International Journal of Digital Information and Wireless Communications (IJDIWC) 6(4): 241-259
The Society of Digital Information and Wireless Communications, 2016 ISSN: 2225-658X (Online); ISSN 2412-6551 (Print)

Assessment of Quality-of-Experience in Telecommunication Services

Demóstenes Z. Rodriguez, Renata L. Rosa, Rodrigo D. Nunes and Emmanuel T. Affonso
Department of Engineering – Federal University of Lavras
Minas Gerais, Brazil
{demostenes.zegarra, renata.rosa}@dcc.ufla.br, {rdantasnunes, emmanuelcomp}@gamil.com

ABSTRACT

In recent years, the number of researches in Quality-of-Experience (QoE) is increasing because this concept is used in several areas. In telecommunication services, the subscriber loyalty depends on the user's QoE level. Thus, service providers need to know the real users' QoE to improve their services offered. In this arena, the study of methods to assess QoE is very important. In this paper is shown the different approaches that compose the QoE concept, from technology, social and human aspects, differentiating it from others well known concepts such as, Quality of Service (QoS) and User Experience (UX). In order to assess user's QoE, subjective tests need to be conducted, and based on their results mathematical models are formulated to establish quality metrics. In this document, two case studies are presented. Firstly, the video streaming service is analyzed, in which the user's QoE is modeled considering the pauses as a degradation factors obtaining a video quality metric. As a second case study, a research regarding voice quality is presented, in which the voice calls in a cellular network are assessed.

KEYWORDS

QoE, MOS, quality metrics, QoS telecommunication service, video streaming, cellular networks, voice and video quality.

1 INTRODUCTION

Nowadays, communication services and their applications are growing. Such offers have raised the interest in providing ever more quality services to the subscribers. A subscriber, who is using a product or service, needs to feel and perceive a good quality associated with reasonable rates. These concerns gave rise to the concept of Quality-of-Experience (QoE), which is very useful for services providers, because the quality information can help to improve the user's satisfaction; therefore, the user's loyalty.

In this context, it is very important that communication services, especially in wireless environments, consider the QoE approach. Hence, to determine the user's QoE, the multimedia signal quality perception needs to be complemented with other criteria set related with sensorial processing, cognitive process and psychological approaches.

In recent years, researchers have tried to improve existing Quality of Service (QoS) methodologies to evaluate communication services in order to provide a most comprehensive approach for user satisfaction assessment [1]-[4].

The concept of QoS is more related to technological aspects, such as, network performance, transmission channel capacity, among others. In general, QoS deals with performance aspects of physical systems. The concept of Service Level Agreement is related to technological and economic aspects, because users pay more for a better service. As stated before, the QoE concept does not only cover technical, economic or social aspects, it is also related to human aspects, such as, their necessities and preferences. As a consequence, the concept of QoE has gained popularity in different research areas, in which one of the most important is the telecommunication services.

Thus, the QoE was defined by IEEE in Standard 3333.1 [5] as the degree of delight or annoyance of the user of an application or service resulting from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service.

Furthermore, the number of research regarding QoE presented an important growing over the last 10 years. For instance, Table 1 presents the number of citations of the words QoE and QoS in abstracts of papers registered in IEEE database, and also the ratio between QoE and QoS is shown. The period of time used for comparison is three years, only the most recent period considers two and a half years.

Table 1. Number of QoE and QoS Citations in Abstracts Obtained in the IEEE Database

Period of Reference (Years)	QoE	QoS	QoE/QoS
2002 - 2004	4	3451	0.12%
2005 - 2007	23	5328	0.43%
2008 – 2010	235	6689	3.51%
2011 - 2013	853	5291	16.12%
2014 – Jun. 2016	826	4051	20.39%

The application areas of QoE concept is very large, that include, communication and multimedia services, educational solutions, medical applications, business models, entertainment services, among others.

In this context, the main objective of this research is to demonstrate the relevance of QoE concept and its applicability in different telecommunication services.

In this research is presented two case studies, both of them are extended contributions of previous works [6]-[9]. Thus, the main contribution of this paper is shown how important is the application of QoE, and how the existent metrics can be enhanced. The motivation to present these case studies is to show the different steps of the development process to model objective quality metrics to be used in communication services.

The first case study is related to assess user's QoE in video streaming service over HTTP/TCP, in which a video quality metric named VsQM is modeled [6], considering psychological approaches. The second case study treats about voice quality assessment during a phone call [7]-[9], in which a new quality indicator is proposed.

The remainder of this work is structured as follows. Section 2 presents an overview of definition of QoE and its applications. Section 3 describes the methods for audio and video quality assessment. Section 4 introduces the first case study regarding video streaming service. Section 5 presents the study about the assessment of phone calls quality in cellular network and its applications. Finally, Section 6 draws the conclusions of this research.

2 DEFINITION AND APPLICATIONS OF QUALITY OF EXPERIENCE

In this section, firstly, we will review different general concepts related to perception, quality, and experience, on an individual basis, in order to define the overall concept of QoE. Later, QoE is compared with the concepts of QoS and User Experience (UX). Finally, the application areas of QoE and their influence factors are introduced.

The terms perception and experience are defined to understand the QoE concept. The process of perception begins with the incidence of the respective stimulus for one or more human sensory organs. Perception is a conscious processing of sensory information to which humans are exposed, and involves two stages:

- Conversion of stimulus from a sensory organ into a neural signal.
- Processing and transmission of neural signals in the central nervous system to the cortex.

The perception is influenced by events stored in memory. As a result, neural features that belong to the same object are associated. According to Cowan [10], Coltheart [11] and Baddeley [12], [13] different memory levels has been identified, each one with their roles in the process of perception, and their duration of storage. Such memories are:

- Sensory memory: the peripheral memory that stores short representations between 150 ms. to 2 s.
- Working memory: stores information lasting up to 10 s. [9]. It is also known as short-term memory.
- Long-term memory: stores information for long times even years or a lifetime.

The process of creation of quality concept may be viewed as a parallel process of cognitive high level associated with the process of experiment. The reflection can be triggered by an external task to assess what has been experienced, during or after the process of experiment. It is important to note that the emotional state of the person as well as your personality plays an important role in the procedures of quality assessment. Based on the concepts described, the term of quality refers to feelings of individual perception, sensory perception and concepts that occur in a particular situation, such as when a person experiences a multimedia service.

According to Jackson [14], the word Qualia can be seen as a property to experience something that cannot be shared by verbal or technical descriptions, thus being an individual and subjective experience. Martens and Martens [15] discuss two existing approaches to understanding quality: (1) objective, rational and oriented to a product and (2) perceptual and subjective. The first approach focuses on the characteristics and properties of an item (product or service) in terms of quality; while the second approach requires human evaluation considering terms of "assessment of excellence". Reeves and Bednar [16] define quality in a more intrusive manner, for instance, "the form that a product or service meets or exceed customer expectations"; this definition comes from the marketing literature. For these reasons, the definition of quality is based on standards, such as, the given by the International Organization for Standardization (ISO) in the ISO 9000:2000 standard [17]: "quality is the ability of a set of features of a product, system or process to meet certain customer requirements, consumers and other stakeholders"

A definition of greater acceptance regarding to QoE concept, stated by ITU-T on P.10 Recommendation is: "QoE is the general acceptance of an application or service, as perceived subjectively by the end user." From this definition we can see two points: (1) Includes full performance of the end-to-end system. And, (2) can be influenced by user expectations and context.

In recent years there have been some criticisms about the definition of QoE. The term "acceptability" included as a basis for the QoE is not the most appropriated as stated by Moller [18] and other scientists; because “acceptance is the result of a decision which is partially based on Quality of Experience". If a more service-oriented view is considered, arguing that quality is based on the comparison of perceptions with expectations. Aspects of expectation have been addressed in a more comprehensive manner in the context of market research, considering the role of the person as a customer. Perception can refer to both the perception during the meeting with the service, and the concept of QoS related to a particular company in terms of satisfaction or dissatisfaction of customers.

Another author defines QoE as "the characteristics of sensations, perceptions and views of people about a particular service or product; these characteristics can be good, fair or bad” [5]. Also, one should emphasize that to determine the user’s QoE, the perceived quality of multimedia signal needs to be complemented by other criteria related to sensory processing, the human cognitive process and psychological approaches.

2.1 Quality of Experience and Quality of Service

The difference between the concepts of QoS and QoE can be shown in Figure 1, in which the QoS must be ensured by the network providers, by managing some parameters such as delay, jitter, bandwidth and packet loss rate; these parameters must be measured and controlled to guaranty a better service. While the QoS is more related to technical aspects, the QoE cares about the customer service; therefore, the QoS can be part of the QoE definition. Also, other features need to be considered, such as the user preference and service costs.

