On Additive and Multiplicative QoS-QoE Models for Multiple QoS Parameters

Tobias Hoßfeld1†, Lea Skorin-Kapov2‡, Poul E. Heegaard3, Martín Varela4, Kuan-Ta Chen5

1 University of Duisburg-Essen, Modeling of Adaptive Systems, Germany
2 University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia
3 NTNU, Norwegian University of Science and Technology, Norway
4 VTT Technical Research Centre of Finland — Communication Systems, Finland
5 Institute of Information Science, Academia Sinica, Taiwan

Abstract

Generic relationships between QoE and QoS have been intensively discussed in literature for single QoS parameters and often found to be logarithmic or exponential. While there are many experimental studies investigating statistically the influence of several parameters on QoE, the generic relationship between them, and how to best model it, have not been discussed so far. For communication networks, however, there is a major interest from different stakeholders to have multi-dimensional QoE models. The contribution of this paper is an analysis of the generic relationship between QoS and QoE for multiple QoS parameters and its implications. We address the question of whether multi-dimensional QoE models for several parameters are additive or multiplicative. In an analytic model and with examples involving HTTP non-adaptive video streaming, we show that a multiplicative model has different properties than the current additive QoE model proposed in ITU-T standards. We want to raise sensitivity in the community on multi-factor QoE models, their properties, and the need for multi-factor studies to confirm the appropriate models.

1. Introduction

The notion of Quality of Experience (QoE) has been widely accepted as a multidimensional concept influenced by a number of system, user, and context factors [1]. From a network or service provider’s point of view, it is important to understand the relationships between QoE and underlying network and application-layer Quality of Service (QoS) parameters, thus providing the input for successful QoE management. As such, QoS parameters represent one of the most business-relevant parameters for network and service providers [2].

In general, QoE can be characterized by a function, which maps the impact of influence factors onto the quality perceived by the end user. Typical mapping functions between QoS parameters (at the network and application level) have been found to be exponential functions because of underlying fundamental logarithmic relationships, which is a result of the Weber-Fechner law, and was successfully applied in [3, 4]. The user’s sensitivity is caused by the perception abilities of the human sensory system (“just noticeable difference”), while the QoS parameter reflects a certain stimulus (e.g., waiting times) which is perceived by the human.

†T. Hoßfeld’s work was supported by the Deutsche Forschungsgemeinschaft (DFG) under Grants HO4770/1-2 (DFG project OekoNet).
‡Lea Skorin-Kapov’s work was supported by the Croatian Science Foundation under project UIP-2014-09-S605 (Q-MANIC).

In contrast, the IQX hypothesis [5, 6] postulates an exponential relationship between QoE and QoS. Thereby, the QoS parameter x quantifies a certain degradation (e.g., stalling as an application-layer QoS parameter in the case of HTTP streaming [7], or packet loss as a network-layer QoS parameter in the case of VoIP [8]). The exponential relationship models a linear relationship between the user’s sensitivity to QoE and the actual level of QoE. This leads to a differential equation which is solved as an exponential function.

Due to the complexity of conducting subjective studies with multiple parameter manipulations, available QoE models to a large extent address the relationships between QoS and QoE focusing on the impact of a single QoS parameter. Taking as an example QoE modeling for HTTP-based video streaming, we note that while a number of proposals exist which model QoE as a function of a single parameter (stalling/re-buffering or initial start-up delay), the question arises how to model the joint impact of multiple QoS parameters on QoE. As a result, the currently available ITU-T Recommendation P.1201 (amd. 2) [8] suggests the use of an additive model whereby degradations resulting from stalling and initial delay are subtracted from a maximum Mean Opinion Score (MOS). The impact of stalling is defined as in [7] following the IQX hypothesis.

This contribution addresses the following question: given empirically derived quality models Q1(x1) and Q2(x2) (each modeling QoE as a function of a single influence factor x1 and x2, respectively), and knowing that Q1 follows the IQX hypothesis, what is the generic structure of a combined model Q(x1, x2)? Is it additive or multiplicative?

In this paper we present a theoretical analysis of a combined generic QoE model which considers multiple QoS parameters, illustrated by a numerical example involving HTTP-based video streaming. In an analytic model and with numerical examples involving HTTP-based video streaming, we show that a multiplicative model has different properties as the current additive QoE model proposed in ITU-T standards. Therefore, it is essential to conduct multi-factor subjective studies in order to confirm that one is better than the other.

2. Background and Related Work

QoE can conceptually be characterized by a function mapping the impact of n influence factors onto the quality perceived by the end user. If we denote influence factors as x, a mathematical expression of the function is as follows: QoE = f(x1, x2, ... , xn).

