

Report from Dagstuhl Seminar 12181

Quality of Experience: From User Perception to Instrumental Metrics

Edited by

Markus Fiedler¹, Sebastian Möller², and Peter Reichl³

1 Blekinge Institute of Technology – Karlskrona, SE, markus.fiedler@bth.se

2 Deutsche Telekom Laboratories, DE

3 FZ Telekommunikation Wien, AT, reichl@ftw.at

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 12181 “Quality of Experience: From User Perception to Instrumental Metrics”. As follow-up of the Dagstuhl Seminar 09192 “From Quality of Service to Quality of Experience”, it focused on the further development of an agreed definition of the term Quality of Experience (QoE) in collaboration with the COST Action IC1003 “Qualinet”, as well as inventories of possibilities to measure QoE (beyond the usual user polls) and to exploit feedback between users and systems that reflects QoE issues. The report furthermore describes the mode of work throughout the seminar, with focus on personal statements by the participants, results of the group works, and open challenges.

Seminar 01.–04.May, 2012 – www.dagstuhl.de/12181


1998 ACM Subject Classification H.5.1 Multimedia Information Systems, H.5.2 User Interfaces

Keywords and phrases Quality of Experience, usability, content, network monitoring, quality measurement, service pricing

Digital Object Identifier 10.4230/DagRep.2.5.1

1 Executive Summary

*Markus Fiedler,
Sebastian Möller
Peter Reichl*

License  Creative Commons BY-NC-ND 3.0 Unported license
© Markus Fiedler, Sebastian Möller, and Peter Reichl

During the recent years, Quality of Experience (QoE) has established itself as a topic of its own for both industrial and academic research. With its focus on the user in terms of acceptability, delight and performance, it is about to take over the role of Quality of Service as key paradigm for provisioning and managing services and networks. As one of the follow-up activities of the Dagstuhl Seminar 09192 “From Quality of Service to Quality of Experience”, this Dagstuhl Seminar 12181 focused on the relation between quality perception and QoE quantification, which is among the most challenging tasks for bringing together the three essential corner stones, i.e. user, technology, and business. In particular, qualitative user perception needs to be translated into quantitative input to dimensioning and control of networks and services. Further, different kinds of feedback flows (acceptance, usage, cost, quality) need to be taken into account. Considering the multidisciplinary nature of this problem with complementary and potentially controversial views, the seminar worked towards metrics and measurement techniques aimed at improving QoE prediction and control. The outcomes are expected to become visible in the future QoE research agenda and corresponding standardisation efforts.



Except where otherwise noted, content of this report is licensed under a Creative Commons BY-NC-ND 3.0 Unported license

Quality of Experience: From User Perception to Instrumental Metrics, *Dagstuhl Reports*, Vol. 2, Issue 5, pp. 1–25
Editors: Markus Fiedler, Sebastian Möller, and Peter Reichl



Dagstuhl Reports

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

2 Table of Contents

Executive Summary

<i>Markus Fiedler, Sebastian Möller, and Peter Reichl</i>	1
---	---

Introduction

Introduction	
<i>Markus Fiedler</i>	4

Overview of Talks


Combination of multi-source observations at the sub-second scale	
<i>Patrik Arlos</i>	4
On-line Estimation of the Quality of Experience	
<i>Åke Arvidsson</i>	4
A User-Centric Service Modeling approach for User Experience Assessment	
<i>Sergio Beker</i>	5
Are engineers from Mars and users from Venus? QoE measurement as an interdisciplinary process	
<i>Katrien R. De Moor</i>	6
A network based method for Video QoE measurement	
<i>Marcus Eckert</i>	7
Waiting for QoE	
<i>Sebastian Egger</i>	7
Quality of Experience instrumentation: Read the user AND the network	
<i>Markus Fiedler</i>	8
QoE IPTV	
<i>Marie-Neige Garcia</i>	9
User-centric troubleshooting	
<i>Riccardo Guerzoni</i>	9
Long-term QoE: How does the overall quality perception of one user evolve over multiple interactions with IP-based services?	
<i>Dennis Guse</i>	10
QoE Assessment for Web-based systems and IP TV	
<i>Richard John Harris</i>	10
You look blocky, is everything alright?	
<i>Helmut Hlavacs</i>	11
QoE: Measuring the Immeasurable?	
<i>Tobias Hossfeld</i>	11
Energy metrics unique enough for smartphone-based video QoE evaluation?	
<i>Selim Ickin</i>	12
Repeatable Results – The Key to Scientific Accuracy	
<i>Lucjan Janowski</i>	13

Impact of physical layer impairments on higher layer QoS parameters <i>Maria Kihl</i>	13
Towards Total Quality of Experience. A Conceptual View on QoE in a communication ecosystem <i>Khalil Laghari</i>	13
Quality of Experience is not only Quality of User Experience <i>Patrick Le Callet</i>	14
How to come to good QoE instrumentation? <i>Sebastian Möller</i>	14
Qo(E) Vadis? Multi-user, multi-service, multi-information, multi-timescales <i>Alexander Raake</i>	15
On Economics of QoE <i>Peter Reichl</i>	16
Automatic QoE assessment <i>Gerardo Rubino</i>	16
Non-intrusive network-based estimation of QoE <i>Junaid Shaikh</i>	17
From modeling QoE to QoE management: challenges for domain-wide QoE-driven resource allocation <i>Lea Skorin-Kapov</i>	18
A Generic Approach for Understanding QoE <i>Martin Varela</i>	19
Factors Influencing Quality of Experience of Commonly-Used Mobile Applications <i>Katarzyna Wac</i>	20
The influence of contextual factors on quality ratings <i>Ina Wechsung</i>	20
Working Groups	
QoE White Paper and Group Work 1: Key Aspects of Experience Perception and Their Subjective evaluation <i>Sebastian Möller, Sebastian Egger, and Markus Fiedler</i>	21
Group Work 2: Measurable Aspects of QoE <i>Peter Reichl, Martin Varela, and Markus Fiedler</i>	22
Group Work 3: Identification of QoE-Related Feedback Loops <i>Markus Fiedler and Riccardo Guerzoni</i>	23
Conclusions and Open Challenges	23
Conclusions and Open Challenges <i>Markus Fiedler</i>	23
Participants	25

3 Introduction

3.1 Introduction

Markus Fiedler (Blekinge Institute of Technology – Karlskrona, SE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Markus Fiedler

Dagstuhl seminars strongly depend on the delegates and their input. In order to give room for both presentations and group discussions during a three-day seminar, the presentations were confined to five minutes and one slide. Each presentation was followed by a short block of questions and answers. In order to truly reflect the delegates' positions with regards to the topic of the seminar, the abstracts have been included in the sequel as-is and in alphabetical order.

The presentation round was followed by the presentation of a QoE White Paper around a QoE definition that has emanated from the Dagstuhl Seminar 09192.

The related discussions of QoE-related definitions and notions were continued and deepened during the first group work entitled “Key aspects of experience perception and their subjective evaluation”. The other two groups discussed “Measurable aspects of QoE” and “Identification of QoE-related feedback loops”. The outcomes of the group works were presented and discussed in the plenum, and excerpts are presented below.

The seminar was concluded with a plenary discussion of follow-up activities.