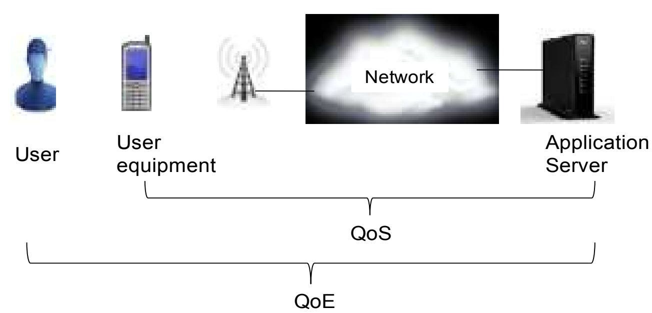

Figure 1. Differences between QoE and QoS in a mobile communication.

Human aspects, such as user preference and service preference, are related with the user’s QoE. Studies consider [19] the user preference for video content, in which the user makes a more critical assessment, giving a low quality score, depending on his or her preference regarding the video content.

2.2 Quality of Experience and User Experience

QoE and User Experience (UX) are new research topics and they have become popular in recent decades. They are also related to development studies and the validation of products and services. In 2010, the ISO introduced its new standard ISO 9241-210 [20], which included the following definition of the term User Experience: "The perceptions and responses of a person resulting from the use of a product, service or system". This standard highlights four UX features:

- UX presents temporal and dynamic aspects; therefore, UX changes through the time;
- UX is dependent on the context in which each situation and experience lived by someone, and this experience can be unique;
- UX is considered subjective and individual;
- The focus of UX is on enjoyment to use a new technology, instead of highlighting problems and difficulties when dealing with a computer.

Studies show that usability methods are not sufficient to measure UX and consequently, new methods of measurement have been studied [21]. Also, aspects like qualities and emotions are not in the scope of Human-computer interaction (HCI).
While QoE has its origin in the Telecommunication area, the UX has its origin in the study of HCI. The origin and evolution of QoE has always focused on industry, in order to avoid users' frustration when they are using certain systems. Also, the approaches of economic aspects and customer loyalty are related to the QoE. This economic dimension is less prominent in the literature regarding to UX. In this context, it is concluded that QoE is much closer to a global customer experience than UX [22].
Another difference between UX and QoE is that UX is more focused on the user. Roto et al. [23] emphasizes that the UX is not related to technology. QoE is considered mostly dependent on QoS [24]. The QoE is focused on application, having a more practical view. Over the past years, several more holistic conceptual frameworks begin to be presented in the literature, for instance, taxonomy focused on QoE [25], the Gr@sp- QoE-framework [26] and the approach of Quality of Experience centered in the user [27].
The QoE is often represented in terms of the index named Mean Opinion Score (MOS). The MOS scale was already used before the QoE become a research topic [28].

2.3 Quality of Experience and its application areas

For example, the IPTV is a service that has several researches that includes QoE [29]-[31]. The area of Cellular Networks also needs to concern about the user's QoE; studies as [32] propose a new Voice Quality Indicator and [33] evaluates voice quality index for mobile systems. Considering all these services, the perceived QoE becomes an indispensable study. There are many application areas that consider the QoE, among them:

- Telecommunication services: covering a large variety of multimedia communications and tradictional fixed and mobile networks.
- Assistive technology: wich goal is design and develop assistive and rehabilitative devices for people with disabilities to improve their QoE.
- Cloud Computing [34]: this new technology must be transparent for the user, the capacity to share, transfer and collaborate with the cloud must maximize the user's QoE;
- Multimedia learning [35]: the use of multimedia contents must be easy and natural with cognitive resources, facilitating the access to information.
- Games: the quality of game software affect directly the user's QoE.
- Sensory Experience [36]: QoE must be multi-dimensional and multi-sensorial, users need to have the impression to be part of the multimedia asset. For this can be used sensory effects (light, wind, vibration) and specific devices and effects (air vaporizer, motion chairs).

2.4 Influence Factors of Quality of Experience

The task of determination of users' QoE in telecommunication services is very complex because the number of Influence Factors (IFs). These factors can be classified in the following categories, human, system and context [37]. Human IFs are regarding to the user's subjectivity, such as, users' preferences [19] and the characteristics of the Human Visual System (HVS). System IFs are related to the characteristics of end-users' devices such as display screen or speaker, processing and energy capacity; and also the characteristics of the network transmission [38]. Context IFs are concerned with the space and time in which a service is used, considering social and economic aspects; for instance, the cost of a phone call, or the characteristics of the physical environment in which a user watches a video. It is worth noting that all of the IFs are interrelated, and each one adding some degradation on the global user's QoE.

3 AUDIO AND VIDEO QUALITY ASSESSEMENT METHODS

In this section, the most relevant methodologies to assess the quality of audio and video signals are briefly described.

3.1 Quality assessment of Audio Signals

Audio quality assessment methods can be classified in two main categories, subjective methods and objective methods. Subjective methods are based on the user's evaluation of the content and objective methods are based on algorithms that contain technical parameters related to network performance or content characteristics.

Subjective methods are also classified in presence and remote tests, which in turn are separated into utilitarian and analytical methods.

Moreover, objective methods can be classified as intrusive, non-intrusive and parametric methods. Intrusive methods require a reference for the evaluation; thus, an original file is used; it is necessary to compare the audio that has arrived to its destination to the original one to evaluate the quality of an audio transmission. One of the most accepted intrusive method is the ITU-T P.862 recommendation [35], most known as Perceptual Evaluation of Speech Quality (PESQ), because its high correlation with subjective tests.

The P.862 algorithm measures the voice distortion effects and the noise in the speech quality. Basically, their algorithm compares an original signal X(t) with a degraded signal Y(t) that is the result of passing X(t) through a communications system [39]. In the first step of PESQ the delays between original input and degraded output are computed and an alignment algorithm is performed. Later, the following stages are implemented, a level alignment to a calibrated listening level, a time-frequency mapping, frequency warping, and compressive loudness scaling.

ITU-T P.862 recommendation only evaluates the effects of one-way speech distortion and noise on speech; then, delays, echo, loudness loss and other impairments related to two-way conversations are not reflected in the PESQ scores. The scenarios for which PESQ had demonstrated acceptable performance are: speech input levels to a codec, transmission channel errors, packet loss and packet loss concealment, bit rates if a codec is a multi-rate codec, transcoding, environmental noise at the transmission side, varying delay in listening only tests, an different techniques of coding, such as waveform codecs, code-excited linear prediction (CELP), adaptive multi-rate (AMR), among others.

On the other hand, non-intrusive methods are based only on the degraded audio file; this method does not require a reference file.

It is important to note that non-intrusive methods have a lesser correlation with subjective test in relation to objective methods. However, non-intrusive methods are recommended for real time quality evaluation, since the only information needed is the audio itself. This characteristic is very important when considered that in most online streaming applications, the original audio is not available.

The most popular nonintrusive objective method is ITU-T P.563 which predicts the speech quality of a degraded signal without a given reference speech signal [40].

The parametric method uses physical measures of the system, including the network, codecs and acoustic parameters. The E-Model is a parametric method that is standardized as ITU-T G.107 Recommendation [41]. This metric tries to predict the audio quality by analyzing some parameters of the network transmission. Although, these metrics are considered state-of-the-art standard, some studies [42]-[44] suggest different algorithms to improve the voice quality assessment metrics.

The subjective tests methods for audio quality assessment are used in tests conducted under laboratory conditions, in which the instructions are explained by a supervisor to assessors. Assessors listen to different audio files and grant an adjective score using different scales. The most popular scale is the five-point Mean Opinion Score (MOS) scale described in the Absolute Category Rating (ACR) method, which is introduced Table 2.

Table 2. Absolute Category Rating (ACR)

Score	Estimated Quality
1	Bad
2	Poor
3	Fair
4	Good
5	Excellent

3.2 Quality assessment of Video Signals

In general, the video quality assessment methodologies can be classified in two groups, subjective and objective methods.

Nowadays, the most popular subjective tests methodologies for video quality assessment are stated in ITU recommendations ITU-R BT-500 [45] and ITU-T P.910 [46]. These subjective tests are conducted in a laboratory environment with special requirements concerning lighting and acoustics conditions. In order to conduct the subjective tests, a supervisor explains the test instructions to the assessors. Then, they score each video assessed using a MOS scale, such as, the scale presented in Table 2 or another quality scale.

Several objective methodologies for video quality assessment are focused on determining the quality of human visual perception considering videos with spatial impairments [47]-[48] or temporal interruptions [49]-[52]. In the last 6 years, some researches [53]-[59] investigate the impact of video resolution changes on the users' QoE.

Hence, many communication service providers measure the user satisfaction to find a manner to improve their services, and other solutions were developed in order to improve existing image and video quality metrics [60]-[63].

Also, measure the quality in audio and video multimedia services is very important, and some other studies [64]-[69] are dedicated to describe the impairments, quality models, components, and metrics. In the next section, additional video quality metrics are treated.