A general approach to the systematic identification of QoE influence factors is proposed in [9], where QoE factors are cat-
Egoritzed into multidimensional IF spaces and further mapped to multiple perceivable quality dimensions. When multiple factors are considered, multidimensional analysis techniques (e.g., Principal Component Analysis, regression techniques) may be used to identify and analyze their impact on QoE (e.g., \cite{10, 11}). A key challenge is understanding the fundamental relationships between multiple QoE influence factors and QoE itself. In the following sections we give a brief overview of related work addressing both additive and multiplicative models, which in certain cases (such as for video streaming) may result in complementary approaches. We also note that models based on various machine learning approaches and neural networks also represent a common approach to QoE modeling, whereby the underlying relationships remain hidden.

2.1. Additive models

The E-model \cite{12} is a commonly used parametric planning model for predicting expected speech quality, taking into account the combined effect of a wide range of impairments transformed onto a perceptual impairment scale. The underlying principle for handling multiple different types of impairments came from the OPINE model proposed by NTT, assuming quality degradation factors are summed on a psychological scale \cite{13, 14}. In an extension to the E-Model, studies on perceived speech quality in cases of random packet loss and additional network impairments for VoIP again showed impairment additivity, although limited in certain cases (e.g., impairment due to packet loss found to be partially masked by additional line noise) \cite{15}. As pointed out in \cite{15}, from the point of view of quality perception, impairment additivity implies distinguishable perceptual features, also referred to orthogonal quality dimensions in a multidimensional perceptual feature space. Wälterman et al. \cite{10} study quality dimensions related to speech transmission, and further model integral listening quality in terms of a weighted linear combination of the identified dimensions.

For video, bit-stream models are given in the ITU-T P.1202-series for video quality estimates. In \cite{16}, the authors use a log-logistic model to model various uni-type impairments of network video quality, and propose an additive log-logistic model for multiple types of impairments. This work has served as a basis for models given in ITU-T Rec. P.1202.2 \cite{17}, whereby overall video quality is estimated as a weighted linear combination of different degradations in the form of compression artefacts, slicing artefacts, and freezing artefacts. ITU-T Recommendation P.1201 \cite{8} provides models for non-intrusive monitoring of audiovisual quality of IP-based video services such as mobile Web browsing \cite{19}.

2.2. Integrating multiplicative terms

While previously mentioned work has focused on additive models, other studies consider multiplicative models in various service scenarios. Studies addressing audiovisual quality have generally proposed a combination of two dimensions (audio and video qualities) leading to the following integration model: $MOS_{AV} = aMOS_A + bMOS_V + \gamma MOS_A \ast MOS_V + \zeta$, where $MOS_{AV}$ refers to overall audiovisual quality and $MOS_A$ and $MOS_V$ refer to audio and video quality, respectively. As summarized in \cite{20}, subjective tests have shown that the multiplicative term between audio and video qualities, with an additive shift, is generally sufficient to estimate audiovisual quality. This was confirmed in a survey comparing integration models \cite{21}, highlighting also the importance of the MOS ranges of audio and video qualities.

A generalized MOS estimation function is proposed in \cite{22}, defining MOS as a weighted product of all variable-specific MOS’ (referring to the MOS’ of individual influence factors) rather than a weighted sum. The rationale behind this approach is that the model should effectively reflect the situation when one variable’s MOS is very low and cannot be compensated by an improvement of other variables. The authors further propose a deterministic QoE model where QoE may be calculated based on multiple QoS factors, with VoIP quality studies confirming exponential relationships and the potential of combining single variable equations in a multiplicative way \cite{23}.

In \cite{24}, the authors go beyond MOS estimation and propose a Generalized Linear Model to estimate the probabilities of particular QoE levels based on given QoE influence factors. In this case, again multiplicative terms between independent variables are considered.

3. Theoretical Analysis

We consider QoE as a function $Q(x)$, $x = \{x_1, \ldots, x_n\}$ which maps $n$ influence factors $x_1, \ldots, x_n$ on QoE. Depending on the type of influence factors, the range of a given factor may be a positive integer (e.g., number of stalls), a positive real number from a certain range of values (e.g., packet loss ratio from $[0; 1]$), or a real value (e.g., jitter).

If we assume that there is a linear relation between the QoE sensitivity $\frac{\partial Q}{\partial x_i}$ of $x_i$ and the QoE function $Q(x)$, then we can say that $x_i$ follows the IQX hypothesis in $Q(x)$ \cite{5}.

$$\frac{\partial Q(x)}{\partial x_i} \propto Q(x)$$  \hspace{1cm} \text{(IQX hypothesis)}$$

For a single QoS parameter $x$, the IQX yields an exponential functions; $Q(x) = \alpha e^{-\beta x}$ when considering normalized values in the range $Q \in [0; 1]$. The parameter $\alpha$ determines the maximum MOS value of $Q_i$, referred to as a maximum parameter. $\beta$ determines the degree of the slope of the QoE curve, i.e. its sensitivity. $\beta$ is referred to as a sensitivity parameter.

