

# A Generic Quantitative Relationship between Quality of Experience and Quality of Service

Markus Fiedler, Blekinge Institute of Technology  
Tobias Hossfeld and Phuoc Tran-Gia, University of Würzburg

## Abstract

Quality of experience ties together user perception, experience, and expectations to application and network performance, typically expressed by quality of service parameters. Quantitative relationships between QoE and QoS are required in order to be able to build effective QoE control mechanisms onto measurable QoS parameters. Against this background, this article proposes a generic formula in which QoE and QoS parameters are connected through an exponential relationship, called IQX hypothesis. The formula relates changes of QoE with respect to QoS to the current level of QoE, is simple to match, and its limit behaviors are straightforward to interpret. It validates the IQX hypothesis for streaming services, where QoE in terms of Mean Opinion Scores is expressed as functions of loss and reordering ratio, the latter of which is caused by jitter. For web surfing as the second application area, matchings provided by the IQX hypothesis are shown to outperform previously published logarithmic functions. We conclude that the IQX hypothesis is a strong candidate to be taken into account when deriving relationships between QoE and QoS parameters.

Information and communications technology (ICT) users find an ever growing set of IP-based applications and access networks from which to choose. With the availability of residential broadband, IP-based digital TV is also offered via Ethernet and asymmetric digital subscriber line (ADSL). Such a multitude of offers makes prices decrease, and competition between service and/or network providers increase. The customer finds itself in a strong position, being able to choose between different competing providers. Given similar pricing schemes, which are a primary decision aid for many users, their subsequent choices are then likely to be influenced by expected and experienced quality (i.e., through personal ratings of the perception and price-worthiness of a service). Consequently, the providers' interest in how users perceive usability, reliability, quality, and price-worthiness has increased. A provider needs to be able to observe and react quickly on quality problems, at best before the customer perceives them and considers churn. Facing this kind of quality competition, the concept of quality of experience (QoE) emerged, combining user perception, experience, and expectations with non-technical and technical parameters such as application- and network-level quality of service (QoS) [1].

We observe a non-uniform view of QoS by different stakeholders. While the International Telecommunication Union (ITU) standards focus on service quality toward the end user, the Internet Engineering Task Force's (IETF's) understanding of QoS relates to the capabilities of the network to provide packet transfer in a better-than-best-effort way. While the ITU view on QoS is user-centric [1], the IETF view on QoS is

network-centric. This raises the question of how network-level QoS measurements and control relate to the user perception of a service. In order to choose appropriate measures to keep user-perceived service quality above an acceptance threshold, a provider needs to know how network-level QoS parameters translate into user-level QoE perception and vice versa.

Obviously, QoE-QoS relationships depend on many aspects and parameters. A typical user-related measure is the mean opinion score (MOS), which can be determined from *subjective* ratings by real users or predicted from *objective* measurements of properties of the delivered *goods* such as audio, video, or files. From the area of usability engineering, reaction time thresholds for user perception are known as follows: 100 ms is roughly the boundary at which a user feels the system reacting instantaneously; less than 1 s keeps the user's thoughts, although a delay is perceived; less than 10 s keeps the user's attention, while exceeding 10 s implies the risk of the user abandoning the activity. The order of magnitude of the latter threshold has been confirmed by studies of user patience regarding web surfing [2].

Yet, no matter the networked application, users expect timely and complete data delivery. Ideally, a networked application should behave as if the enabling network was completely transparent. Streaming applications must maintain fluidity, conversations and interactive applications shall expose short response times, and downloads need to be able to efficiently *stream* large amounts of data in order to minimize waiting times [3]. Basic QoS problems on the network level relate to non- or late delivery, which might entail re-ordering. Thus,

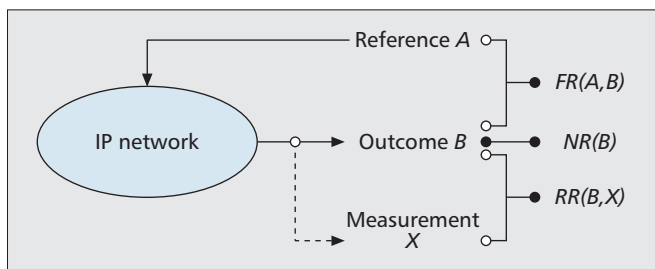


Figure 1. Illustration of the different quality metrics to derive QoE-QoS relationships.

they affect the timely behavior of the application and possibly even the appearance of the content.

Obviously, *generic QoS problems* (e.g., loss, delay, jitter, re-ordering, throughput limitations) imply *generic QoE problems* (e.g., glitches, artifacts, excessive waiting times). Thus, it is of interest to investigate *generic relationships between QoE and QoS*, which is the core contribution of this article. It presents a nice, simple, unified, easy-to-match, and practicable formula expressing an exponential dependency of QoE on QoS. It is thus applicable for online in-service classification of QoE problems based on QoS observations, which is of interest to service providers and network operators. Our results enable QoE control mechanisms that build on QoS monitoring.