4 FIRST CASE STUDY: QOE IN VIDEO STREAMMING SERVICE

This case study is based on a previous work [6] and its goal is to show how a new metric that quantifies the user's QoE in a video streaming session is modeled. Hence, our motivation to present this case study is show the development process of a video quality metric, considering the different steps involved, such as, the identification of key degradation factors in this specific service, initial subjective tests, mathematical model definition, technical implementation, and finally, the validation tests.

It is worth noting that the proposed metric can be applied in many realistic applications, such as, adaptive video streaming, in which the video coding characteristics depends on the video quality assessed in conjunction with other parameters as network capacity at end user device.

Nowadays, most of video streaming services are running over HyperText Transfer Protocol (HTTP), which uses Transmission Control Protocol (TCP). In order to minimize the network congestion effects, TCP implemented different congestion control mechanisms [70]-[72]. When TCP detects packet losses in the network, the number of transmitted IP packets decreases, and if this new rate is smaller than the playback rate, the player takes all the buffer information and then

enters into a rebuffering process. In this rebuffering time period, no information is displayed and this negatively affects the user's QoE.
A customized player was implemented to extract information regarding the player buffer states during a video streaming. These states are the application layer parameters and they indicate: the number of pauses and their frequency, mean pause length and temporal location.
In subjective tests of video quality, the human perception on the quality of the tested material is quantified by a score and the global quality of the service is evaluated according these results. Results of subjective tests are very important, because product improvements are based on users' requirements [73].
It is important to note that objective metrics such as: Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM) [47], Video Quality Metric (VQM) [48] and algorithms based on Region of Interest [74] or visual attentions maps [75], [76] are not indicated for video streaming running over TCP, because they do not take into account the characteristics of degradations on the temporal domain. Some solutions based on application parameters consider the temporal degradations, specifically the number and duration of the temporal interruptions [49],[50]. In this case study, subjective test results, from experimental tests, are related with the application layer parameters and, as a result, the metric named Video streaming Quality Metrics (VsQM) was established. This approach considered the temporal location of each temporal interruption or pause. Therefore, QoE does not only depend on the number of pauses and their mean period of time, as stated in [49]-[59]. For video streaming service over TCP, the temporal location of each pause must be considered.
Moreover, the proposed metric is used in a useful scenario, in which a feedback mechanism sends the quality metric score from end user device to the video server. This scenario can be used for different purposes such as monitoring, reports or as input of a Rate Determination Algorithm (RDA) to improve the user satisfaction or the network performance.

4.1 Subjective Test Methods for assessing Video Quality

Nowadays, subjective test methodologies do not consider the effects of temporal interruptions, such as the pauses. Especially, in the video streaming service over HTTP/TCP, degradations do not happen in the spatial domain, because TCP guarantees the packet delivery.
In general, the video impairments can appear in the temporal or spatial domain. In Figure 2 the temporal and spatial impairments are presented using TCP and UDP as transport protocol, respectively. It can be observed that the user's QoE is affected in different ways.

Figure 2. Effects of transport protocols in the video quality: (a) Impairment using TCP. (b) Impairment using UDP in a packet loss scenario.

4.2 Limitation of Current Subjective Test Methods for assessing Video Streaming Quality

As stated before, most of the subjective test methods are described in the ITU recommendations: ITU-R BT-500 [45] and ITU-T P.910 [46]. Furthermore, others works compare these subjective methods [77]-[81].
ITU-R BT-500 described the following methods: Double Stimulus Impairment Scale (DSIS), Double Stimulus Continuous Quality Scale (DSCQS), Single Stimulus Continuous Quality Evaluation (SSCQE) and Simultaneous Double Stimulus for Continuous Evaluation (SDSCE).
In the DSIS method the test sequences are presented in pairs: the first stimulus presented in each pair is always the source reference, while the second stimulus is the impairment video [82].
DSCQS method requires the evaluation of two test videos. One of each pair is unimpaired while the other video might or might not contain impairment; but the assessors do not know which video is the reference. Also, the position of the reference picture is changed in pseudo random order [82].

The SSCQE methodology is recommended for longer video sequences. The original video is not used as reference to reproduce viewing conditions that are similar to real situations [45]. The assessors give a score at each certain period of time during the overall video; therefore, there is not a sole global score for a video sequence test.

SDSCE has been developed taking as reference the SSCQE method, in which the presentation of video sequences, and the rating scale had some variations [45].

The ITU-T P.910 recommendation introduces the following methodologies, Absolute Category Rating (ACR), Absolute Category Rating with Hidden Reference (ACR-H), Degradation Category Rating (DCR) and Pair Comparison method (PC).

The ACR methodology is a category judgment, in which the test sequences are presented one at a time and are rated independently on a category scale. The ACR-H methodology is a variant of ACR, in which the assessor does not know which is the original video sequence [46].

The DCR methodology is characterized because the first stimulus presented in each pair is always the source reference, while the second stimulus is the same source with some impairment [46].

In PC methodology the video sequences for testing are also presented in pairs; however, both of them are representing different impairments [46].

Another method named Assessment Methodology for Video Quality (SAMVIQ) is introduced in [83], [84] and it considers some variants of the ITU-T methodologies. Table 3 presents the main parameters of the previous described methodologies. The parameters used to compare these methods are (1) the video length; (2) the explicit reference, the assessor knows which are the original and impairment video; (3) the hidden reference, the original video is presented but the assessor does not know; (4) simultaneous stimuli, in which two videos are presented at the same time; and (5) continuous quality scale, the assessor score several times during a sole video sequence.

As presented in Table 3, the focus of all these methods, except SSCQE, is to assess the effects of spatial degradation, because they consider a video length around 10 seconds. These methods are more useful to evaluate the performance of a particular encoder. Thus, the effects of degradation caused by pauses cannot be properly evaluated by these methods. Also, it is important to note that, in some cases, the length of pauses is almost equal to the total video length.

Table 3. Parameters of Video Quality Assessment Methods

Methods	Video Length (s)	Explicit Refe-rence	Hidden Refe-Rence	Simulta-neous Stimuli	Continuous Quality Scale
DSIS	10	Yes	No	No	No
DSCQS	10	No	Yes	No	No
SSCQE	300	No	No	No	Yes
SDSCE	10	No	No	Yes	Yes
SAMVIQ	10	Yes	Yes	No	No
ACR	10	No	No	No	No
ACR-HR	10	No	Yes	No	No
DCR	10	Yes	No	No	No
PC	10	No	No	No	No

As stated before, the SSCQE methodology is indicated for longer videos. However, grant a score during a pause would not be reasonable, because there is not visual information. Also, there is not any spatial degradation in the video frames because the streaming service uses the TCP.

4.3 The Proposed Subjective Test Methodology

In subjective tests performed in this case study, the variability of the cognitive processes of assessors is considered. They have different cognitive characteristics, such as, attention, speed in information processing, short-term and long-term memory, prior knowledge about technology, and even preferences of video content.

The main differences of the proposed subjective test methodology in relation to the described in Table 3 are the length of the test video sequences, and the global score given by the assessors at the end of the video test. Therefore, the proposed method is in accordance to the current video streaming services.

The subjective video tests were conducted in a laboratory using the recommendation introduced in Table 4 [6]. These recommendations try to establish a realistic environment of the video streaming service.

The results demonstrated that a metric for assessing video quality in streaming service have to consider the temporal location of each pause in the video. Based on this criterion several test scenarios or impairment videos were built to create pauses at different instants of the video and with certain duration.

Table 4. Considerations for Conducting Subjective Tests

Item	Considerations
1	In order to preserve the temporal effect of pauses video sequences are longer than 10 seconds. Video sequences with two and four minutes were chosen taking as a reference the top videos in the most popular *video-sharing* service.
2	Assessors watch the videos according to their preference, and the times they deemed necessary.
3	Considering the variability of attention of assessors, the instructions were specific. Thus, assessors know that video degradation is only due to the presence of pauses.
4	The assessors had different speeds in processing information, hence no limited time to score is considered. Also, the tests were performed individually.
5	Considering the characteristics of assessors' memory, they could watch the videos as many times as each of them considered necessary.

4.3.2 The Proposed Video Quality Model

Firstly, the concept of a video temporal segmentation is introduced, and the following temporal segments are defined: (a) segment A, initial video segment; (b) segment B, first intermediate segment; (c) segment C, second intermediate segment; and segment D, final video segment.

The proposed metric is named *VsQM* and it was determined by the following parameters: number of pauses, pauses length and weight of the temporal segments.

Figure 3, adapted from [6], helps to understand the proposed metric, at which the video playback time was of T_D seconds. In this scenario, four segments were established, and six pauses of different durations were distributed randomly. The number of segments could be increased, but to calculate the degradation weight of each segment, more test video sequences would be necessary.