4 Overview of Talks

4.1 Combination of multi-source observations at the sub-second scale


Patrik Arlos (Blekinge Institute of Technology – Karlskrona, SE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Patrik Arlos

Synchronizing measurements from different systems; network, services (app, supporting app, etc...), and human. There is very little usage knowing that at 09:35:01 the user reported/signaled a problem, when there were problems tagged by the network at 09:25:10, 09:35:10 and 09:36:10. At the same range, the services were reporting problems at 09:20:18 and 08:45:10. Then to add to this, how do you add observations from cameras, microphones, EEK etc...

4.2 On-line Estimation of the Quality of Experience

Åke Arvidsson (Ericsson Research – Stockholm, SE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Åke Arvidsson

In cellular networks, the fraction of data traffic surpassed the one of voice traffic in Q4 2009, and by Q1 2012 it has grown to become on the order of three times larger.

The air interface is a limited resource subject to variable demand and variable capacity and most of this often applies to the backhaul as well. This means that cellular networks are likely to encounter temporary bursts of congestion in which case operators have to make intelligent decisions about priorities. In doing so, maintaining QoE should be an important aspect.

The notion of QoE, traditionally measured as MOS, does, however, involve a range of aspects such as, e.g.,

- the terminal (screen, buttons etc.),
- the application (functionality, design etc.),
- the encoding (image resolution, audio fidelity etc.),
- the content (degree of interest, production quality etc.) and
- the presentation (disturbances due to loss, delay etc.).

Since only the last point is directly applicable to managing bursty congestion, it is difficult to rely on MOS in this context. A further complication is that applications and expectations are diverse and subject to constant change whereas MOS measurements are limited and time consuming. Moreover, the special setting around any noticeable attempt to estimate MOS may introduce a bias.

For this reason we are interested in the part of QoE that relates to loss and delay only, and to measure it by observing live traffic. To this end we suggest using a binary QoE metric, acceptable or unacceptable.

For example, acceptable web response times are those which do not prompt users to abort their requests by clicking on stop or other links and we note that both response times and user abortions can be measured in the network. After accounting for the fact that not all abortions are related to response times, we can thus assess the “acceptability” of a certain response time as the fraction of users that finds it acceptable. Congestion control may then be tuned to maximise acceptability and operators may set targets for this value. (An interesting remark is that such a target in many aspects is similar to the classical grade-of-service target in circuit switched networks.)

4.3 A User-Centric Service Modeling approach for User Experience Assessment

Sergio Beker (Huawei Technologies – München, DE)


License  Creative Commons BY-NC-ND 3.0 Unported license
© Sergio Beker

In the context of ever reducing profit margins, operators are turning to Customer Experience Management (CEM) as a strong market differentiator. Current CEM approaches are based on Key Quality Indicators (KQIs) representation of the service performance. Although claiming to be user-oriented, the service modeling frameworks such as eTOM (TMF) and ITIL keep a per-service view, and as such, they are still network-centric. User Experience Indicators (UXIs) would be better suited to represent the user experience, and Customer Experience Indicators (CEIs) to represent the customer outcome of it. The traditional approach to assess UX is to run subjective tests and then to correlate the user answers to the technical measurable aspects of the service. The scope of application for such techniques is narrow around the original context of the test, results in high costs for the operator and proves invasive for the user. In true operational or commercial contexts, a modeling approach would

be preferred. Also, the capability of modeling user experience from network observation alone is a key challenge. The User-Centric Service Modeling (USM) approach models the user experience by taking into account the user interaction with the service. This user centered view allows to include the different aspects of the service and the usage context in estimating the per-user- per-service-per-session User Experience, and to derive the corresponding User and Customer Experience Indicators. Also, by following the customer through the daily interaction with the service, the service-related and non- service related aspects, as well as the user experience with time can be modelled. This pioneering framework has been awarded three patents and is setting the industry standards at ETSI. A platform integrating the USM concepts is under development at Huawei’s European Research Center in Munich.

4.4 Are engineers from Mars and users from Venus? QoE measurement as an interdisciplinary process

Katrien R. De Moor (Gent University, BE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Katrien R. De Moor

From a network engineering and application provision point of view, there is a huge need for unambiguous information: hard numbers, objective data, control. At the other end of the chain are the technology users, who ‘as human beings’ can be predictable in some aspects, but who can also be highly ambiguous and unpredictable in others. In a user’s mind, an experience is not captured or recorded in numerical ratings, it is recorded in thoughts, feelings, stories, images, Bridging these differences in ‘language’ is in my opinion not ‘mission impossible’, but it requires an interdisciplinary approach, it requires ‘translators’ (e.g., psychologists, cognitive scientists, social scientists, . . .).

As previous research has already shown, (the quality of) users’ experiences are dynamic and variable, influenced by a wide range of both human and technical factors and strongly bounded by multilayered contextual aspects. Moreover, there is no such thing as ‘the user’. A person can e.g., be a hardcore online gamer and at the same time an absolute laggard when it comes to mobile services. How to deal with this complexity? There are different levels and stages related to QoE measurement, which in my opinion need a tailored instrumentation. Different steps in this approach could be:

1. Identification of relevant, measurable QoE features (cfr. Qualinet white paper definition: perceivable, recognized and nameable characteristic of the individual’s experience of a service which contributes to its quality), as well as possible influencing factors for a specific service or set of services. E.g., seeking to understand the context in which an application is used, understanding user behavior for specific service with specific affordances, . . . ; In terms of instrumentation, different types of input (objective and quantitative as well as subjective and qualitative) and measurements are needed.
2. Isolation of specific influencing factors to investigate if and if yes, how they interplay and may influence users’ QoE and the features it is composed of. Which patterns can be detected? This typically requires measurements in controlled environments or settings that are at least to some degree ‘controllable’.
3. Extrapolation to more realistic user environments (less control, but higher ecological validity). Understanding QoE from a user perspective requires understanding QoE in the user’s natural ‘habitat’. This is especially crucial in the case of e.g., mobile applications and services.

4. QoE measurement is not a goal in itself, it should be the basis for action: insights and findings should be translated into actionable input for the different stakeholders in the QoE ecosystem (e.g., operators, network providers, ...)

About myself/how I got to the QoE topic: With a background in social sciences, rolling into the field of QoE some years ago was literally a bit of a ‘culture’ shock. The first QoE-project I worked on was a very strange experience. A bit comparable to arriving in an completely unknown country, in which people speak a language you don’t understand, which uses conventions and signs you are not that familiar with. At first there is a huge barrier to meaningful interaction and you might start thinking maybe I should just go back home, to what I am familiar with. But slowly but surely I tried to learn some of the basics of the language and started to realize that some bits and pieces of knowledge in my social science backpack might be somehow relevant after all. Ultimately, QoE is and should be about people. Although the field is already undergoing a major shift to really start from the user’s perception, a major challenge is still related to (1) meaningful interaction: between users, operators and providers, researchers from different backgrounds and with a different expertise; and (2) translating insights arising from that interaction into ‘actionable’ and tangible input at different levels.