For the sake of simplicity, we only consider two factors $x_1$ and $x_2$ and their corresponding QoE models $Q_1(x_1)$ and $Q_2(x_2)$. We further assume that $Q_1$ follows IQX, let $Q_1(x_1) = \alpha_1 e^{-\beta_1 x_1}$ and $Q_2(x_2)$ be a general QoE function.

All multi-factor QoE models in the following suffer from the same problem: they strongly depend on weighting factors.
that need to be determined by curve fitting based on results from multi-factorial studies. Such studies are exactly what’s missing.

3.1. Additive QoE Model

The additive model assumes the weighted sum of both influence factors with some weighting factors \( w_i \geq 0 \) and \( \sum_i w_i = 1 \).

\[
Q_a(x_1, x_2) = w_1 Q_1(x_1) + w_2 Q_2(x_2) \tag{1}
\]

According to the IQX hypothesis, the sensitivity of QoE depends on the actual QoE value. Therefore, we need to consider the partial derivative w.r.t. \( x_1 \).

\[
\frac{\partial Q_a(x_1, x_2)}{\partial x_1} = -w_1 \beta Q_2(x_1) \neq Q_a(x_1, x_2) \tag{2}
\]

Eq.(2) means that the sensitivity of QoE w.r.t \( x_1 \) does not depend any more on its actual value. Therefore, the additive QoE model destroys the IQX property, although \( Q_1 \) follows an exponential function.

3.2. Multiplicative QoE Model

We consider now a multiplicative QoE model and assume again normalized QoE values, \( Q_i \in [0; 1] \).

\[
Q_m(x_1, x_2) = a_{12} Q_1(x_1) \cdot Q_2(x_2) \tag{3}
\]

The IQX hypothesis is still valid then for the impairment factor \( x_1 \), independent of function \( Q_2 \), as the QoE sensitivity of \( x_1 \) depends on the overall value \( Q_m \).

\[
\frac{\partial Q_m(x_1, x_2)}{\partial x_1} = -\beta a e^{-\beta x_1} Q_2(x_2) = -\beta Q_m(x_1, x_2) \tag{4}
\]

3.3. Linear Regression QoE Model

A generalization of the additive and multiplicative QoE model is a linear regression which includes both terms.

\[
Q_r(x_1, x_2) = Q_a(x_1, x_2) + Q_m(x_1, x_2) \tag{5}
\]

Also in this case, the QoE sensitivity of \( x_1 \) does not depend any more on the actual QoE value.

3.4. Linear Regression of Impairment Factors

When several impairment factors are influencing the same stimuli (or QoE dimension) like stalling, then those impairment factors may be combined with a linear regression model. For example, there may be the same underlying cause for stalling which jointly influences the stalling frequency but also the stalling duration.

\[
x = v_1 x_1 + v_2 x_2 + v_{12} x_1 x_2 \tag{6}
\]

If the QoE model for the parameter \( x \) follows the IQX, then this results in a multiplicative model (see Section 4.1) and all parameters \( x_i \) follow IQX.

3.5. Observation and discussions

The key observation is that if \( Q(x) \) follows the IQX hypothesis, then QoE functions should be combined in a multiplicative way. However, in practice, it is not clear which is the best model \( Q(x) \). Therefore, multi-factor experiments are needed to obtain an appropriate QoE model. The additive or multiplicative models discussed in this section are only two examples with very different properties to emphasize the importance of understanding the underlying mechanisms. A multi-factor QoE model might need to take into account that (i) impairment factors and their corresponding QoE might be correlated (both cross-correlated and longitudinal over time), (ii) it has non-additive or non-multiplicative structure, (iii) it changes over the impairment factor sampling space, e.g., at the extremes it might be max/min function otherwise an additive or multiplicative model. As the number of parameter combinations is growing for each additional factor \( x_i \), sophisticated sampling strategies and parameter selections are required in the subjective studies [25], in order to limit the costs for the subjective studies. The sampling strategies should thereby consider the models to be tested.

4. Case: HTTP Streaming QoE Model

We apply the analysis from the previous section to a simple example involving a QoE model for non-adaptive HTTP streaming. In a Gedankenexperiment, we combine existing QoE models for stalling as well as for initial delay in an additive and multiplicative way. This allows to compare a multiplicative QoE model with the current ITU-T Recommendation advising the use of an additive model.