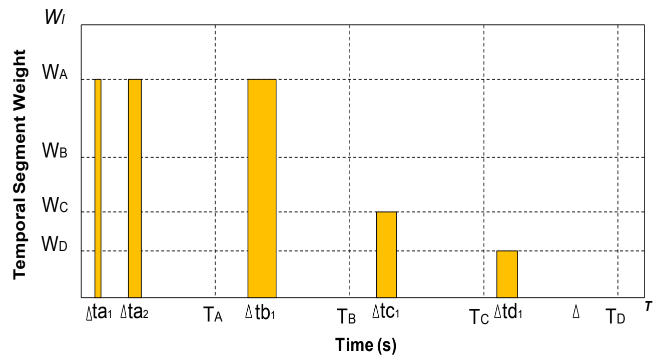

Figure 3. Parameters used to determine the VsQM metric [6].

Then, the VsQM metric is modeled as following (1) [6]:

$$VsQM = \sum_{i=1}^{k} \frac{N_i L_i W_i}{T_i} \qquad (1)$$

Where: N_i is the number of pauses; L_i is the average length of pauses, in seconds, that occurs in the same temporal segment; W_i is a weigh factor which represents the degradation of each segment; T_i is de duration of each segment; k is the number of temporal segments; in this work four segments were considered.

The results of subjective video tests [6], permitted to determine the weights of each segment represented by W_A, W_B, W_C and W_D.

Also, the VsQM values were mapped to a 5-point scale using an exponential function. VsQM at MOS scale is named as $VsQM_{MOS}$, which is described in [6], and it is presented in (2).

$$VsQM_{MOS} = C \exp\left(-\sum_{i=1}^{k} \frac{N_i L_i W_i}{T_i}\right) \qquad (2)$$

Where: $VsQM_{MOS}$ is the MOS index expressed in a 5-point scale, C is a constant for scaling purposes, and the other variables are the same that were described in (1).

Some studies [85], [86] states that the exponential function is the most correlated with subjective test results.

In this case study were performed 20 different test scenarios. The result of each test scenario is represented by a MOS index, for example, the following equation corresponds to scenario 1 ($VsQM_{MOS-1}$):

$$Ln(VsQM_{MOS-1}) = Ln(C) - \frac{W_A N_A L_A}{T_A} - \frac{W_B N_B L_B}{T_B} - \frac{W_C N_C L_C}{T_C} - \frac{W_D N_D L_D}{T_D} \quad (3)$$

Considering the 20 scenarios and (3), an over determined linear system with 2 variables and 20 equations was obtained. To solve this equation system, the least squared method, specifically the pseudo-inverse, was used. Where C is a constant and W_X is the weight of temporal segment "X" to be determined. Also, for each scenario, the variables MOS_X, N_X, T_X and L_X are known. This equation linear system is represented by:

$$\begin{bmatrix} 1 & t_{1,2} & \cdots & t_{1,5} \\ 1 & t_{2,2} & \cdots & t_{2,5} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & t_{20,2} & \cdots & t_{20,5} \end{bmatrix} \begin{bmatrix} Ln(C) \\ W_A \\ W_B \\ W_c \\ W_D \end{bmatrix} = \begin{bmatrix} Ln(VsQM_{MOS-1}) \\ Ln(VsQM_{MOS-2}) \\ \\ \\ Ln(VsQM_{MOS-20}) \end{bmatrix} \quad (4)$$

In which: $t_{1,2}$ to $t_{1,5}$ represent the first scenario; $t_{2,2}$ to $t_{2,5}$ represent the second scenario and so on.

4.3.3 Test Scenario and Application

A video server, a video client and a network emulator were used in the test scenario. The network emulator is used to insert network impairments. The HTTP/TCP protocols were used.
Different impairments, such as, reduction of bandwidth and packet were considered. These impairments are responsible of pauses with different lengths.
In order to monitor the buffer behavior, a customized player that captures all the events related to buffering and playing status was used and it is described as follows [6].

4.3.3.1 Customized Player

The buffer states allow measure the following parameters: (a) number of pauses; (b) length of each pause, which corresponds to the duration of rebuffering state; (c) frequency of pauses; and (d) temporal location of each pause.

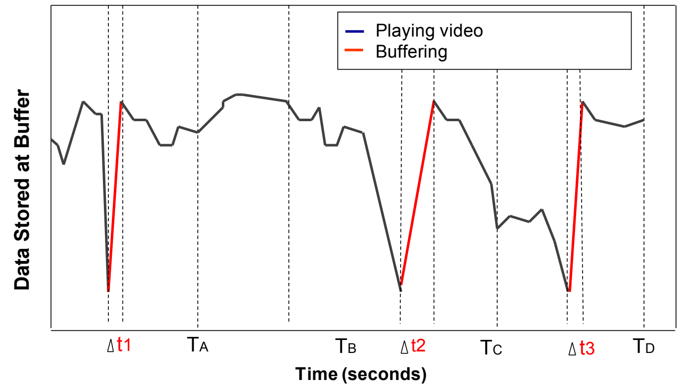

Figure 4. Events traced on the customized player ("R": Rebuffering status and "P": Playing status).

Subjective test results shows the relevance of the temporal location parameter; for instances, five pauses at the beginning of the video affects the video quality in a different way in relation to five pauses at end of the video.

4.3.3.2 Video Data Set

As stated before, twenty different impairment models were built for each video content type considered. The content types used were sport, news and documental; thus, in total 60 impairment video were created following one of the 20 impairment models; five of these impairment models (S1, S2, S3. S4 and S5) are depicted in Figure 5, which is adapted from [6].
The main characteristics of the three original videos are: video and audio format followed H.264/ACC standard, spatial resolution of 640x360, temporal resolution of 30 fps and video length of 240 seconds. Figure 6 presents a snapshot of the content information of each original video.

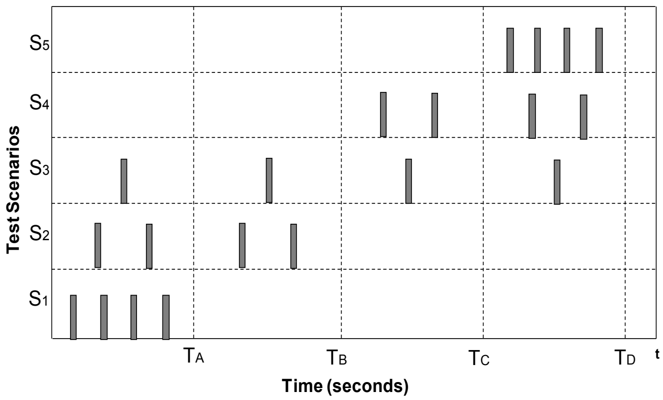

Figure 5. Test Scenarios used as Reference.

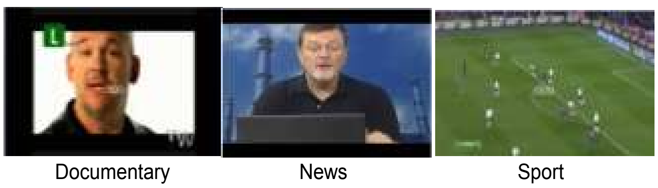

Figure 6. Snapshots of the three video content types used as test material.

As can be observed in Figure 6, the snapshots of the documentary and sport video content types present three circles at the center of each figure. These circles were used to represent a temporal interruption or pauses. In total, 60 test sequences were created, each one with different number and temporal distribution of pauses. These videos were stored in the personal computers used in the experimental tests.

4.3.3.3 Service Application Scenario

Figure 7, adapted from [6], introduces an useful scenario, in which, VsQM metric was sent from the end user's device to the service provider using a feedback mechanism.

In this work, VsQM value is obtained automatically in a pre-defined time period. Depending on this period, it is possible to use VsQM as input of an RDA; thus, the number of users utilizing the service is increased.

For non-real-time applications, the video quality metric can be used to prepare reports or to perform operations and maintain tasks. The feedback mechanism was implemented using a socket interface [87].

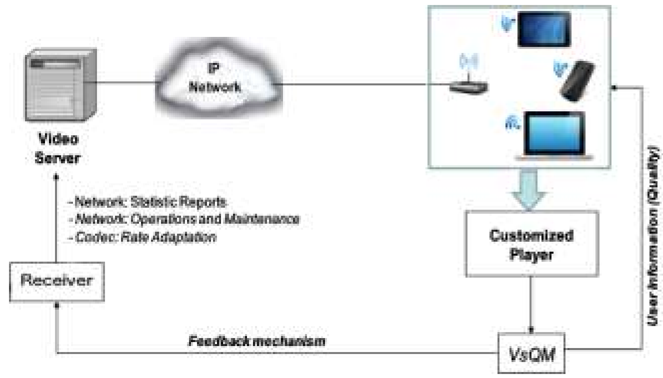

Figure 7. Application Scenario with Feedback Mechanism for the VsQM metric.

4.3.4 Results and Final Considerations

In the subjective tests, 96 evaluators participated, who reported to have no vision problems. Each video had at least fifteen MOS scores. The tests were performed in the same laboratory environment.