The remainder of the article is structured as follows. The next section elaborates on different types of quality metrics relevant to QoE and QoS evaluations. We then present and discuss the qualitative relationship between QoE and QoS. We then quantify the latter into an exponential formula and validate it for three case studies addressing different QoE parameters such as voice quality and user reactions to download times and throughput limitations. In the latter two cases, the original logarithmic matching formulae are outperformed by the proposed exponential matching. The final section concludes and points out directions for future work.

## Quality Comparisons and Classification of Metrics

The derivation of QoE-QoS relationships builds on quality comparisons between:

- The so-called *reference*, by which we mean undistorted content such as image or video, or an undistorted service such as a download activity
- The *outcome* of the transmission in form of a potentially distorted image or video, or a delayed download activity

A distortion of the outcome may impact the quality of the content (e.g., image quality) and/or timing (e.g., fluidity of a video, download times, service activation time). Then the QoE relates to the remaining quality of the outcome after such a distortion.

References play an important role when it comes to rating the quality of the outcome. Evidently, the closer the quality of the outcome comes to that of the reference, the better the QoE. In the optimal case both match, and the network between can be considered transparent. The impact of distortions can, among other methods, be expressed with aid of utility functions [4].

For the sake of quality comparison, there are different measurement methods and observation levels. We can distinguish between:

- A *communication situation*, in which merely the outcome is available
- A *laboratory situation*, in which both reference and outcome are available and can be compared offline with great effort and in great detail

There are two basic measurement options. *Subjective tests* are carried out by a test panel of real users. While many and possibly even diverging views on the quality of the outcome can be taken into account — entailing accurate results and a good understanding of the QoE and its sensitivity — this type of test can be time-consuming and costly, since the tests have to be conducted by a large number of users for statistically relevant results. *Objective tests* are carried out by an algorithm on behalf of a real user, trying to imitate or predict user perception based on key properties of the reference and/or the outcome. Typically, subjective quality tests form the basis for perceptual objective test methods. Objective tests can follow psychophysical approaches and engineering approaches, a detailed description of which is found in [5]. For voice over IP (VoIP), the Perceptual Evaluation of Speech Quality (PESQ) standard objectively evaluates and quantifies voice quality of voice-band speech codecs. It uses a psycho-acoustic and cognitive model to analyze and compare the reference and the outcome. PESQ allows for repeatable and automated measurement processes, yielding statistically significant results [6].

Depending on the object of interest, we can observe content and related network traffic on different levels. Observation on the *application level* implies examination of the payload, which makes it possible to get a detailed picture of the content, and on the timing of reference and outcome. Problems with the latter may arise from network nodes and links and the network stacks in the end systems, as well as from the implementation of the application. Additionally, measurements on the *network level* may be conducted. This means investigation of the flow of packets in terms of completeness, timeliness, and pattern analysis regarding bursty losses or correlated delays. For testing the IQX hypothesis, we observe both levels to derive QoE-QoS relationships.

Metrics can be classified according to the following scheme [5]:

**Full reference (FR) metrics:** Both reference and outcome are available, and allow for detailed subjective and objective comparisons of images, videos, download times (on the application level), packet traces (on the network level), and so on. Concretely, this means extraction, evaluation, and comparison of QoE and QoS parameters on any level in an offline manner, which is most interesting for deriving QoE-QoS relationships. A prominent example of an FR metric is PESQ [6].

**No reference (NR) metrics:** Quality information has to be extracted from the outcome, as no reference is available. This is a typical online situation with focus solely on the resulting quality as perceived by the end user (evaluated through observations and questions) or its representative (an algorithm). In a networking context, NR metrics usually lack the possibility of discerning between quality problems stemming from the very reference and additional disturbance by the network. This is a hindrance to deriving QoE-QoS relationships aiming at capturing the impact of the network. On the other hand, user-acceptance-related QoE parameters building on generalizable experiences such as image quality (exposure, sharpness, contrast, colors) or response times may well serve as a basis for QoE-QoS relationships; see the investigations made in [3, 7].