4.1. Stalling

The QoE of HTTP streaming depends mainly on the actual number of stall events \( N \) and the average length \( L \) of a single stall event [26]. A QoE model combining both key influence factors into a single equation \( Q_S(L, N) \) is provided in [7] and found to follow the IQX hypothesis [5], see Figure 1. The model provides MOS estimates on a 5-point ACR scale. We normalize the model into the domain \([0; 1]\) and obtain

\[
Q_S(N, L) = e^{-\alpha LN - \beta N} \tag{7}
\]

By means of linear regression, the stall pattern \( \alpha LN + \beta N \) is captured as input for the exponential function [7] with \( N \) being the dominant factor; \( LN \) captures thereby the total stall duration. Please note that the stall pattern regression is a complementary approach to the linear regression model (Sec. 3.3) which may allow to combine also different QoE dimensions (like initial delay vs. service interruption due to stalling); the stall pattern regression maps several parameters on the same dimension (which is the stall pattern here).

![Figure 1: Results from the subjective study [26] lead to exponential fitting functions between MOS and number \( N \) of stalls. Each curve depicts a different average duration \( L \) of stall events. Without stalling, maximum QoE is reached.](image-url)
4.2. Stalling and Initial Delay

We consider now in addition the impact of initial delays $D$ on QoE, as derived in [7] with parameters $a$ and $b$ leading to MOS values in $[1; 5]$ which are again normalized in $[0; 1]$: \[Q_1(D) = -a \log_{10}(b \cdot D + c) \] \hfill (8)

However, since no subjective studies on the joint influence on QoE are available, we compose the additive and the multiplicative QoE model based on $Q_1$ and $Q_2$ in order to show the different behavior, as depicted in Figure 2. Figure 3 shows the results for the additive and multiplicative QoE model in comparison with the ITU-T model. It can be seen that the ITU-T model $Q_{ITU}$ (as the simple additive model $Q_a$) leads to counter-intuitive results. Related work [27] shows that initial delays only have limited impact on QoE – in contrast to stalling. With stalling being the dominant factor (e.g. [27]), it seems reasonable that the user’s sensitivity to stalling depends on the actual QoE level. This is however not true for the additive model and the ITU-T model. The ITU-T model leads to very optimistic QoE values in case of low initial delays. The additive model leads to very pessimistic QoE values in case of high initial delays, potentially overestimating the impact of the initial delays. The multiplicative model (and the not shown linear regression model) seems to capture properly the insights from literature - nevertheless, multi-factor user studies need to demonstrate the applicability of the model.

5. Concluding remarks

Our aim with this paper is to draw attention to multi-factor QoE models, their properties, and the need for multi-factor studies. It is not straightforward to combine existing single-parameter QoE models into a multidimensional QoE model. Thus, the observations $Q_1(x_1)$ and $Q_2(x_2)$ do not allow to constitute $Q(x_1, x_2) = f(Q_1(x_1), Q_2(x_2))$. It may be tempting just to assume an additive model or a multiplicative model for $f$. However, the two models lead to strongly different properties.

We provided a concrete example for HTTP video, where conflicting results exist (e.g. ITU-T standard P.1201 (amd. 2) [8] vs. [7, 18]). We showed that the multiplicative model will preserve the IQX hypothesis and the assumption that QoE sensitivity depends on the actual QoE level. The additive model however destroys the IQX property although this was found in a subjective study for $Q_1$ (i.e. in which only stalling factors were varied, but no initial delays). We note, however, that the multifactor QoE model cannot be derived from observations $Q_1$ and $Q_2$. Thus, existing recommendations advocating the use of a combined additive QoE model when considering multiple QoE influence parameters should be reconsidered.

There are also more models of course available, e.g. linear regression $w_1Q_1 + w_2Q_2 + w_{12}Q_1Q_2$, e.g. $f = \min$, e.g. numerical approaches based on machine learning, e.g. integration functions [20], etc. The intention of the paper is to raise awareness to researchers to investigate the underlying principles and generic relationships of QoE modeling. This understanding of generic relationships is important to make progress in multi-factor QoE modeling and gain deeper understanding of the underlying fundamental principles. In order to obtain a deeper understanding of QoE, we need to go beyond black box approaches such as basic regressions, fittings, and machine learning approaches.

In conclusion, subjective tests are clearly required to test which kind of multi-factor QoE model is appropriate for a given service scenario. Due to the growth of the parameter space where each parameter adds one dimension of complexity, it is required in practice to come up with proper sampling approaches (e.g. [28]) and parameter selections (e.g., adaptive approaches in crowdsourcing [25]). This is another relevant direction for future work, as the sampling approaches or adaptive CS may consider the underlying model in order to come to an efficient implementation of the test design.
6. References