To analyze the results, the average MOS value of the three video content types for the same scenario was considered. The values of the constant C and temporal segments weights were obtained: WA, WB, WC and WD. Figure 8 shows the weigh factor values [6].

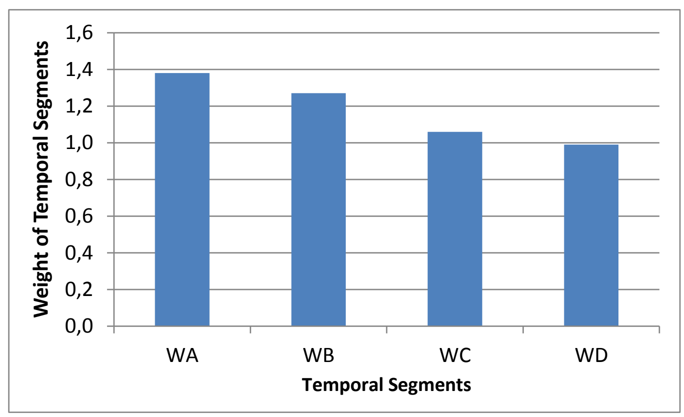

Figure 8. Weight of Temporal Segments: W_A, W_B, W_C and W_D .

It is important to note that pauses at the beginning of the video, have a higher negative effect on the user QoE.

Figure 9 shows the relation between the proposed metric [6] and the subjective test results. The exponential model is very confident, because the maximum error obtained was 0.013 at 5-point MOS scale and a Pearson Correlation Coefficient of 0.96.

Hence, the $VsQM_{MOS}$ metric can be used in real video applications, such as, Dynamic Adaptive Streaming over HTTP (DASH) [88], in which the video resolution transmitted depends on the network capacity at end user, but these algorithms can be improved with other parameters, such as video quality index [89]-[92].

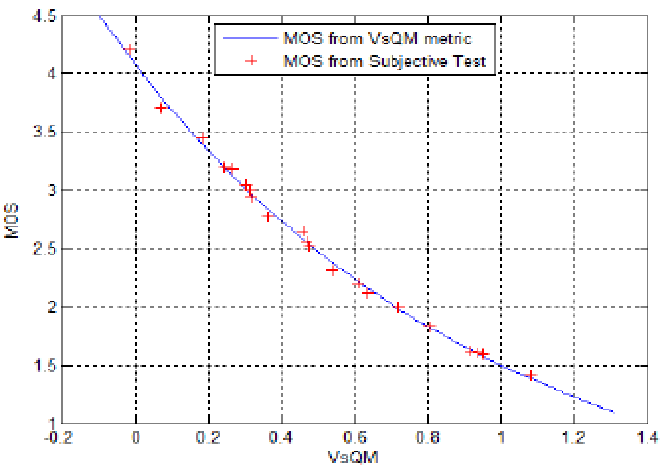

Figure 9. Relation between Subjective MOS results and VsQMMOS estimation.

Subjective test results of video quality showed the relevance of considering the temporal location parameter in a video quality metric. It was demonstrated that pauses at the beginning of the video have a higher negative effect on user's QoE in relation to the intermediate and final parts of the video. Furthermore, the concept of the effect of recent memory is not applicable in the streaming video service.

The VsQM was determined considering the approach of temporal location parameter, and that also considers the number of pauses in a specific video segment, and the duration of each pause. Results show how the MOS index values, calculated from $VsQM_{MOS}$, change if the temporal locations of pauses are considered or not. Furthermore, a function to map the proposed metric values into a 5-point MOS scale was showed.

5. SECOND CASE STUDY: QOE ASSESSMENT OF VOICE CALL SERVICE IN A CELLULAR NETWORKS

This case study is based on the following previous works [8], [9] and it intends to demonstrate the relevance of considering the voice quality index to improve the users' QoE. In a voice call service, the users' QoE depends on different factors [98], mainly the perceived voice quality at end user device, and also the service rate.

Nowadays, in cellular networks there are several services, such as, video conference, location based services, social network and others applications. However, the greatest amount of economic gains of cellular operators corresponds to the voice services. Although, mobile operators are supervised by national regulatory agencies of telecommunications services, the quality of voice services do not have an acceptable quality in all the geographic areas where the cellular operator provides the service. The operation tasks, such as, drive tests, generate high costs for the cellular operator and also these tasks are not enough to discover network coverage problems.

This problem occurs mainly because the parameters of quality indicators - KPI (Key Parameter Indicators) used by regulatory agencies do not reflect the real user's satisfaction. These KPI parameters are not related with the voice signal quality. Specifically, in Brazil, the National Agency of Telecommunications (ANATEL) supervises the performance of mobile operators with 12 indicators [93], and none of them consider the user's QoE. Thus, there is not a KPI that monitors the voice quality after the voice call has been initialized.

In this case study, a network topology to assess voice quality is presented. This network topology is based on the 3GPP project called Minimization Drive Test (MDT) [94]. On the other hand, there are studies [95], [96] that present solutions for monitoring the quality of voice calls, in which network parameters are collected to detect coverage problems. However, these solutions are difficult to be implemented in commercial cellular networks.

The objective of this case study is to demonstrate the importance of including the MOS index [97] in the 3GPP - MDT project. Because, the RF parameters, such as signal strength reception, co-channel interference (C/I), and others, not always are correlated with the voice quality. Thus, MDT not only will be focused to discover coverage problems; additionally, MDT will be able to monitor the voice call quality.

5.1 Parameter Indicators used for assessing the Service Quality of Cellular Networks in Brazil

According to ANATEL reports [99], the number of prepaid users represents about 77.5% of the total cellular users of cellular networks in Brazil and the main service used for prepaid subscribers is the voice call service.
The quality control of cellular networks operators in Brazil is supervised by ANATEL. Indicators are important to evaluate system performance [100]. The ANATEL resolution number 335 of April, 2003 [93] established the definitions, methods and frequency of collection of Personal Mobile Services (PMS) quality indicators. Table 5 introduces the indicators and quality targets for PMS.

Table 5. Key Parameter Indicators of PMS In Brazil.

Index	Description	Target Value
PMS 1	Rate of complaints	1%
PMS 2	Rate of coverage and congestion	4%
PMS 3	Rate of call completion by call centers	98%
PMS 4	Attendance by telephone / electronic service	95%
PMS 5	Rate of completed calls	67%
PMS 6	Rate of call set-up	95%
PMS 7	Rate of call drops	2%
PMS 8	Rate of user response	95%
PMS 9	Rate of response to requests for information	95%
PMS 10	Rate of personal service to the user	95%
PMS 11	Rate of Assistance to the user accounts	5%
PMS 12	Rate of failures recovery	95%

As can be seen from Table 5, none of the PMS indicators correspond to the voice signal quality. These PMS are more related to the service assistance of call center of cellular operators. Also, some PMS treat the cases when the phone call was not established due congestion or coverage problems.

5.2 Overview of the Minimization Drive Tests Solution

To improve the cellular network performance, field engineers perform tasks called drive tests. Thus, radio frequency (RF) parameters are collected to discover some coverage holes or weak coverage areas. It is important to note that the drive tests cannot be performed in all the coverage areas, because some areas are access restricted. Also, drive test tasks are expensive in both time and money.
In this context, MDT solution deals with the two problems mentioned above, because user equipment's (UE) from the real subscribers are used to collect RF parameters. As a consequence, the costs regarding to drive tests are reduced considerably, and the network measurements can be performed in all different places of the network coverage area in short periods of time.

5.2.1 Main characteristics of MDT solution

The main characteristics and functionalities of the MDT are [39]:

- There are two MDT modes to capture the network parameters: (i) The logged MDT, where the UE captures the RF parameters, then they are stored for a certain period of time before the data is sent to the MDT server. This MDT mode is performed when the UE is in idle state. (ii) Immediate MDT is referred when the UE captures and reports immediately the RF parameters to the MDT server. This mode of MDT is performed when the UE is in active state.
- Collection and report of network parameters by the UE. The measurement logs captured for the UE consist of multiple events logged in different timestamps.
- Network operators can choose specific geographical regions to perform the MDT measurements.

- The UE geographical location is important to determine the regions with a weak coverage, this information depends of the UE capability.
- Timestamp in the measurement logs, which need to be correlated with every event of the parameters captured.
- The network operator can select some UE based on their capabilities.

5.2.2 MDT architecture

Figure 10, adapted from [5], depicts the MDT network architecture defined in the 3GPP Rel. 10. A brief description of the network nodes in this figure is presented in following.

- Home Location Register (HLR) / Home Subscriber Server (HSS) are the data base of the subscriber profiles in a cellular network.
- Signal Transfer Point (STP), which is a signaling point that is responsible only for the routing functions within an SS7 network.
- Mobile Switching Center (MSC) - Serving GPRS Support Node (SGSN) / Mobility Management Entity (MME). The MSC is a telephone exchange that makes the connection between mobile users within the same or different mobile and fixed networks. The SGSN is a main component of the GPRS network, which handles all packet switched data within the network.
- Radio Access Network (RAN) / Radio Network Controller (RNC). The RAN is the base station controller in LTE networks.
- MDT Server is the node in charge of collects all network parameters from de UE.
- Operations, Administration, and Maintenance (OAM).