**Reduced reference (RR) metrics:** For reference and outcome, the same set of parameters are derived and compared. For instance, QoE on the application level can be described by the hybrid image quality metric (HIQM) [5], and QoS on the network level can be represented by throughput variations and losses [4]. Such parameters often have their roots in FR research as a means of summarizing and interpreting results. However, as they represent key QoE and QoS parameters in a very condensed manner, they can be applied in an online in-service scenario by transmitting them between source and sink, and subsequently comparing them in order to find out

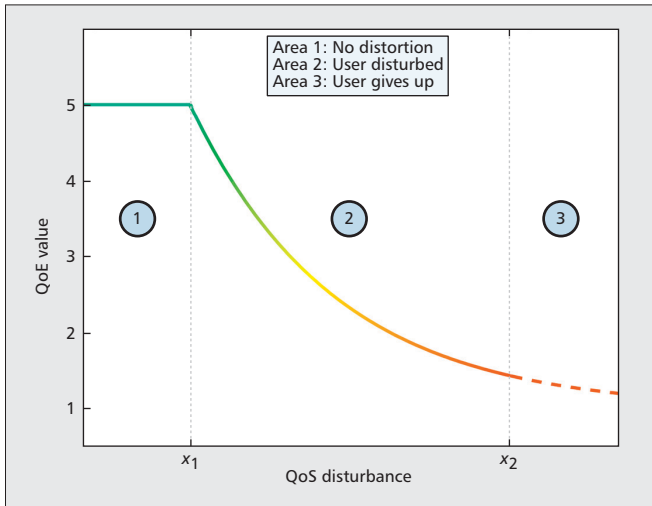


Figure 2. General shape of the mapping curve between QoS and QoE.

about quality problems. Thus, they represent promising candidates to formulate QoE-QoS relationships [4].

Figure 1 illustrates the different FR, NR, and RR quality metrics and their required inputs. For the FR metric, the reference  $A$  as well as the outcome  $B$  are available, allowing to estimate the QoE by  $FR(A, B)$ . For the NR metric, only the outcome  $B$  is available, yielding  $NR(B)$ . For the RR metric, in addition to the outcome  $B$ , the measured parameters  $X$  are available on the receiver side and potentially even communicated from the sender side [4]. Thus, the quality is estimated as  $RR(B, X)$ .

### Qualitative Relationship between QoE and QoS

We now turn our focus onto a qualitative schematic relationship describing the impact of QoS problems on QoE, illustrated in Fig. 2. On the x-axis, the QoS disturbance is denoted, while the y-axis indicates a QoE value (e.g., an opinion score). Although this relationship is basically independent of the type of metric discussed above, we now focus on a situation in which the network accounts for QoE reductions between reference and outcome [4].

The QoE of the outcome of the transmission as a function of QoS disturbance is split in several areas, separated by thresholds  $x_1$  and  $x_2$ .

*Area 1: constant optimal QoE* — For a vanishing QoS disturbance (i.e., in case of a transparent network), the user considers the QoE equivalent to that of the reference. A slight growth of the QoS disturbance may not affect the QoE at all. For instance, small delays and delay variations may be eliminated by a jitter buffer, without the user noticing the additional delay. A similar effect is observed for web surfing delays up to half a second. Even if web pages were delivered faster than what is given by threshold  $x_1$ , to the average user it would not matter.

*Area 2: sinking QoE* — When the QoS disturbance exceeds a certain threshold  $x_1$ , the former quasi-optimal QoE level cannot be maintained anymore. As the QoS disturbance grows, the QoE and thus the user satisfaction sinks. In case of a high QoE, a certain additional QoS disturbance might have a considerable impact on the QoE, while for low QoE, the particular additional QoS disturbance might not be that critical anymore. Consequently, as the QoE sinks, its negative gradient is expected to do so as well.

*Area 3: unacceptable QoE* — As soon as the QoS disturbance reaches another threshold,  $x_2$ , the outcome of the trans-

mission might become unacceptably bad in quality, or the service might stop working because of technical constraints such as timeouts. A user might give up using the service at that point; this is illustrated by the dashed line.

While threshold  $x_1$ , due to its technical nature, represents a sharp threshold that very well may be collocated with the y-axis, threshold  $x_2$  may be user-dependent, an example of which is discussed for the cancellation rate of web browsing users in the next section.

### The Exponential Relationship between QoE and QoS

In this section we present, demonstrate, and discuss a fundamental functional relationship between QoE and QoS parameters, the IQX hypothesis (exponential interdependency of quality of experience and quality of service). It is motivated and presented later. The use of such a QoE-QoS relationship is straightforward; by inserting measured QoS values into the corresponding exponential formula, their impact on QoE can be assessed immediately. Thus, the formulae presented and discussed in the sequel enable QoE threshold surveillance and control.

In the first place, the IQX hypothesis is formulated as described earlier with the QoE parameter representing the level of satisfaction and the QoS parameter reflecting the level of disturbance. Typically, the QoE parameter and user perception decrease when the QoS parameter increases. The IQX hypothesis is tested for two different services, VoIP and web browsing, and the corresponding results are shown later in this section for the impact of loss, jitter, and re-ordering on VoIP, and for the impact of response times on user ratings.

