and S. Egger, "From packets to people: Quality of experience as a new measurement challenge," in *Data Traffic Monitoring and Analysis*, vol. 7754, E. Biersack, C. Callegari, and M. Matijasevic, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 219–263.
- [15] S. Chikkerur, V. Sundaram, M. Reisslein, and L. J. Karam, "Objective video quality assessment methods: A classification, review, and performance comparison," *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 165–182, Jun. 2011.
- [16] P. Juluri, V. Tamarapalli, and D. Medhi, "Measurement of quality of experience of video-on-demand services: A survey," *IEEE Commun. Surv. Tutorials*, vol. 18, no. 1, pp. 401–418, Jan. 2016.
- [17] M. Alreshoodi and J. Woods, "Survey on QoE/QoS correlation models for multimedia services," *International Journal of Distributed and Parallel Systems (IJDPs)*, vol.4, no.3, May 2013.
- [18] D. De Vera, P. Rodriguez-Bocca, G. Rubino, "Automatic quality of experience measuring on video delivering networks," *ACM SIGMETRICS Perform. Eval. Rev.*, vol. 36, no. 2, pp. 79–82, Aug. 2008.
- [19] P. Charonyktakis, M. Plakia, I. Tsamardinos, and M. Papadopoulou, "On user-centric modular QoE prediction for VoIP based on machine-learning algorithms," *IEEE Trans. Mob. Comput.*, vol. 15, no. 6, pp. 1443–1456, Jun. 2016.

- [20] M. Volk, J. Sterle, U. Sedlar, and A. Kos, "An approach to modeling and control of QoE in next generation networks [Next Generation Telco IT Architectures]," *IEEE Commun. Mag.*, vol. 48, no. 8, pp. 126–135, Aug. 2010.
- [21] J. Hosek, P. Vajsar, L. Nagy, M. Ries, O. Galinina, S. Andreev, Y. Koucheryavy, Z. Sulc, P. Hais, and R. Penizek, "Predicting user QoE satisfaction in current mobile networks," in *IEEE International Conference on Communications (ICC)*, 2014, pp. 1088–1093.
- [22] E. Liotou, D. Tsolkas, N. Passas, and L. Merakos, "Quality of experience management in mobile cellular networks: key issues and design challenges," *IEEE Communications Magazine*, vol. 53, no. 7, pp. 145–153, July 2015.
- [23] S. Möller: *Quality Engineering - Qualität Kommunikationstechnischer Systeme*, Springer, 2010.
- [24] Recommendation ITU-T G.107 (2014), *The E-model: a computational model for use in transmission planning*.
- [25] Recommendation ITU-T G.108 (1999), *Application of the E-model: A planning guide*.
- [26] R. G. Cole and J. H. Rosenbluth, "Voice over IP performance monitoring," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 31, no. 2, pp. 9–24, Apr. 2001.
- [27] Recommendation ITU-T G.1070 (2012), *Opinion model for video-telephony applications*.
- [28] S. Thakolsri, S. Khan, E. Steinbach, and W. Kellerer, "QoE-driven cross-layer optimization for high speed downlink packet access," *Journal of Communications*, vol. 4, no. 9, pp. 669–680, Oct. 2009.
- [29] P. Ameigeiras, J. J. Ramos-Munoz, et al, "QoE oriented cross-layer design of a resource allocation algorithm in beyond 3G systems," *Computer Communications*, vol. 33, no. 5, pp. 571–582, 2010.
- [30] A. Khan, L. Sun, E. Jammeh, and E. Ifeachor, "Quality of experience driven adaptation scheme for video applications over wireless networks," *IET Communications*, vol. 4, no. 11, pp. 1337–1347, 2010.
- [31] A. Khan, Sun Lingfen, E. Ifeachor, "Content clustering based video quality prediction model for MPEG4 video streaming over wireless networks," in *IEEE International Conference on Communications (ICC)*, 2009, pp. 1–5.
- [32] T. Hoßfeld, R. Schatz, E. W. Biersack, and L. Plissonneau, "Internet video delivery in YouTube: From traffic measurements to quality of experience," in *Data Traffic Monitoring and Analysis*, E. Biersack, C. Callegari, and M. Matijasevic, Eds. Springer Berlin Heidelberg, 2013, pp. 264–301.
- [33] P. Casas, R. Schatz, and T. Hossfeld, "Monitoring YouTube QoE: Is your mobile network delivering the right experience to your customers?," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2013, pp. 1609–1614.
- [34] P. Szilagyi and C. Vulkan, "Network side lightweight and scalable YouTube QoE estimation," in *IEEE International Conference on Communications (ICC)*, 2015, pp. 3100–3106.
- [35] T. Hoßfeld, M. Seufert, C. Sieber, and T. Zinner, "Assessing effect sizes of influence factors towards a QoE model for HTTP adaptive streaming," in *IEEE International Workshop on Quality of Multimedia Experience (QoMEX)*, 2014, pp. 111–116.
- [36] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hoßfeld, and P. Tran-Gia, "A survey on quality of experience of HTTP adaptive streaming," *IEEE Commun. Surv. Tutorials*, vol. 17, no. 1, pp. 469–492, Jan. 2015.
- [37] F. Wamser, S. Deschner, T. Zinner, P. Tran-Gia, "Investigation of different approaches for QoE oriented scheduling in OFDMA networks," in *Mobile Networks and Management*, Eds. Springer, 2013, vol. 125, pp. 172–187.
- [38] Y. Wang and X. Lin, "User-provided networking for QoE provisioning in mobile networks," *IEEE Wirel. Commun.*, vol. 22, no. 4, pp. 26–33, Aug. 2015.
- [39] T. Hoßfeld, R. Schatz, and S. Egger, "SOS: The MOS is not enough!," in *IEEE International Workshop on Quality of Multimedia Experience (QoMEX)*, 2011, pp. 131–136.