4.5 A network based method for Video QoE measurement


Marcus Eckert (TU Chemnitz, DE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Marcus Eckert

A new method for monitoring the quality of Internet video streaming is presented. The method fully relies on network-centric measurements: it determines the buffered play-out time during progressive download by the evaluation of TCP segment timings in measurement traces. Results from tests carried out in a real mobile network show good agreement with user perceived quality measured directly at the end device. The performance of the method is compared to results of other approaches.

4.6 Waiting for QoE

Sebastian Egger (FZ Telekommunikation Wien, AT)


License  Creative Commons BY-NC-ND 3.0 Unported license
© Sebastian Egger

Recent work on QoE sets out to identify natural psychophysical relationships between the network (stimulus) and user perception (response), with network- induced waiting time being a specific example of such a stimulus which directly affect user satisfaction and thus QoE. Especially in the context of interactive data services, QoE is determined by such waiting times to a large extent, a fact which has led to the catchy notion of WWW as World Wide Wait. A large share of services e.g. file downloads, E-Mail browsing, picture viewing or basic web browsing is characterized by an information request from user side and respective waiting times until the request is fulfilled. The past shift from UDP media streaming to TCP media streaming (e.g. youtube.com) has extended the relevance of waiting times also to the domain

of online video services. Therefore the following questions are of particular importance to the QoE community: Which waiting times are sufficient to ensure a certain degree of user satisfaction? Are the waiting times translatable between different services?

4.7 Quality of Experience instrumentation: Read the user AND the network

Markus Fiedler (Blekinge Institute of Technology – Karlskrona, SE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Markus Fiedler

My background is teletraffic modelling and analysis in communication networks, which means that I entered the QoE domain from the network measurement, modelling and analysis point-of-view. Triggered by Swedish industry to look for indicators of bad user performance that might cause user churn, our first QoE-related Quality of Service (QoS) study was a bottleneck indicator [1].

During the last decade, QoE has taken over the role that QoS was supposed to have, with focus on multimedia applications and spatial distortions. During classical QoE studies, users have to sit still and rate QoE under very well defined circumstances – this might be repeatable, but also far from real-life situations.

We realise that presence or lack of quality is affecting user's behaviour towards systems, meaning that the way a user interacts with a system mirrors performance and QoE issues. User reactions to application service performance as seen from two-way measurements in user interfaces have a lot to tell and should be exploited to a much larger extent than what is the case today. Furthermore, recent results reveal links between energy consumption patterns and QoE issues (cf. Selim Ickin's contribution). On the other hand, temporal distortions in data flows are visible within the network and can be detected without even bothering the end users [2]. Though having been claimed for many years, efficient feedback channels between users and providers/operators are still missing [3].

Interestingly enough, classical teletraffic modelling and analysis did not follow the above-mentioned shift from QoS to QoE, although they are well prepared to capture (so-far untended) temporal QoE issues. Based on experience with the search for user-perceived bottlenecks through Comparative Output/Input Analysis (COIA) [4] and with recent modelling efforts of mobile connectivity, both based on an underlying fluid model, we postulate the potential of queuing models to quantitatively describe QoE issues [5]




References

- 1 M. Fiedler, K. Tutschku, P. Carlsson, and A. Nilsson. *Identification of performance degradation in IP networks using throughput statistics*. Proc. 18th International Teletraffic Congress (ITC 18), pp 399–408, Berlin, Germany, Sept. 2003.
- 2 J. Shaikh, M. Fiedler, D. Collange, and P. Arlos. *Modeling and analysis of web usage and experience based on link-level measurements*. To be presented at the 24th International Teletraffic Congress (ITC 24), Cracow, Poland, Sept. 2012.
- 3 M. Fiedler. *Quality feedback flows in future networks*. elektrotechnik und informationstechnik (e&i), Springer, Vol. 7-8.2009, pp 269–273.
- 4 M. Fiedler, K. Wac, R. Bults, and P. Arlos. *Estimating performance of mobile services from comparative output-input analysis of end-to-end throughput*. To appear in IEEE Transactions on Mobile Computing.

- 5 M. Fiedler. *Teletraffic models for Quality of Experience assessment*. Tutorial at 23rd International Teletraffic Congress (ITC 23), San Francisco, CA, Sept. 2011.

4.8 QoE IPTV

Marie-Neige Garcia (TU Berlin, DE)

License    Creative Commons BY-NC-ND 3.0 Unported license
© Marie-Neige Garcia

Models developed within the standardization groups such as the International Telecom Union (ITU) and the Video Quality Expert Group (VQEG) output video/audio/audiovisual quality (e.g. MOS) scores for short- time (10 s) audiovisual sequences, while providing a score for a whole session (e.g. a whole TV program in case of IPTV) may also be necessary. Moreover, the same score is provided for all types of user and context (e.g. for IPTV: test lab vs. home environment). These short-term quality scores are not sufficient for knowing if a customer is satisfied with the service/product she/he subscribed to. If we want to find the link between the quality scores provided by the quality metrics and user satisfaction, several points have to be addressed:




- How do we identify what the dimensions underlying QoE for a given application are? For IPTV, these dimensions may be for instance the perceived quality, accessibility/stability and usability/joy of use. Subjective tests are required for identifying these QoE dimensions. Multidimensional analysis (e.g. semantic differential scaling followed by Principal Component Analysis) and interview-based methods are examples of mix of qualitative and quantitative test methods which could be applied in that respect.

- How do we make QoE-models ecologically valid, user- and context- dependent? One requirement is bringing subjective testing to the field. In addition, new subjective tests should be designed for addressing long audiovisual sequences (as with the SSCQE of the ITU-R BT-500 Rec.) but with more appropriate task so that the subjects focus on the content and not on quality anymore (in the latter case, subjects are becoming too sensitive to impairments, see Staelens, 2010, IEEE Trans. On Broad.) For making QoE-models user-dependent, the most relevant criteria (personality? degree of expertise? Demographic data?) for characterizing the users should be identified. At last, we need to adapt the models when the 'technical' context is changing (e.g. IPTV is becoming interactive but models are trained in the context of a one-way (non- interactive IPTV services)?

This statement tried to identify requirements for modeling QoE in terms of subjective testing and in the context of a concrete application (IPTV).

4.9 User-centric troubleshooting

Riccardo Guerzoni (Huawei Technologies – München, DE)

License    Creative Commons BY-NC-ND 3.0 Unported license
© Riccardo Guerzoni

The speech introduces a novel approach to current and future ICT networks troubleshooting, designed around user-centric criteria and QoE modeling. The proposed methodology for symptoms localization and root causes identification, denoted as User Centric Troubleshooting

(UTS), goes beyond the classical hierarchical relations between key performance indicators (KPI) and key quality indicators (KQI).

In the proposed process, the triggering point is the detection of performance deteriorations either in quality of experience (QoE) or in related QoS parameters. The input data may be customer complaints or QoE models. (It is well known that Network KPIs are not reliable parameters.) When any of the reference parameters goes below the corresponding threshold, all affected sessions and related pattern of anomalies are identified and aggregated by correlating Transaction Data Records (TDR), derived from network protocols, and Session Data Records (SDR), attained from upper layers protocols. The recurrence of an anomalies pattern (detectable by a clustering algorithm) and related KPI distribution make it possible to identify the root cause of the QoE/QoS issue, which triggered the diagnosis process, exactly like an expert engineer would operate. The information carried by the TDRs can be organized in classes of anomalies, standardizing the diagnosis parameters fed to the clustering algorithm. A pivoting step completes the analysis, identifying the principal components among the KPI dimensions (network context, service context, user context).