In the second place, the QoE parameter reflects a negative user reaction, while the QoS parameter is associated with the availability of resources. Also, here, the QoE parameter sinks when the QoS parameter increases, but this stands for improved user perception. In this case we test the IQX hypothesis with the cancellation rate as the QoE parameter and throughput as the QoS parameter.

#### The IQX Hypothesis for QoE

At large,  $QoE = \Phi(I_1, I_2, \dots, I_n)$  is a function of  $n$  influence factors  $I_j$ . In this contribution we focus on a single influence factor,  $I = QoS$ , in order to derive the fundamental relationship  $QoE = f(QoS)$ .

In general, the subjective sensibility of the QoE is more pronounced the higher this experienced quality is. If the QoE is very high, a small disturbance will strongly decrease the QoE. On the other hand, if the QoE is already low, a further disturbance is not perceived significantly. This relationship can be motivated considering a restaurant QoE: If we dined in a five-star restaurant, a single spot on the clean white tablecloth would strongly disturb the atmosphere. The same incident would go unnoticed in a simple tavern.

Against this background, we assume that the change of QoE depends on the current level of QoE, given the same amount of change of the QoS value, but with a different sign. Assuming a linear dependence on the QoE level, we arrive at the following differential equation:

$$\frac{\partial QoE}{\partial QoS} \sim -(QoE - \gamma). \quad (1)$$

The solution to this equation is easily found as an exponential function, which expresses the basic relation of the IQX hypothesis:

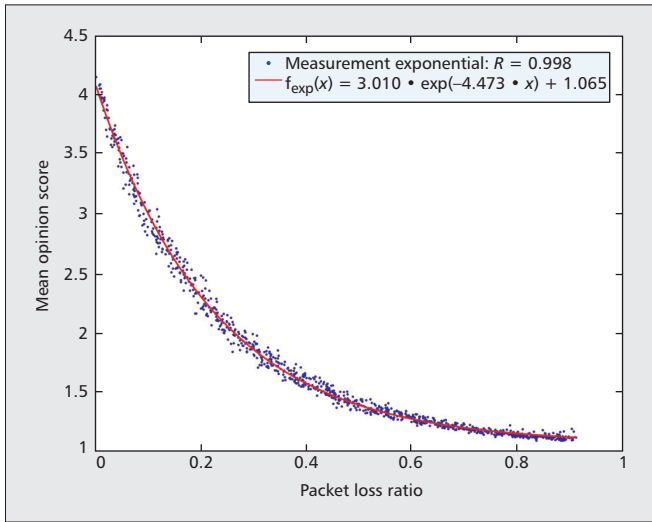


Figure 3. Measurement results and obtained mapping function  $f_{\text{exp}}(p_{\text{loss}})$  between packet loss ratio  $p_{\text{loss}}$  and QoE for the iLBC codec.

$$QoE = \alpha \cdot e^{-\beta \cdot QoS} + \gamma. \quad (2)$$

At this point, it is observed that the logarithmic approaches found in [3, 7] and discussed later imply the differential equation

$$\frac{\partial QoE}{\partial QoS} \sim -\frac{1}{QoS}. \quad (3)$$

In contrast to the IQX hypothesis (Eq. 1), the change of QoE depends on the reciprocal QoS value.

#### Voice Quality Affected by Loss, Jitter, and Reordering

For demonstrating the mapping between QoE and QoS of a VoIP service, we use as an example the free *Internet low bit rate codec* (iLBC) [8] used by Skype. The measurement sessions are detailed in [9]. For emulating loss and jitter in the network, we used Nistnet v. 2.0.12c (<http://www-x.antd.nist.gov/nistnet/>). The potentially disturbed audio stream is captured from the audio application SJPhone v. 160.299 (available at <http://www.sjlabs.com/sjp.html>) into an audio .wav file. This outcome is compared to the undistorted reference audio file using the PESQ method. The resulting PESQ value is mapped into a subjective MOS value according to ITU-T Recommendation ITU-T P.862.1 [6], taking on the values 1 = bad, 2 = poor, 3 = fair, 4 = good, and 5 = excellent. While PESQ is a full-reference method  $FR(A, B)$ , the IQX hypothesis yields a reduced reference metric  $RR(B, X)$ , where  $X$  represents the QoS distortion and is derived from comparing the amount of data received with the amount of data sent, the difference of which indicates loss of data on its way from user A to user B. We vary the packet loss probability from 0 to 90 percent in steps of 0.9 percent.