Also, a target geographic area with a possible weak coverage can be analyzed.

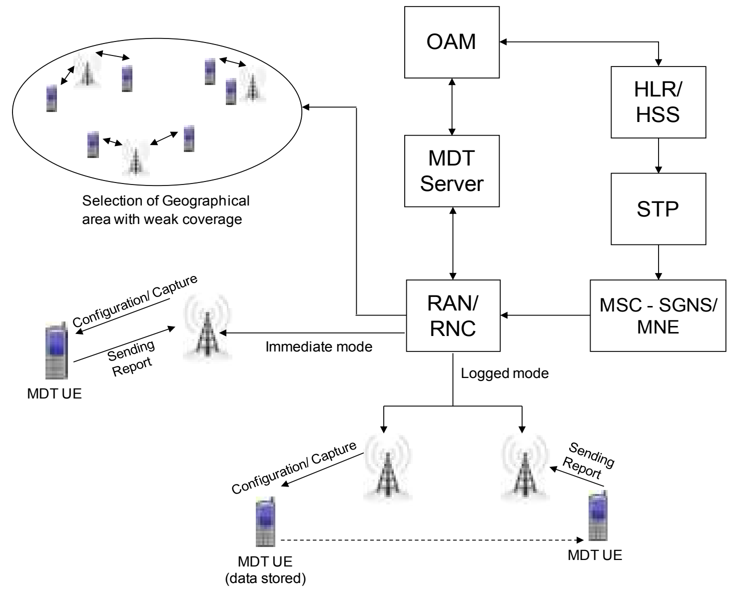

Figure 10. Network Architecture of the MDT solution.

5.2.3 Proposed Network Solution Architecture and Methodology to Assess Voice Quality

In order to analysis and determine the voice quality index, the E6474A (Wireless Measurement Software) and VQT (Voice Quality Test) tools were used. These tools use the ITU-T Recommendation P.862 [39]. A Personal Computer (PC) with the VQT tools installed is connected to an audio card and an UE. The server sends and receives the audio files over a cellular network under test.

The mobile phone originates a phone call to the quality server, which process estimation of voice quality is started using P.862 algorithm [39].

It is worth noting that the P.862 algorithm is considered as standardized state-of-the-art.

Figure 11 shows the test scenario used for capturing the MOS scores in different coverage areas of a commercial cellular network. Additionally, the RF parameters were also measured and collected for analysis purposes.

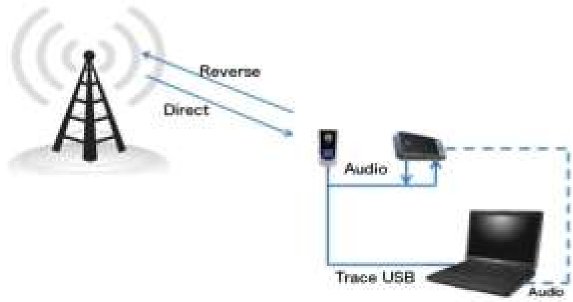

Figure 11. Representation of the test scenario for Voice Quality Assessment in a Cellular Network.

Figure 12, adapted from [9], shows simplified network architecture of the proposed system, based on [94], in which is included the MOS index.

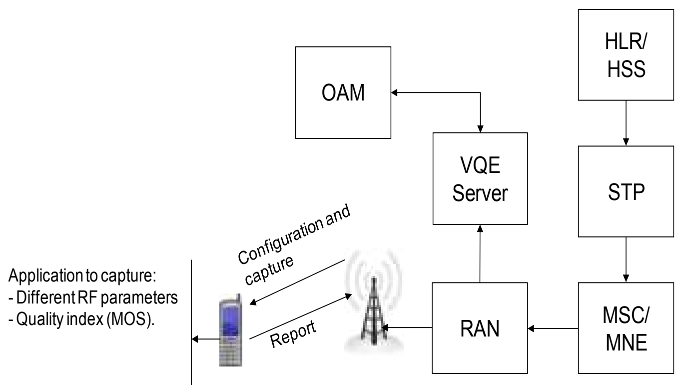

Figure 12. Simplified Network Architecture of MDT including the MOS index.

It is worth mentioning that is necessary to install an application in the cellular phone. This application runs a voice quality algorithm to determine an MOS index. For real-time applications, this algorithm needs to be performed without any voice reference. ITU-T Recommendation P.563 [40] is a non-intrusive objective quality metric, which accomplishes these requirements.

6. CONCLUSIONS

The application of QoE concept is important to everyone who works in the development of product and services, including researchers, administrators and network operators, service and content providers and product manufacturers. Nowadays, the research about QoE are increasing, because it is applicable in several areas such as electrical, electronic, computer, telecommunication engineering, computer science, information technology, professional in the areas of ergonomic and usability, among others. Hence, a considerable number of new researches are expected in the next years, considering the increasing number of works in the last 10 years.

The assessment of the users' QoE is very complex because the QoE concept is based in different influence factors, which are classified into three categories: human, system and context. The first one is especially hard to be evaluated, because it is related to the human subjectivity and each person is different to another. In this context, the subjective tests are really important to determine the user's satisfaction, because during these tests can be identified and determined the main impairment factors that are degrading the user's QoE. Based on the results of subjective tests and considering the key parameters of the service evaluated, a mathematical model can be defined establishing a relation between these parameters and the subjective scores.

Thus, in the first case study, the experimental results indicated that the main impairment factors were the number, duration and the frequency of the pause happened during a video streaming session. Using these impairment factors, or parameters, a mathematical model was defined which output represents the user's QoE index. In the second case study, we presented how a voice quality index can be used to improve the global user's QoE in the voice telephone service.

REFERENCES

1. Chitra, K. and Senkumar, M. R.: Hidden Markov model based lightpath establishment technique for improving QoS in optical WDM networks, In Proc. of 2nd International Conference on Current Trends in Engineering and Technology, pp. 53-62, Cairo, Egypt (2014).
2. Oddershede, A. and Carrasco, R.: Methodology to evaluate and improve the QoS ICT networks in the healthcare service, In Proc. of International Symposium on Communication Systems Networks and Digital Signal Processing, pp. 871-875, Newcastle, United Kingdom (2010).