The aggregated per user per service (PSPU) SQM approach proposed by the UTS framework links the users segments impacted by the QoE deterioration to the root cause analysis, enabling the network operator to allocate efficiently the budget for the network optimization, working on the issues that are actually affecting the users perception and keeping the churn under control.

Experimental results showed the proposed solution to be an essential component for efficient user centric Customer Service Assurance.

4.10 Long-term QoE: How does the overall quality perception of one user evolve over multiple interactions with IP-based services?


Dennis Guse (TU Berlin, DE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Dennis Guse

Shifting the perspective from technical parameter-based to perceptive quality estimation will allow tailoring service quality and network performance perfectly suiting to the user's need, i.e. provide an optimal QoE. In my research, I am focusing on the quality experience over multiple distinct interactions with one or more IP-based services. The user integrates his quality perceptions into an overall perceived quality, which influences his behavior like reusing the service. Understanding these effects would give us the capability to provide not only satisfying single service interactions, but going for real-life long-term QoE.

4.11 QoE Assessment for Web-based systems and IP TV

Richard John Harris (Massey University, NZ)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Richard John Harris


Subjective assessment for measuring the Quality of Experience may require time-consuming and often expensive methods, and yet, quantitative and accurate user scores are desired. In order to obtain valid correlation between analytical model and user scores, assessment based

on networking perspectives and human perception is required. In our study, we have utilized orthogonal arrays using the Taguchi approach to construct an experiment to characterize the application as well as network performance metrics in our QoE assessment model for web-based systems.

For our study of QoE in IPTV and related systems we have divided the model into three basic components, viz: the Content Producer, the Network and the Customer Premises Equipment and selected metrics that influence QoE from each of these categories to study and develop models to be integrated into a single platform.

4.12 You look blocky, is everything alright?


Helmut Hlavacs (Universität Wien, AT)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Helmut Hlavacs

Non-verbal communication is known to convey a significant part of the message between human beings during their interactions. Among other types of body language, facial expressions represent an important medium to evaluate the mood and feelings in the person one is interacting with. Nowadays, more and more distant communication is made possible by the fast growth of networks and devices such as mobile phones, tablets and traditional computers. For a few years now, we have been facing a rapid development of video communication, which allows distant persons to exchange either live or recorded messages in high fidelity. However, all distant communication has to go through a chain of treatments that can alter the quality of the delivered message. From the capture of the message using a video camera, through compression and transmission, then through decompression, post-treatment and finally to display, each step of the delivery process can introduce degradations in the data and deteriorate the message. The impact of the multimedia processing and delivery channel on the human ability to recognize facial expressions is therefore quite important and currently researched by us.

4.13 QoE: Measuring the Immeasurable?

Tobias Hossfeld (Universität Würzburg, DE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Tobias Hossfeld

My personal background is performance analysis in communication networks. As focus of this abstract, we consider a user-centric service and network management. From the view-point of a network or service provider, the goal may be to offer a good QoE to its customers while at the same time to optimize the costs for QoE management. Hence, the provider requires some indicators and instruments to quantify the users' (dis-)satisfaction with a service. The question arises whether QoE is the right path for this? Whether QoE can be measured in real-life system or whether QoE is too complex? Whether other (measurable) metrics instead of QoE can be exploited?


The intention of this abstract is to provoke discussions beyond QoE and to ask critically how to utilize QoE, how to implement QoE metrics for a certain purpose in practice, what are the next steps for QoE research in general.

One simple solution to overcome these challenges above is to go into a different direction than QoE. Instead of monitoring and estimating the user perceived subjective quality of a service, we may have a closer look at the user reactions and the user behavior. For example, the actions of an user (where and when he clicks on which elements of a web page) give a clear answer. E.g. when there is a problem with a web page, the user may reload the page or start another web page in parallel in a different tab of the Internet browser application. Another example is to measure how long and whether the user pauses a video to avoid stalling for example. This may show that the user is not happy with the current network situation and overcomes it. For an Internet service provider, the user reaction (especially, when the user churns) may be more important than the user perception! Furthermore, there is a known gap between user perception and user reaction.

In summary, as a next step of QoE research, the consideration and analysis of the user behavior (complementing the QoE perspective) seems to be important. Then, this user feedback should be taken into account in the service delivery accordingly.

4.14 Energy metrics unique enough for smartphone-based video QoE evaluation?

Selim Ickin (Blekinge Institute of Technology – Karlskrona, SE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Selim Ickin

Mobile applications and services are used in daily life activities, to support the needs for information, communication or leisure. Mobile video applications, so called apps, are of the most energy/network demanding ones and need to be investigated further with the goal to achieve video quality close to the overall perceived quality of the applications running at fixed networks and terminals. In this talk, we will be presenting one unique metric, energy consumption, as a bellwether to locate both user activity levels and the video application statuses on a smartphone. Most popular video applications work based on transmission-controlled streaming, i.e., video plays without any jumps in the video frames. However, in this type of streaming, a stalling event, so-called freeze, is a common impairment that occurs during a video play-out and it is considered as a key influence factor in user's recent perceived video quality. An instantaneous increase in the network layer metric values such as delay and packet loss rate may not always be a herald for an interruption of a video streaming by a stalling event, due to its dependency on the content/size of the jitter buffer. Therefore, not only focussing on traditional QoS metrics, but investigation of other metrics that represent the state of the video streaming in the application level is necessary.


Our previous study tells that battery lifetime in smartphones is one of the most influential factors for overall smartphone-based QoE. Monitoring QoE especially on these energy limited devices is a challenge and needs smarter QoE monitoring frameworks. More recently, we have investigated that, also by continuous instantaneous per-application/per-service basis energy consumption measurements, it might be possible to detect anomalies such as stalling events in video applications. We have recently investigated that the frequency and the duration of the freezes during the runtime of a mobile phone based video player application are likely revealed by energy consumption values. As our hypothesis, we propose an energy consumption metric that is likely correlated with the stalling events in transmission-controlled video streaming.

Abnormal fluctuations of instantaneous energy consumption metric can indicate the

stalling events experienced in the phone display during a user's video streaming session. In this way, instead of using hard to deploy and high energy demanding network level monitoring tools, we can facilitate the built-in energy measurement tools on mobile handheld devices to locate those anomalies. Once a suspicious behaviour such as stalling events are detected through energy measurements, deeper analysis of underlying QoS metrics, sensors, and all other user dependent traces can be enabled, which are expected to provide a more energy friendly future QoE monitoring framework.

4.15 Repeatable Results – The Key to Scientific Accuracy


Lucjan Janowski (AGH Univ. of Science & Technology – Krakow, PL)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Lucjan Janowski

As community we produce lots of metrics, indicators and other type results, the problem is that they are not checked by others. The key to have meaningful results is to be able to repeat them. It is not QoE community problem but in general IT. Orange lab took 99 % accurate traffic classification algorithm and it classified 50 % of traffic as unknown. In order to change it AGH would like to propose a joint effort of creating a journal focused on QoE problems. This journal will publish not only currently the most important trends in QoE research but also metrics validation, repeating subjective experiments conditions, and other articles focused on checking other researchers' results. In addition the journal could provide a platform to review a paper not only by checking its text but also results.