Figure 3 shows the obtained MOS values dependent on the measured packet loss ratio  $p_{\text{loss}}$  for the conducted experiments and for the applied model  $f_{\text{exp}}(p_{\text{loss}})$ . Each dot represents a single MOS-from-PESQ measurement for a given packet loss probability. For  $p_{\text{loss}} \rightarrow 1$ , the QoE in terms of MOS approaches its minimum of one from above. The model function  $f_{\text{exp}}(p_{\text{loss}})$  is retrieved by means of nonlinear regression. We used the optimization toolbox of Matlab to find an optimal fitting function (i.e., the unknown parameters  $\alpha$ ,  $\beta$ ,  $\gamma$  in Eq. 2) such that the normal error  $E$  is minimized. The normal error is defined as the sum of the residuals  $r_i$  for all  $n$  measurements  $(x_i, y_i)$  with a measured packet loss  $x_i$  and a measured MOS  $y_i$ :

$$E = \sum_{i=1}^n r_i, r_i = f_{\text{exp}}(x_i) - y_i. \quad (4)$$

We obtain the following fit for the chosen iLBC voice codec according to the IQX hypothesis:

$$QoE = 3.010 \cdot e^{-4.473 \cdot p_{\text{loss}}} + 1.065. \quad (5)$$

The goodness of fit for the model function  $f_{\text{exp}}(x)$  can be measured with different metrics. The coefficient of correlation  $R$  between the model function and the measured data or the coefficient of determination  $R^2$  should approach one for a close-to-perfect match. In contrast, the mean squared error

$$MSE = \frac{1}{n} \sum_{i=1}^n r_i^2$$

or the normalized mean squared error  $NMSE = MSE/\text{Var}[y_i]$  should approach zero. In fact, Eq. 5 yields  $R = 0.998$ ,  $R^2 = 0.995$ ,  $MSE = 0.003$ , and  $NMSE = 0.005$ , meaning that all metrics indicate an almost perfect match. Further details are found in [9], where we also show well fitting exponential results for the G.711 voice codec.

It has to be noted that the packet loss is only one impairment factor indicating QoS problems. For a general quantification of the QoE, additional factors such as jitter have to be considered. In the context of this work, jitter is used as a general term to describe the variation in the end-to-end delays of IP packets between sender and receiver, a common effect in packet-switching networks and quite disturbing in real-time communications. Although jitter does not cause a packet to get lost in the network, it is possible that the packet is received too late at the application layer and thus considered lost from the application's perspective. We can thus expect excessive jitter to degrade the QoE in a similar way as real data loss. In order to defeat such problems and guarantee continuous playout of audio transmitted over the network, jitter buffers are used at the receiver. The maximum jitter that can be countered is equal to the buffering delay introduced before starting the playout of the voice stream.

Nistnet is configured with the average and standard deviation of the end-to-end delay as input parameters and delays each individual packet according to a normal distribution with the given parameters. It has to be noted that the generated end-to-end delays of consecutive packets are independent of each other. Hence, it is possible that packets overtake each other, which means that a packet  $q$  sent after packet  $p$  arrives at the receiver before packet  $p$  (i.e.,  $t_q < t_p$ ). We use the *type-p reordered ratio* [10] to quantify the jitter, defined as the percentage of packets in the received stream that are reordered and obtained from comparing the sequence numbers  $X$  at sender and receiver, respectively. Voice codec and speech samples were the same as described above. In our measurements we use an average end-to-end delay of 90 ms and vary its standard deviation  $\sigma$  from 0 to 90 ms. Figure 4 shows the measurement results. Each dot in the diagram represents a single measurement. The  $x$ -value shows the measured type-p reordered ratio and the  $y$ -value the corresponding MOS. We use the same nonlinear regression techniques mentioned earlier to obtain an optimal fitting function  $f_{\text{exp}}(x)$  between the QoE in terms of MOS and the QoS in terms of type-p reordered ratio. The coefficient of correlation  $R = 0.993$  shows a very good match between the measurement data and the applied exponential model. We get similar results when applying different metrics for quantifying the jitter of the end-to-end path. Such metrics are, for example, the standard deviation of the one-way delays, the interpacket delay variation, or more complex reordering metrics like type-p reord-late-time.

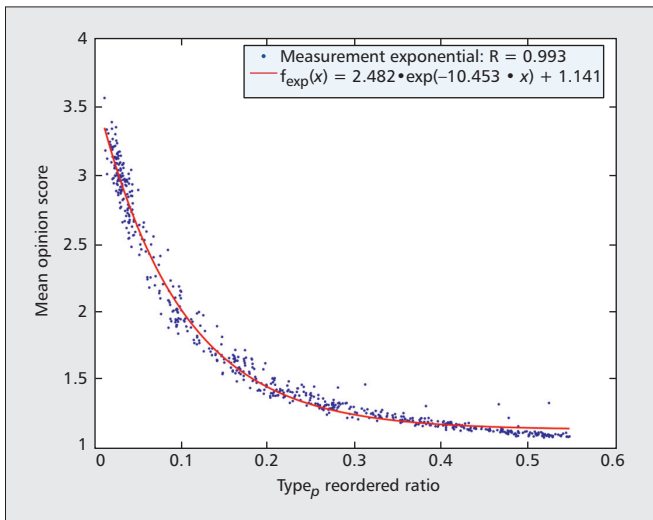


Figure 4. Measurement results for the iLBC codec and the obtained exponential mapping function  $f_{\text{exp}}(x)$  between the QoE and the type- $p$  reordered ratio  $x$  as QoS parameter.