3. Gorlatch, S., Humernbrum, T. and Glinka, F.: Improving QoS in real-time internet applications: from best-effort to Software-Defined Networks, In Proc. of International Conference on Computing, Networking and Communications, Honolulu, HI, U.S., pp. 189-193(2014).
4. Seno, L., Valenzano, A. and Zunino, C.: A dynamic bandwidth reassignment technique for improving QoS in EDF-based industrial wireless networks, In Proc. of 13th IEEE International Conference on Industrial Informatics, pp. 892-899, Cambridge, United Kingdom (2015).
5. Zapater, M. N. and Bressan, G.: A Proposed Approach for Quality of Experience Assurance of IPTV, In Proc. of First International Conference on the Digital Society, pp. 25-25, Guadeloupe (2007).
6. Rodriguez, D., Abrahão, J., Begazo, D., Lopes, R. and Bressan, G.: Quality metric to assess video streaming service over TCP considering temporal location of pauses," IEEE Trans. on Consumer Electron., vol. 58, no. 3, pp. 985-992 (2012).
7. Rodriguez, D., Pivaro, G. and Sousa, J. : Apparatus and method for evaluating voice quality in a mobile network, US. Patent number US 9,078,143 B2 by US Patent and Trademark Office (2015).
8. Rodriguez, D., Lopes, R. and Bressan, G.: A billing system model for voice call service in cellular networks based on voice quality, in Proc. of IEEE International Symposium on Consumer Electronics, Hsinchu Taiwan, pp. 89-90 (2013).
9. Rodriguez, D. and Bressan, G.: Improving the minimization drive tests using voice quality index, in Proc. of International Workshop on Telecommunications, Minas Gerais, pp. 10-13 (2013).
10. Cowan, N.: On short and long auditory stores, Psychol Buletin, vol. 96, no.2, pp.341–370 (1984).
11. Coltheart, M.: Iconic memory and visible persistence, Percept Psychophysics, vol. 27, no. 3, pp. 183–228 (1980).
12. Baddeley, A.: Human memory—theory and practice, Taylor & Francis: Psychology Press, East Sussex (1997).
13. Baddeley, A.: Working memory: looking back and looking forward, Nat. Rev. Neurosci., vol. 4, no. 10, pp. 829–839 (2003).
14. Jackson, F.: Epiphenomenal qualia, Philos. Quaterly, vol.32, no.127, pp. 127–136 (1982).
15. Martens, H. and Martens, M.: Multivariate analysis of quality, Wiley, Chichester, England, pp. 445 (2001).
16. Reeves, C. and Bednar, D.: Defining quality: alternatives and implications, Acad. Manage Rev., vol.19, no. 3, pp. 419–445 (1994).
17. International Organization for Standardization: ISO 9000:2000, Quality management systems: fundamentals and vocabulary (1999).
18. Möller, S.: Quality engineering, Springer, London (2010).
19. Rodriguez, D. Z., Rosa, R. L. and Bressan, G. : Video quality assessment in video streaming services considering user preference for video content, IEEE Trans. on Consumer Electron., vol. 60, no. 3, pp. 436-444 (2014).
20. ISO 9241-210: Ergonomics of human system interaction-part 210: human-centered design for interactive systems, International organization for standardization (ISO), Geneva (2010).
21. Bargas-Avila J. and Hornbæk, H. : Foci and blind spots in user experience research, Interactions, vol.19, no. 6, pp. 24–27 (2012).
22. Lai-Chong, E., Roto, V., Hassenzahl, M., Vermeeren, A. and Kort, J. : Understanding, scoping and defining user experience: a survey approach, in Proc. of the International Conference on Human Factors in Computing Systems, NY, U.S., pp 719–728 (2009).
23. Roto, V., Lai-Chong, E., Vermeeren, A. and Hoonhout, J. : User experience : bringing clarity to the concept of user experience," White Paper (2011).
24. Callet, P. Le, Möller, S., Perkis, A.: Qualinet white paper on definitions of quality of experience, In Proc. Europ. net. on quality of experience in multimedia systems and services, White Paper, Lausanne (2012).
25. Möller, S., Engelbrecht, K.-P., Kühnel, C., Wechsung, I., Weiss, B.: A taxonomy of quality of service and quality of experience of multimodal human-machine interaction, in Proc. of the International Workshop on Quality of Multimedia Experience, San Diego, CA, pp. 7–12 (2009).
26. Geerts, D., Moor, K. De, Ketyko, I., Jacobs, A.,. Bergh, J. V, Joseph, W., Martens, L., Marez, L. De.: Linking an integrated framework with appropriate methods for measuring QoE, in Proc. of Quality of Multimedia Experience, Trondheim, pp 158–163 (2010).
27. Jumisko-Pyykkö, S.: User-centered quality of experience and its evaluation methods for mobile television, PhD. dissertation, Tampere University of Technology, Tampere (2011).
28. Rothauser, E.H., Chapman, W.D., Guttman, N., Nordby, K.S., Silbiger, H.R., Urbanek, G.E., Weinstock, M.: IEEE recommended practice for speech quality measurements, IEEE Trans. Audio Electroacoust., vol. 17, no.3, pp.225–246 (1969).
29. Rodriguez, D. and Bressan, G.: Video Quality Assessments on Digital TV and Video Streaming services using Objective Metrics, IEEE Latin America Transactions, vol. 10, no.1, pp. 1184-1189 (2012).
30. Rosa, R. L., Rodríguez, D. Z., Souza, V. A. and Bressan, G.: Recommendation system based on user profile extracted from an IMS network with emphasis on social network and digital TV, in Proc. of the Latin America Networking Conference, New York, US, pp. 40-47 (2011).
31. Rodriguez, D. and Bressan, G.: Video Quality Assessment on Digital TV and Video Streaming services using Objective Metrics, in. Proc. International

Information and Telecommunication Technologies, Florianopolis, Brazil (2011).
32. Souza, J., Rodriguez, D. Z. and Cavalcante, A.: Proposal of a new voice quality indicator for the Brazilian mobile telecommunications system based on the end user quality experience, In Proc. of National Conference on Telecommunications, Arequipa, Peru (2011).
33. Stuber, G. L.: Principles of Mobile Communication, Kluwer Academic Publishers, 1st ed., Norwell, MA, USA (1996).
34. Hoßfeld, T., Schatz, R., Egger, S., Fiedler, M., Masuch, K. and Lorentzen, C.: Initial delay vs. interruptions: between the devil and the deep blue sea," International Workshop on Quality of Multimedia Experience, pp.1-6, Yarra Valley, VIC (2012).
35. Mayer, R. E.: Multimedia Learning, Cambridge University Press, New York, USA, pp. 320 (2009).
36. C. Timmerer, M. Waltl, B. Rainer, H. Hellwagner, “Assessing the quality of sensory experience for multimedia presentations,” Signal Processing: Image Commun.,vol. 27, no. 8, pp. 909-916, Sept. 2012.
37. Reiter, U., Brunnström, K., Moor, K. De, Larabi, M.-C., Pereira, M., Pinheiro, A., You, J. and Zgank, A.: Factors influencing quality of experience, in Quality of Experience, Eds. S. Moller and A. Raake, London: Springer, pp. 55-72 (2014).
38. Atzori, L., Floris, A., Ginesu, G. and Giusto, D.: Quality perception when streaming video on tablet devices, J. Vis. Commun. Image R., vol. 25, no. 3, pp. 586-595 (2014).
39. ITU-T Rec. P.862: Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, ITU-T, Geneva, Switzerland (2005).
40. ITU-T. Rec. P.563: Single-ended method for objective speech quality assessment in narrow-band telephony applications, ITU-T, Geneva, Switzerland (2004).
41. ITU-T. Rec. G. 107: The E-model: a computational model for use in transmission planning, ITU-T, Geneva, Switzerland (2000).
42. Hines, A., Skoglund, J., Kokaram, A. and Harte, N.: Visqol: the virtual speech quality objective listener, In Proc. of International Workshop on Acoustic Signal Enhancement, pp. 1-4, Aachen, Germany (2012).
43. Malfait, L., Berger J. and Kastner, M.: P.563&8212; The ITU-T Standard for Single-Ended Speech Quality Assessment, IEEE Transactions on Audio, Speech, and Language Processing, vol. 14, no. 6, pp. 1924-1934 (2006).
44. Thiede, T., Treurniet, W. C., Bitto, R. , Schmidmer, C., Sporer, T., Beerends, J. G. and Colomes, C. : PEAQ - The ITU Standard for Objective Measurement of Perceived Audio Quality, Journal of the Audio Engineering Society 48(1-2), 3–29 (2000).
45. ITU-R Rec. BT.500: Methodology for the subjective assessment of the quality of television pictures, ITU-T, Geneva, Switzerland (2012).
46. ITU-T Rec. P.910: Subjective video quality assessment methods for multimedia applications, ITU-T, Geneva, Switzerland (2008).
47. Wang, Z., Bovik, A. C. and Sheikh, H.: Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing, vol. 13, no. 4 (2004).
48. Pinson, M. and Wolf, S.: A new standardized method for objectively measuring video quality, IEEE Trans. on Broadcasting, vol. 50, no. 3, pp. 312–322 (2004).
49. Porter, T. and Peng, X. H.: An objective approach to measuring video playback quality in loss networks using TCP, IEEE Communications Letters, vol. 15, no. 1 (2011).
50. Mok, R., Chan, E. and Chang, R.: Measuring the quality of experience of HTTP video streaming, in Proc. International Symposium on Integrated Net. Management, Dublin, Ireland, pp. 485-492 (2011).
51. Xue, Y., Erkin, B. and Wang, Y. : A novel no-reference video quality metric for evaluating temporal jerkiness due to frame freezing, IEEE Trans. on Multimedia, vol. 17, no. 1, pp. 134-139 (2015).
52. Khan, A., Sun, L., Fajardo, J., Liberal, F. and Ifeachor, E. : Where's the music? Comparing the QoE impact of temporal impairments between music and video streaming, in Proc. International Workshop on Quality Multimedia Experience, Klagenfurt, Austria, pp. 64-69(2013).
53. Hsio, Y., Chen, C., Lee, J. and Chu, Y. : Designing and implementing a scalable video-streaming system using an adaptive control scheme, IEEE Trans. Cons. Electron., vol. 58, no. 4, pp. 1314-1322 (2012).
54. Mok, R., Luo, X., Chan, E. and Chang, R. : QDASH: A QoE-aware DASH system, in Proc. Multimedia System Conference, N. Carolina, US, pp. 11-22 (2012).
55. Chen, C., Choi, L., Veciana, G. De, Caramanis, C., Heath, R. and Bovik, A.: A dynamic system model of time-varying subjective quality of video streams over HTTP, in Proc. IEEE Acoustics, Speech and Signal Processing Conference, Vancouver, Canada, pp. 3602-3606 (2013).
56. Oyman, O. and Singh, S.: Quality of experience for HTTP adaptive streaming services, IEEE Communications Magazine., vol. 50, no. 4, pp. 20-27 (2012).
57. Essaili, A., Schroeder, D., Staehle, Shehada, D. M., Kellerer, W. and Steinbach, E. :Quality-of-experience driven adaptive HTTP media delivery, in Proc. of IEEE International Conference on Communications, Budapest, Hungary, pp. 2480–2485 (2013).
58. Ni, P., Eg, R., Eichhorn, A., Griwodz, C. and Halvorsen, P. :Spatial flicker effect in video scaling, in Proc. of International Workshop on Quality Multimedia Experience, Mechelen, Belgium, pp. 55-60 (2011).