4.16 Impact of physical layer impairments on higher layer QoS parameters


Maria Kihl (Lund University, SE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Maria Kihl

An increasing demand from e.g. real-time multimedia applications (IPTV, OTT) adds strains on especially DSL based access links. Advanced physical layer monitoring tools are deployed, however, the mapping from physical layer impairments to network layer QoS parameters is still rather unknown. In this presentation we argue why it is important to study DSL link impairments and their impact on QoS parameters on higher layers.

4.17 Towards Total Quality of Experience. A Conceptual View on QoE in a communication ecosystem

Khalil Laghari (Télécom SudParis – Evry, FR)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Khalil Laghari

With clear understanding on how human behavior is shaped and what disparate factors could influence his/her needs and expectations in particular context, it is possible to get

more reliable and authentic view on QoE. A communication ecosystem brings together various domains such as technical aspects, business models, human behavior and context. For each aspect of a communication ecosystem, various models have been developed. However, few models have been designed to integrate all aspects of a communication ecosystem to understand human behavioral needs in a detailed and structured way. While existing models have produced the basic sketch of QoE modeling, more concepts and inter-domain mapping are to be incorporated in order to have a clear picture of QoE in a communication ecosystem.

4.18 Quality of Experience is not only Quality of User Experience


Patrick Le Callet (Université de Nantes, FR)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Patrick Le Callet

Since the late 90's, QoE has reached its great momentum in several communities. Nevertheless, clear understanding of its relative concepts is still missing. QoE to be practicable needs contextualisation. An obvious witness relies on the effort raised by the UX (User Experience) community. Assessment of the user experience is a key aspect and with this respect QoE is certainly related to UX. Nevertheless, QoE encompasses more scenarios Experience is not necessarily attached to the user of services or applications and encompasses much more than the relationship between a user and a service in many cases: visiting an expo, watching a movie, wine tasting The contents care and the emotion that one be suffered to convey.

4.19 How to come to good QoE instrumentation?


Sebastian Möller (TU Berlin, DE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Sebastian Möller

In this talk, 10 steps are highlighted which are necessary to define instrumentation concepts for QoE: (1) Define what you understand by Quality of Experience (Qualinet white paper). (2) Analyze relevant perceptual dimensions (features) of QoE. (3) Structurize the space of features (see the example in Möller et al., QoMEx 2011, for speech services). (4) Investigate models to predict desired features (several models are still missing). (5) Analyze how you can obtain the necessary input information (what is possible?). (6) Analyze possible user behavior (carry out data analytics). (7) Investigate whether and how you can predict such behavior. (8) Define what you do instrumentation for (monitoring, adaptation, charging, offers, etc.). (9) Decide which features and which user actions contribute to that aim. (10) Define an instrumentation approach.

4.20 Qo(E) Vadis? Multi-user, multi-service, multi-information, multi-timescales

Alexander Raake (TU Berlin, DE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Alexander Raake

The majority of current QoE research is focusing on unidimensional, utilitarian measures of QoE, expressed for an average user of an individual service. Typically, such approaches are taking a system-, that is, object-oriented perspective, where the main goal is the QoE-based performance evaluation of objects such as services, applications or multimedia systems. For practical application, user tests are complemented or fully replaced by instrumental models used to predict QoE as perceived by users. Such predictions are usually based on model input information such as system parameters or multimedia output signals.


These approaches must be extended in different directions:

- For current models, the underlying input information describes the system or partially also the transmitted content, but mostly does not include contextual information or user behavior data. To overcome the limited ecological validity, and exploit all relevant sources of information on user-perceived QoE, multi-level-information approaches must be adopted including both context and behavioral information. Research challenges will lie in identifying the optimal set of information, and developing strategies for handling missing information.
- Most QoE models are agnostic to the kind of user, although QoE is individual to a given user, and largely depends on her personality and current state. For different applications, it is desirable to overcome this limitation by making models user-specific. Here, an appropriate demographic and role-related user classification as well as behavioral information will need to be exploited by future approaches, and included in a respective user-model.
- The still sparse examples of multi-timescale QoE assessment, such as in the case of call quality models for speech QoE of telephony services, must be complemented, to go beyond quality predictions for a single, mostly short-term time scale in the 5-16 sec range. Temporal pooling of QoE features and QoE episodes across different time-scales is not well understood, and a highly relevant topic for further QoE-research.
- If QoE models for an individual service are used for service optimization, the across-service perspective of subscribers is normally not reflected. As a complement, a multi-service perspective must be adopted, catering for how users integrate QoE across different services.

These research topics have to be addressed by finding a good balance between the achieved degree of ecological validity and the practical applicability for deployable QoE assessment solutions.

4.21 On Economics of QoE


Peter Reichl (FZ Telekommunikation Wien, AT)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Peter Reichl

Probably, QoE can be considered the perfect example for a research topic in the interdisciplinary research field which has recently been termed “Communication Ecosystems” (cf. Kilkki, Laghari et al.), which unites the technological (T), economical (E), user (U) and context (C) perspectives on communication systems as orthogonal and/or competing forces. While the interface U-T has made significant progress by now, both “technoeconomics” (T-E) and “socioeconomics” (U-E) are still lagging behind, while context basically provides an additional orthogonal component. Our main goal is to revisit the economical part of this sketched ecosystem. To this end, we focus on the methodological role of economics as a means of modeling the interaction between users as well as between user and technological environment. Following a hierarchical ecosystem model, competition takes place within a layer, while user utility is maximized by employing lower layer resources.

4.22 Automatic QoE assessment

Gerardo Rubino (INRIA – Rennes, FR)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Gerardo Rubino

1/ The PSQA (Pseudo-Subjective Quality Assessment) technology.

In the past, we developed a no-reference parametric assessment methodology for the evaluation of the perceptual quality of a voice or video + audio communication when the channel includes a packet network. PSQA works by measuring specific QoS metrics and specific application-based parameters, and then by invoking a particular function that maps all these variables into a MOS-like number. These metrics and parameters must then be accessible, and with almost no cost (that is, they must be measured efficiently, or available “for free”). The function is built using statistical learning tools, working on data coming from subjective testing sessions, and a mix of random and quasi-random sampling techniques to prepare the sequences to evaluate in those sessions. Once built, PSQA works in real time, if useful or necessary.

We claim that PSQA is accurate enough for any network-oriented usage (typically, for network monitoring or for network control). The procedure is network- and application-dependent: if the network and/or the application evolve, or change, the measuring module must be developed again from the beginning. However, this effort is done only once, before putting PSQA at work. Once built, the module just measures the instantaneous perceptual quality at time t .

We are currently in the process of putting PSQA in industry, in a “2.0” version, where we are integrating the experience cumulated over some years of application in several different domains, together with improvements on the original techniques.

2/ From perceptual quality to QoE

Compared to the case of voice or video communications, QoE covers an immensely larger range. This makes the idea of extending the parametric approach we followed for the simpler

perceptual quality assessment of voice or video communications difficult to apply as such. Said in other terms, this huge universe makes difficult to achieve the same accuracy than PSQA in case of, say, a given Web service, or when considering the generic class of Web services as a whole. Instead, we are exploring two ideas.