### Mapping of Weighted Session Time to Perceived Web Browsing Quality

The first example for checking the IQX hypothesis for web browsing is based on ITU-T Recommendation G.1030, “Estimating End-to-End Performance in IP Networks for Data Applications” [3]. It applies perceptual models to gauge user satisfaction (i.e., QoE) with the end-to-end performance (i.e., QoS) in terms of response and download times, measured in the network or calculated from the HTTP transaction times. Response and download times in a web session were manipulated, and users were asked to evaluate the perceived quality according to the five-point MOS scale presented before.

In [3] it is stated that the expected maximal session time will dominate the perceived quality and the user’s rating. Therefore, the network context (fast, medium, and slow network) was taken into account through different maximal session times, in particular 6 s, 15 s, and 60 s. Furthermore, weighted session time  $t_w$  was introduced in order to improve the correlation between service time and MOS; details are found in [3].

The web session consisted of three steps, reflecting a typical search-for-information situation involving:

- Requesting a search page
- Typing and submitting a query
- Retrieving the results

The users were asked to type in the same search query in every session. In total, 49 experiments were conducted for each of the three network contexts with varied response and download times. Here, the testing users were distinguished into two separate groups, trained experts and untrained naive users. No references were available (i.e.,  $NR(B)$  was obtained).

Figure 5 shows the measurement results from [3] for the fast network and the entire user set, experts and naive users. Each point in the graph represents a single experiment with the weighted session time on the x-axis and the MOS on the y-axis. In [3] a logarithmic mapping function  $f_{\text{log}}(x)$  is proposed that yields a coefficient of correlation  $R = 0.954$ . Note that the logarithmic model leads to MOS values above 5 for  $t_w < 0.62$  s and to MOS values below 1 for  $t_w > 13.48$  s. This is indicated in Fig. 5 by the dashed line segments.

Testing the IQX hypothesis results in a coefficient of correlation  $R = 0.966$ , which denotes a slight improvement. An advantage of using the exponential curve as the mapping between QoS and QoE is the fact that  $f_{\text{exp}}(x)$  is lower-bounded by  $\gamma = 1.390$  for large weighted session times in contrast to

$f_{\text{log}}(x)$ . However, for very small session times,  $t_w < 0.50$  s, the exponential function also leads to MOS values above 5. This example nicely demonstrates the threshold  $x_1$  of the principal shape of the mapping function between QoE and QoS introduced earlier. An exemplary user is completely satisfied if the session time is around half a second. If the data is delivered faster than that, the user is not able to value this better service quality. Thinking of network provisioning, this opens the possibility to save resources, as it is not necessary to provide better QoS for maintaining the same QoE. This illustrates the potential impact of QoE and the paradigm change in telecommunication networks accompanied by the consideration of QoE instead of QoS.

### Cancellation Rate of Web Browsing Users

The presented measurements here are taken from [7] in which Khirman and Henriksen measured the level of user dissatisfaction with web content delivery quality, which is a no-reference measure  $NR(B)$ . We use these results to check the IQX hypothesis in the context of web browsing in a second scenario with different parameters. As a QoS parameter, the delivery bandwidth is used. The QoE is expressed as the cancellation rate of web requests.

Khirman and Henriksen used a passive network-attached sniffing device that collected packets traveling across a specific network link. Afterward, they applied reverse engineering to the captured packets to get information about the states of TCP connections and extract details of the application layer transactions. The data collector was installed in a commercial ISP network with public Internet access. Eighty percent of the traffic was generated by customers using dial-up modem connections up to 56 kb/s; 20 percent of the traffic was generated by customers using high-speed connections. More details regarding the measurements can be found in [7].