59. Rodriguez, D., Wang, Z. , Rosa, R. and Bressan, G.: The impact of video-quality-level switching on user quality of experience in dynamic adaptive streaming over HTTP, EURASIP Journal on Wireless Communications and Networking, vol. 216, no. 1, pp. 1-15 (2014).
60. Lin, W. and Kuo, C.-C. J.: Perceptual visual quality metrics: A survey, Journal of Visual Communication and Image Representation, vol. 22, no. 4, pp. 297–312 (2011).
61. Moorthy, A. K. and Bovik, A. C.: Visual quality assessment algorithms: what does the future hold? Multimedia Tools and Applications, vol. 51, no. 2, pp. 675–696 (2011).
62. Chikkerur, S., Sundaram, V., Reisslein, M. and Karam, L. J.: Objective video quality assessment methods: A classification, review, and performance comparison, IEEE Transactions on Broadcasting, vol. 57, no. 2, pp. 165–182 (2011).
63. Chandler, D. M.: Seven challenges in image quality assessment: past, present, and future research, ISRN Signal Processing, vol. 2013, pp.1-53 (2013).
64. Hands, D. S.: A Basic Multimedia Quality Model, IEEE Transactions on Multimedia, vol. 6, no. 6, pp. 806–816 (2004).
65. Garcia, M. N., Schleicher, R. and Raake, A.: Impairment-factor-based audiovisual quality model for IPTV: Influence of video resolution, degradation type, and content type, EURASIP Journal on Image and Video Processing, pp. 1–14 (2011).
66. Yamagishi, K. and Gao, S.: Light-weight audiovisual quality assessment of mobile video: ITU-T Rec. P.1201.1, In Proc. of IEEE 15th International Workshop on Multimedia Signal Processing, Pula, Italy, pp. 464-469 (2013).
67. Pinson, M. H., Ingram, W. and Webster, A. : Audiovisual Quality Components, IEEE Signal Processing Magazine, vol. 28, no. 6, pp. 60-67 (2011).
68. Martinez, H. B. and Farias, C.: Full-reference audio-visual video quality metric, Journal of Electronic Imaging", v. 23, p. 061108 (2014).
69. Jones, C. and Atkinson, D. J.: Development of opinion-based audiovisual quality models for desktop video-teleconferencing, In Proc. of Sixth International Workshop on Quality of Service, Napa, CA, pp. 196-203 (1998).
70. Hiroki, O., Hisamatsu, H. and Noborio, H. : Design and evaluation of hybrid congestion control mechanism for video streaming, in Proc. of the IEEE International Conference on Computer and Information Technology, pp. 585-590, Pafos, Cyprus (2011).
71. Hisamatsu, H., Hasegawa, G. and Murata, M.: Non bandwidth-intrusive video streaming over TCP, in. Proc. of the International Conference on Information Technology, pp. 78-83, Las Vegas, USA (2011).
72. Cai, L., Shen, X., Pan, J. and Mark, J.: Performance analysis of TCP friendly AIMD algorithms for multimedia applications, IEEE Trans. on Networking, vol. 7, no.2, pp. 339-355 (2005).
73. Park, H-J. and Har, D-H.: Subjective image quality assessment based on objective image quality measurement factors, IEEE Trans. Consumer Electron., vol. 57, no. 3, pp. 1176-1184 (2011).
74. Kwon, H., Han, H., Lee, S., Choi, W. and Kang, B. : New video enhancement preprocessor using the region-of-interest for the videoconferencing, IEEE Trans. Consumer Electron., vol. 56, no. 4, pp. 2644-2651 (2010).
75. You, J. Perkis, A., Gabbouj, M. and Hannuksela, M. M.: Perceptual quality assessment based on visual attention analysis, in Proc. of ACM Int. Conf. Multimedia, pp. 561–564, Beijing, China (2009).
76. Noorthy, A. K. and Bovik, A. C.: Visual importance pooling for image quality assessment, IEEE J. Select. Topics Signal Processing, vol. 3, no. 2, pp. 193–201 (2009).
77. Chikkerur, S., Sundaram, V., Reisslein M. and Karam, L. J.: Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison, IEEE Transactions on Broadcasting, vol. 57, no. 2, pp. 165-182 (2011).
78. Pinson, M. H. and Wolf, S.:Comparing subjective video quality testing methodologies, SPIE Video Communications and Image Processing, pp. 573–582 (2003).
79. Janowski L. and Pinson, M. H.: Subject bias: Introducing a theoretical user model, in Proc. of Sixth International Workshop on Quality of Multimedia Experience, pp. 251–256, Singapore (2014).
80. Streijl, R. C., Winkler, S. and Hands, D. S.: Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives, Multimedia Systems, vol 22. No2, pp. 213–227 (2016).
81. Dawson, W. E. and Brinker, R.: Validation of ratio scales of opinion by multimodality matching, Perception & Psychophysics, vol. 9, no. 5, pp. 413–417 (1971).
82. Tominaga, T., Hayashi, T., Okamoto, J. and Takahashi, A.: Performance comparisons of subjective quality assessment methods for mobile video, In Proc. Second International Workshop on Quality of Multimedia Experience (QoMEX), Trondheim, Norway, pp. 82-87 (2010).
83. Blin, J.: New quality evaluation method suited to multimedia context: Samviq, in Proc. of the International Workshop on Video Processing and Quality Metrics, Arizona, U.S., pp. 1-6 (2006).
84. Redi, J., Liu, H., Alers, H., Zunino, R. and Heynderickx, I.: Comparing subjective image quality measurement methods for the creation of public databases, in Proc. of SPIE-IS&T Electronic Imaging, SPIE vol. 1, no. 7529, California, USA, pp. 1-11 (2010).
85. Hosfeld, T., Biedermann, S., Shatz, R., Platzer, A.:The memory effect and its implications on Web QoE modeling, in *Proc. of 23rd International Teletraffic*

Congress (ITC), San Francisco,U.S., pp. 103–110, (2011).

86. Aroussi, S., Bouabana-Tebibel, T., Mellouk, A.: Empirical QoE/QoS correlation model based on multiple parameters for VoD flows, in *Proc. of Global Communications Conference*, pp. 1963–1968, California, U.S. (2012).
87. Rodríguez, D. and Arjona, M.: VoIP quality improvement with a rate-determination algorithm, in Proc. of International Workshop on Telecommunications, Sao Paulo, Brazil, pp. 30-34 (2009).
88. ISO/IEC IS 23009-1: Information Technology – Dynamic adaptive streaming over HTTP (DASH) (2012).
89. Chen, C., Choi, L., Veciana, G. de., Caramanis, C., Heath, R., and Bovik, A., "Modeling the time-varying subjective quality of HTTP video streams with rate adaptations," IEEE Transactions on Image Processing, vol. 23, no.5, pp. 2206-2221 (2014).
90. Chen C., Zhu, X., Veciana, G. de., Bovik, A. and Heath, R.: Adaptive video transmission with subjective quality constraints, in Proc. of IEEE International Conference on Image Processing, Paris, France, pp. 2477-2481 (2014).
91. Singhal, C., De, S., Trestian, R. and Muntean, G-M.: Optimization of user-experience and energy-efficiency in wireless multimedia broadcast, IEEE Transactions on Mobile Computing, vol. 13, no. 7, pp. 1522-1535 (2014).
92. Rodriguez, D., Rosa, R. L, Alfaia, E., Abrahão J. and Bressan, G.: Video Quality Metric for Streaming Service Using DASH Standard, IEEE Trans. on Broadcasting., vol. 62, no. 3, pp. 1-12 (2016).
93. ANATEL Resolution No 335: Regulation of personal Mobil Service Indicators - SMP, Brazil (2003).
94. 3GPP TS 37.320 version 10.4.0 Release 10: Radio measurement collection for Minimization of Drive Tests (MDT) (2012).
95. Diethorn J.: Purposeful receive-path audio degradation for providing feedback about transmit- path signal quality, US PATENT number USA 7,542,761, by Avaya Technology. (2009).
96. Sarkar: Voice quality on a communication link based on customer feedback, US PATENT number US 7,542,761, by ATT. (2009).
97. ITU-T Rec. P.800: Methods for subjective determination of transmission quality, ITU-T, Geneva, Switzerland (1996).
98. Herrero, R.: Encapsulation of Real Time Communications over Restrictive Access Networks, International Journal of Digital Information and Wireless Communications, vol. 6, no. 3, pp. 173-183 (2016).
99. National Agency of Telecommunication (Anatel) Annual Tech. Report of 2014 (2015).
100. Freitas, V., Uren V., Brewster, C., Gonçalves, A.: Ontology for Performance Measurement Indicators' Comparison, International Journal of Digital Information and Wireless Communications, vol. 6, no. 2, pp. 87-96 (2016).