- (i) First, we are looking at a procedure based on the assumption that the target, the system to be assessed from the QoE point of view, exists in a large number of instances. The idea is to request specific users, at specific points in time, to provide an opinion about the QoE of the system, and to integrate the collected information in a way such that we can elaborate an automatic and accurate QoE measuring tool.
- (ii) Second, we are considering the idea of building a QoE metric with values in a multi-dimensional space equipped with some mathematical structure.

We are currently working on the first point above, point (i), focusing on the mathematical foundation of the approach. The point is that the subjective views given by the users of the measured system will not have the same value (the same “quality”) as the scores obtained from a controlled subjective testing experiment. The other side of the coin is the fact that we expect a large number of opinions, must larger than the number that can be obtained using panels of users in the lab. The situation is similar (at least, formally) as what one encounters in Monte Carlo, where for estimating the integral of $f()$ in some interval, we only need to answer the question “is $f(x) > u?$ ” given x and some random value u , for many pairs (x, u) . The question is quite simple, but the result can be an extremely accurate answer to a non-trivial question. Technically, we are looking at what happens with accelerated Monte Carlo where the system’s dynamics is completely changed in order to reach efficiency criteria. We are also looking at Quasi-Monte Carlo methods, where instead of random u values with use quasi-random ones, that is, weak discrepancy sequences. Here the goal is some regularity in the sample (instead of something mimic randomness), a tool we already use in PSQA.

Our first objective is to develop a first set of mathematical tools able to be combined into a measuring technique. Then, we must test the idea on well-chosen specific cases, including some of the scenarios we are identifying in the QuEEN project.

4.23 Non-intrusive network-based estimation of QoE

Junaid Shaikh (Blekinge Institute of Technology – Karlskrona, SE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Junaid Shaikh

The number and types of applications on the Internet are increasing. Each application has its own performance requirements to work smoothly on the network. On a network with the same available resources, the same user can have different experiences based on the type of application used. Therefore, dimensioning a network for all types of applications according to the same criterion may either lead to the congestion (and user churn) or waste of resources (under-utilization of available capacity). Both situations are undesirable for the network operators. It triggers the need for QoE-aware management of networks, in order to organize networks dynamically based on the real-time estimation of QoE.

Several assessment models have been proposed to estimate QoE. Most of them are intrusive and require knowledge of the content reference. In contrast, the network operators require non-intrusive methods, which allow models to be implementable on the network-level


without having much knowledge about that reference. The methods should be able to monitor QoE passively in real-time, based on the information readily available on network level. Considering the high-speed networks, methods should also be fast and practically implementable.

Our work is based on the non-intrusive methods to infer QoE based on the objective indicators obtained from the network traffic. The methods take into account ON and OFF phases of user traffic on the network. They do not require deep packet header information. Amongst others, these methods capture the temporal aspects of QoE and locate those time instances where the problems occur. By using these methods, the frequency and duration of user-perceived problems could be visualized at varying time scales.

Once developed, the above-mentioned methods are intended to be used in the live networks to estimate QoE in the quasi-real-time.

4.24 From modeling QoE to QoE management: challenges for domain-wide QoE-driven resource allocation

Lea Skorin-Kapov (University of Zagreb, HR)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Lea Skorin-Kapov

While a large number of approaches to date employing utility-based resource allocation concepts have linked technical performance metrics (associated with QoS) with perceived service quality, the advent of QoE-related research has led to advancement in the understanding of QoE metrics, focusing on subjective quality perception and subsequent human psychological and behavioral models. It has been widely accepted that QoE is a multidimensional construct, comprised of multiple user perceived quality dimensions. In the domain of QoE-driven service optimization, the concept of utility functions can hence be applied to include such multiple dimensions (e.g., ease-of-use, efficiency, comfort, satisfaction, visual quality, willingness to pay, perceived value for money, etc.) and their impact on an overall (integral) QoE. Furthermore, it should be noted that different dimensions are relevant for different types of services (e.g., conversational voice, streaming video, cloud applications, multiplayer games, etc.).


Based on QoE models and estimation methods, we can attempt to correlate QoE dimensions at a given point in time with QoE influence factors (related to application/service parameters, allocated resources, user parameters, context parameters). By understanding this correlation, and the user, service, and network constraints which are pertinent, we can formulate QoE optimization problems related to tuning (where possible) influence factors (e.g., resource-related, application configuration-related) to maximize QoE.

A question we raise is how to go from our estimation of QoE in a multidimensional QoE space (having so far considered a single user and a single service), to performing domain-wide QoE optimization, whereby we are looking for a global solution in a multi-user, multi-service QoE space, constrained by available network resources, user subscriptions, different user preferences and capabilities, operator policy, etc.? We discuss some challenges related to QoE-driven resource allocation in a domain-wide scenario. Not only do we have multiple dimensions of QoE that need to be considered for different services (and media flows in the case of multimodal services), but we also have multiple simultaneous user sessions. In practice, formulation of the objective function for optimizing resource allocation may differ

depending on whose interests are being considered (e.g., user, network operator, or both in the case of multiobjective optimization). Different examples include: (1) maximizing the (weighted) sum of QoE values across end users, expressed generally as functions of QoS parameters, (2) maximizing number of “satisfied” users, i.e., with QoE above a certain threshold, (3) maximizing operator profit, by minimizing operator costs, or (4) a combination of the previous objectives.

4.25 A Generic Approach for Understanding QoE

Martin Varela (VTT Technical Research Centre of Finland – Oulu, FI)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Martin Varela

QoE research has so far been mainly focused on media services and applications, such as (Internet) telephony, video, etc. The concept of QoE, however, is applicable (and important!) to a much broader domain of services and applications, which are becoming ever more important in everyday life, as more and more activities take place online. Online collaboration, social networking, e-banking, and a myriad other applications, often web-based and residing in the Cloud, are now an integral part of everyday life. Oftentimes, applications which used to be local to the user’s devices, now reside somewhere else, with a network, hosting platform and application platform in between them and the users, all of which introduce potential degradations in the way the user experience the service.

Given the ubiquity of these applications and the increasing importance of their role in our lives, it makes sense to try and understand how we experience our interactions with them (i.e. what is ‘their QoE’). This poses several challenges, starting with basic ones such as understanding what quality means for users in the context of each of these services.

In the context of the Celtic+ QuEEN project, and also together with Qualinet colleagues, we are currently working on a conceptual framework for understanding QoE for any service. The goal is to be able to reason about QoE and exploit it in different ways (e.g. SLA negotiation/enforcement, QoE-driven network management, etc.)

To this end, we consider QoE as a multi-dimensional construct which depends on several (with ‘several’ » ‘a few’) factors, and we strive to understand the relations between these factors and the many aspects (or dimensions) of QoE. For this we use different tools, such as Lea Skorin-Kapov’s ARCU model for classifying QoE-influencing factors, and a layered, compositional view of quality for understanding the relations between different parts of the service-user system.

We further consider that services often comprise multiple modalities, which play differently on the overall (integral) QoE of the service (i.e. they can relate differently to different dimensions of QoE), and which may change over time within an usage session (e.g. video-conferencing with screen-sharing facilities).