Figure 6 shows the relationship between cancellation rate of HTTP objects and corresponding delivery bandwidths per object. Only objects entailing at least 8 kbytes of data to be transferred were considered. This results in 373,050 object requests of which 22,903 (i.e., 6.1 percent) were cancelled. Note that only low-range delivery bandwidth up to 120 kb/s is considered due to the fact that the majority of users had dial-up connections at that time. Every point in this graph represents the cancellation rate for a bin of 7461 objects with a

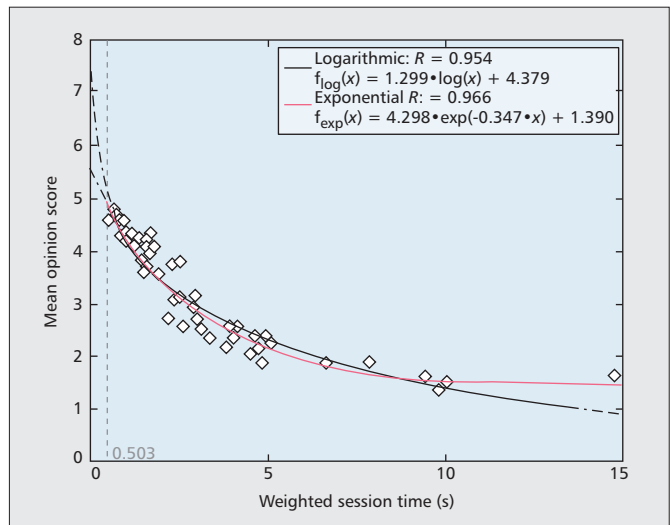


Figure 5. Measurement results for web browsing in a fast network taken from G.1030 [3] and comparison of logarithmic model  $f_{\text{log}}(x)$  and exponential model  $f_{\text{exp}}(x)$ .

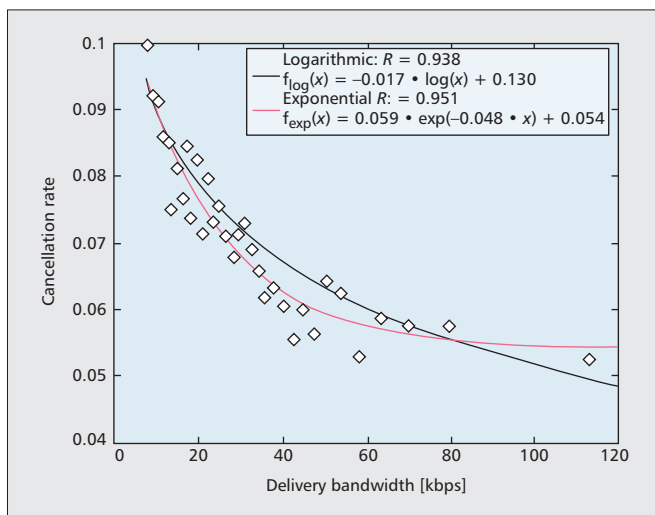


Figure 6. Measurement results for web browsing taken from Khirman and Henriksen [7] and comparison of logarithmic model  $f_{\log}(x)$  and exponential model  $f_{\exp}(x)$ .

similar delivery bandwidth. The  $x$ -value of a point denotes the average delivery bandwidth of these objects; the  $y$ -value represents the cancellation rate, which is the ratio of the number of cancelled objects divided by the total number of requests in this bin. In order to determine if an object is cancelled, the object size advertised by the server and the actual size of the delivered object are compared.

Khirman and Henriksen proposed a logarithmic fitting function  $f_{\log}(x) = -0.017 \log x + 0.130$  for the cancellation rate dependent on the delivery bandwidth  $x$  in [7]. The resulting coefficient of correlation is  $R = 0.938$ . Applying the IQX hypothesis to these measurements leads to a slightly better coefficient of correlation  $R = 0.951$  for the exponential model  $f_{\exp}(x) = 0.059e^{-0.048x} + 0.054$ .

In Fig. 6 we see that the exponential function decreases more strongly in the beginning. At the end, however, the shape of the exponential curve is more flat than for the logarithmic function in this area and slowly approaches the asymptote  $\lim_{x \rightarrow \infty} f_{\exp}(x) = \gamma = 0.054$ . Note that the logarithmic function — in contrast to the exponential one — is not bounded, as  $\lim_{x \rightarrow \infty} f_{\log}(x) = -\infty$ . Altogether, the IQX hypothesis appears to be more appropriate for modeling cancellation rate as the QoE parameter than the original logarithmic formula with respect to delivery bandwidth as the QoS parameter.

## Conclusions and Outlook

In this article we present and evaluate the IQX hypothesis, which is a natural and generic exponential relationship between user-perceived QoE and network-caused QoS. In contrast to formerly presented logarithmic relationships that formulate the sensitivity of QoE as a function of (reciprocal) QoS only, the IQX hypothesis expresses the sensitivity of QoE as a function of QoE itself. Furthermore, it shows superior convergence behavior for extreme values of QoS parameters.

Along the way, we review the common reference models and provide a general view on user perception as a function of disturbances on the application and network levels. We demonstrate the feasibility of the proposed exponential relationship through a couple of case studies, addressing among others voice quality as a function of loss, jitter, and reordering; user ratings as a function of response times (both expressing QoE as a function of QoS disturbance); and cancellation rates of web surfers as a function of access link bandwidth (expressing QoE disturbance as a function of

QoS). The proposed exponential relationship provides approximations of better quality than the original logarithmic approximations.

Future work will address different kinds of streaming and interactive applications and combinations thereof. In this context combinations of different QoS disturbances such as loss and delay will need to be investigated. Our results will enable QoE control mechanisms that build on QoS monitoring.

## Acknowledgments

The sponsorship of this research by the European FP7 network of Excellence “Euro-NF” through the Specific Joint Research Project “QoEWeb” and by the Swedish Knowledge Foundation through the project “QoEMoVi” (d-nr 2006/0208) is thankfully acknowledged. The work has been partly conducted in the project G-Lab, funded by the German Ministry of Education and Research (FKZ 01 BK 0800, G-Lab). The authors alone are responsible for the content of the article.

## References

- [1] ITU-T Rec. P.10/G.100 (incl. Amendment 2), “Vocabulary for Performance and Quality of Service,” 2008.
- [2] A. Bouch, A. Kuchinsky, and N. Bhatti, “Quality Is in the Eye of the Beholder: Meeting User’s Requirements for Internet Quality of Service,” tech. rep. HPL-2000-4, HP Labs, Palo Alto, CA, Jan. 2000.
- [3] ITU-T Rec. G.1030, “Estimating End-to-End Performance in IP Networks for Data Applications,” Nov. 2005.
- [4] T. Hoßfeld *et al.*, “Measurement and Analysis of Skype VoIP Traffic in 3G UMTS Systems,” 4th Int’l. Wksp. Internet Performance, Simulation, Monitoring, Measurement, Salzburg, Austria, Feb. 2006.
- [5] U. Engelke and H.-J. Zepernick, “Perceptual-Based Quality Metrics for Image and Video Services: A Survey,” *NGI ’07*, Trondheim, Norway, May 2007, pp. 190–97.
- [6] ITU-T Rec. P.862.1, “Mapping Function for Transforming P.862 Raw Result Scores to MOS-LQO,” Nov. 2003.
- [7] S. Khirman and P. Henriksen, “Relationship between Quality-of-service and Quality-of-Experience for Public Internet Service,” *3rd Passive Active Measurement Wksp.*, Fort Collins, CO, Mar. 2002.
- [8] S. Andersen *et al.*, “Internet Low Bit Rate Codec (iLBC),” IETF RFC 3951 (Experimental), 2004.
- [9] T. Hoßfeld *et al.*, “Testing the IQX Hypothesis for Exponential Interdependency between QoS and QoE for Voice Codecs iLBC and G.711,” *18th ITC Specialist Seminar Quality of Experience*, Karlskrona, Sweden, May 2008.
- [10] A. Morton *et al.*, “Packet Reordering Metrics,” *RFC 4737 (Proposed Standard)*, Nov. 2006.

## Biographies

MARKEUS FIEDLER (markus.fiedler@bth.se) received his doctoral degree in electrical engineering/ICT from Universität des Saarlandes, Saarbrücken, Germany, in 1998. Since then he has been with Blekinge Institute of Technology, Karlskrona, Sweden, holding a Docent degree in telecommunication systems since 2006. Within the School of Computing, he performs and supervises research on quality of experience, seamless communications, network virtualization, service chains, and networks of the future (NF). He is leading and participating in several national and European projects. He is serving on the Steering Board of the European Network of Excellence Euro-NF and coordinating its specific joint research projects.

TOBIAS HOSSFELD (hossfeld@informatik.uni-wuerzburg.de) studied computer science and mathematics at the University of Würzburg, Germany. He finished his Ph.D. on performance evaluation of future Internet applications and emerging user behavior in 2009. Currently, he is heading the Overlays, P2P, and QoE research group at the Chair of Communication Networks in Würzburg. His main research interests cover network virtualization, social networks, self-organization mechanisms in overlay networks, and P2P systems, with a special focus on mobile environments like 3G and 4G, as well as investigations on QoE for Internet applications like Skype, Joost, and YouTube.

PHUOC TRAN-GIA (trangia@informatik.uni-wuerzburg.de) heads the Department of Distributed Systems within the Institute of Computer Science, University of Würzburg, Germany. His current research areas include architecture and performance analysis of communication systems, and planning and optimization of communication networks. He has been active in several management committees of European research projects. He is currently working with the European Union authorities and the German Ministry of Research and Education on funding strategies and initiatives toward next generation Internet. He is coordinator of the project German-Lab (G-Lab), aiming to foster experimentally driven research to exploit future Internet technologies.