As of this writing, we have a first approach to formalizing the concepts above, and we are currently working on incorporating the temporal/multi-modality aspects to the model. The main challenge, however, lies in finding efficient ways of constructing a mapping from the quality-influencing factors, and QoE itself, taking into account all of the above. We have, in the past, constructed such mappings in mono-modal applications (VoIP, video) for single dimensions of QoE (e.g. with PSQA). More recently, we have successfully created such a mapping for another mono-modal application (video), considering several QoE dimensions in the output, as well as an integral QoE estimation. Making this work in a more generic

setting, however, with larger number of influencing factors and QoE dimensions, presents a non-trivial challenge in terms of the amounts of data needed to be able to use statistical approaches to create the mappings. We are currently working towards such a more complex use case, for a web-based service.

4.26 Factors Influencing Quality of Experience of Commonly-Used Mobile Applications


Katarzyna Wac (University of Geneva, CH)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Katarzyna Wac

Increasingly, we use mobile applications and services in our daily life activities, to support our needs for information, communication or leisure. However, user acceptance of a mobile application depends on at least two conditions: the application's perceived experience and the appropriateness of the application to the user's context and needs. Yet, we have a weak understanding of a mobile user's Quality of Experience (QoE) and the factors influencing it. We present a week long, 29 Android phone users study, where we collected both QoE and underlying network's Quality of Service (QoS) measures through a combination of user, application and network data on the user's phones. We aimed to derive and improve the understanding of users' QoE for a set of widely used mobile applications in users' natural environments and different daily context. We present data acquired in the study and discuss implications for mobile applications design.

4.27 The influence of contextual factors on quality ratings

Ina Wechsung (TU Berlin, DE)

License  Creative Commons BY-NC-ND 3.0 Unported license
© Ina Wechsung

User eXperience (UX) is widely understood as a highly context-dependent construct based on individual perceptions. This position implies that to achieve meaningful measurements of UX the context needs to be taken into account.

Research indicates that judgment and decision making involves two systems, the cognitive-rational and the intuitive-emotive system (e.g. Kahneman 2003). Compared to the emotive system, the cognitive-rational system is more analytic, logical, abstract, active, controlled, rule-based and slower; it is the deliberate mode of judgment [1]. The intuitive-emotive system on the other hand is characterized by automatic, associative, effortless, and often emotionally charged operations; it is the automatic mode of judgment, and is also shown to be more context-specific than the rational system [2, 3]. The heuristic, context-specific, emotive system determine preferences and judgments unless the cognitive system intervenes [1].

To gain a better understanding of the role of context in the judgmental process, we investigated whether findings from cognitive psychology could be transferred to the judgment of interactive systems. For example we found the mood congruency effect [4] to be also applicable to HCI: the better the mood of the participants, the better the ratings for perceived hedonic quality. Furthermore, we showed that increasing mental workload by introducing a parallel task decreased the perceived pragmatic quality. In our most recent study, we

compared ratings collected “online” during a field trial with ratings assessed after the usage period. While quantitative ratings of overall quality were similar, the qualitative data differed: comments collected during usage were more specific with respect to certain negative or positive aspects of the apps performance. Participants often only reported problems, not judgments. Comments collected after usage were often rather general, however they also contained an affective appraisal of the experience. Thus the remembered user experience does not necessarily represent a one-to-one reflection of the actual user experience.

The results reported above show that although laboratory studies that aim to strictly eliminate contextual factors might be appropriate for performance evaluation, such settings are certainly not the best approach for meaningful assessments of the multi-faceted concept UX.

References

- 1 Kahneman, D. (2003) A Psychological Perspective on Economics, *American Economic Review*, 93(2), pp. 162-168.
- 2 Epstein, S., Pacini, R., Denes-Raj, V., and Heier, H. (1996) Individual differences in Intuitive-Experiential and Analytical-Rational Thinking Styles. *Journal of Personality and Social Psychology*, 71(2), pp. 390-405.
- 3 Denes-Raj, V. and Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgement. *Journal of Personality and Social Psychology*, 66(5), pp. 819-829.
- 4 Schwarz, N. and Clore, G. L. (1983) Mood, misattribution and judgement of well-being. Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45(3), pp. 513-523.

5 Working Groups

5.1 QoE White Paper and Group Work 1: Key Aspects of Experience Perception and Their Subjective evaluation

Sebastian Möller, Sebastian Egger, and Markus Fiedler

License © ⓘ ⊖ Creative Commons BY-NC-ND 3.0 Unported license
© Sebastian Möller, Sebastian Egger, and Markus Fiedler

Main reference “Qualinet White Paper on Definitions of Quality of Experience (2012). European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)”, Patrick Le Callet, Sebastian Möller and Andrew Perkis, eds., Lausanne, Switzerland, Output version of the Dagstuhl seminar 12181, 2012. (last seen 2012-07-21).

URL <http://www.qualinet.eu>

On the end of the first day, after the participants’ presentations, the QoE White Paper prepared by the Qualinet group, led by Sebastian Möller and Patrick Le Callet, was presented. The QoE definition in this document (“degree of delight”) originates from the Dagstuhl Seminar 09192 “From QoS to QoE”. The subsequent discussion dealt with the relation between QoE and User eXperience (UX); the concept of experience; the relationship of QoE and the Communication Ecosystem; influencing factors on user level; quality features on service level; and the relationship between QoS and QoE.

These discussion points led to group work, which resulted in an updated version of the Qualinet White Paper, in particular with respect to the concepts of experience and quality of experience, as well as with respect to multimedia learning. This updated version was presented in the Friday morning wrap-up session.

During this session, it was agreed to compile an Output Version of the White Paper until 18.05.2012, work led by Sebastian Möller, with the following tasks:

- Start a discussion in a larger group around the term “experience” and clarify whether this also include perception and judgment processes and user states; potentially provide an update of the definition in Section 2 of the White Paper;
- Check whether the definition of QoE needs to be modified given a new definition of experience; check whether the sentence related to telecommunication services can be excluded from the definition to reflect that QoE exists also without telecommunication services;
- Provide a tentative text to be added to Section 4 which clarifies the different roles users might take in a communication ecosystem, and how this relates to QoE (Kalevi Kilkki could provide support);
- Provide an update of the paragraph of Section 4 which related to multimedia learning, so that it better reflects QoE; provide a short explanation of what the “level of service” means for the QoE features in Section 6.

In the meantime, these targets were all reached, and the Output Version of the Qualinet White Paper is now available through the Qualinet Web Site [1].

References

- 1 “Qualinet White Paper on Definitions of Quality of Experience (2012). European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)", Patrick Le Callet, Sebastian Möller and Andrew Perkis, eds., Lausanne, Switzerland, Output version of the Dagstuhl seminar 12181, 2012. <http://www.qualinet.eu> (last seen 2012-07-21).

5.2 Group Work 2: Measurable Aspects of QoE

Peter Reichl, Martin Varela, and Markus Fiedler

License  Creative Commons BY-NC-ND 3.0 Unported license
© Peter Reichl, Martin Varela, and Markus Fiedler





In parallel with group work 1, this group discussed issues and challenges related to measurable QoE-affecting factors. The results were summarised in a table, listing features and related properties such as measurability, relevance and challenges. A summary of the results is provided hereafter:

- Personality, health aspects and mood are measurable in general (personality rather indirectly). Challenges relate to how to measure them, in particular in field measurements.
- Experience and demographic data are knowable and can be measured both in- and directly. No particular challenges were identified.
- Expectations are indirect measures. They are deemed the more relevant and the more difficult to estimate, the longer the underlying time perspective becomes. Challenges are amongst others found in dependencies on applications and in the roughness of the estimates.
- Technical factors can be measured for the most part in a quite straightforward way. Environmental factors, as far as they are relevant, might be difficult to measure in the field. Socio-economic factors can be estimated, while it is difficult to capture their influence in models.
- Group and role factors can be estimated and are knowable. Role factors are deemed more influential and can be assessed through social graph analysis.

- Mobility factors are directly measure. Task factors can also be measured indirectly and might be difficult to assess.
 - Cost factors are knowable, while emergency factors cannot be measured.
- A use case exemplifying the above factors was prepared by Tobias Hossfeld.

5.3 Group Work 3: Identification of QoE-Related Feedback Loops

Markus Fiedler and Riccardo Guerzoni

License     Creative Commons BY-NC-ND 3.0 Unported license
© Markus Fiedler and Riccardo Guerzoni

This group work was also performed in parallel to the other two group works, and was structured as follows:





1. *History and motivation.* One of the strongest drivers of QoE is the risk of user churn in face of perceived problems. It is important to find out about users' real opinions and their dynamics in order to pave proactive ways to avoid churn. As questionnaires add noise to customer experience, objective measurements for the entire population are needed.
2. *Classification.* A set of explicit feedbacks (e.g. push an “anger” button; call the support; body response from face expression via exclamations and device abuse to violence), implicit feedbacks (e.g. more and longer activity; increased usage frequency and degree of completed transactions; higher spendings; word of mouth etc.) and hybrids (e.g. timings in the user interface; stop/reload buttons) were identified.
3. *Inventory.* It was discussed which QoE-related feedback a user expects, which feedback facilities a user might miss, and which feedback should be provided from the operator's point of view. Obviously, there is a risk of over-polling the user. Feedback should correlate to the possibilities of the user to control the situation. It is definitely dependent of the type of service. A faithful user would consider no news as good news and value notifications of “turbulence ahead”, which is particularly important for transactions that involve several steps.
4. *Construction of new feedback loops.* This part, which seems to be best done based on case studies, still needs to be addressed, e.g. for some interactive service. From the manufacturer's point of view, questions on how to use QoE evaluation for network management and how to assure that the investments into quality improvements feed back positively into the accounts of the network operators are of high importance.

The sub-group planned to follow up on this topic with a survey paper (during Summer/Autumn 2012).

6 Conclusions and Open Challenges

6.1 Conclusions and Open Challenges

Markus Fiedler (Blekinge Institute of Technology – Karlskrona, SE)

License     Creative Commons BY-NC-ND 3.0 Unported license
© Markus Fiedler

The first part of the last – half – day of the seminar (i.e. Friday morning) was devoted to summaries and discussion of the outcomes of the group works in the plenary. For group 2,

a case study was presented. Group 3's feedback scenario triggered a lively exchange of ideas. The omnipresent challenge of how to provide feedback to one's users was stressed in particular by a representative of a vendor. Further issues addressed were how to make users aware of quality, how to avoid cheating by users, and how to tackle privacy issues when it comes to monitoring user traffic. The discussion regarding group 1's work circled around the notions of experience (as opposed to events, and in relationship to anticipations) and quality (in relationship to *qualitas*; absence of the temporal aspect). In principal, everybody using the term QoE should answer the question "How does my usage of the QoE limit a general QoE definition", as for instance provided in the forthcoming QoE White Paper [1], cf. also Section 5.1.

The final session on Friday morning was devoted to the discussion of generic items of joint interest, such as

- a specific journal on QoE, following the proposal of Lucjan Janowski presented in Section 4.15;
- standardisation (realising the orthogonality of the communities and particularities of standardisation work, such as obsession about details); in this context, an upcoming ETSI workshop on QoE [2] and work on a QoE framework and an communication ecosystem was announced by Sergio Beker;
- follow-up work on the QoE White Paper and the documentation of this Dagstuhl Seminar (with corresponding deadlines);
- ideas for the organisation of a follow-up seminar, such as mobilisation of the intercontinental communities already in the application stage, and the use of 30-second elevator pitches and pre-prepared contributions to the group discussions;
- brainstorming about intermediate activities in order to keep the momentum.

It remains to point out that the reduction of the seminar from five to three days in combination with an extensive discussion of the notion of QoE left (much) less time for the treatment of the instrumentation topic as compared to the initial planning. On the other hand, new ideas for future directions of a deepening of the QoE topic in the Dagstuhl context emanated from those discussions. Two-and-a-half months after the end of the seminar, a set of outcomes can be reported (the Qualinet White Paper; joint work on papers; discussions between academia and industry). Thus, there are good reasons to assume that the "Dagstuhl QoE community" will remain active and visible.

References

- 1 Qualinet White Paper on Definitions of Quality of Experience (2012). European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), Patrick Le Callet, Sebastian Möller and Andrew Perkis, eds., Lausanne, Switzerland, Output version of the Dagstuhl seminar 12181, 2012. <http://www.qualinet.eu> (last seen 2012-07-21).
- 2 ETSI workshop on QoE. http://www.etsi.org/WebSite/NewsandEvents/2012_STQWORKSHOP.aspx (last seen 2012-07-21).

Participants

- Patrik Arlos
Blekinge Institute of Technology –
Karlskrona, SE
- Ake Arvidsson
Ericsson Res. – Stockholm, SE
- Sergio Beker
Huawei Techn. – München, DE
- Katrien R. De Moor
Gent University, BE
- Marcus Eckert
TU Chemnitz, DE
- Sebastian Egger
FZ Telekommunikation Wien, AT
- Markus Fiedler
Blekinge Institute of Technology –
Karlskrona, SE
- Marie-Neige Garcia
TU Berlin, DE
- Riccardo Guerzoni
Huawei Techn. – München, DE
- Dennis Guse
TU Berlin, DE
- Richard John Harris
Massey University, NZ
- Helmut Hlavacs
Universität Wien, AT
- Tobias Hofffeld
Universität Würzburg, DE
- Selim Ickin
Blekinge Institute of Technology –
Karlskrona, SE
- Lucjan Janowski
AGH Univ. of Science &
Technology – Krakow, PL
- Maria Kihl
Lund University, SE
- Khalil Laghari
Télécom SudParis – Evry, FR
- Patrick Le Callet
Université de Nantes, FR
- Sebastian Möller
TU Berlin, DE
- Alexander Raake
TU Berlin, DE
- Peter Reichl
FZ Telekommunikation Wien, AT
- Gerardo Rubino
INRIA – Rennes, FR
- Junaid Shaikh
Blekinge Institute of Technology –
Karlskrona, SE
- Lea Skorin-Kapov
University of Zagreb, HR
- Martin Varela
VTT Technical Research Centre
of Finland – Oulu, FI
- Katarzyna Wac
University of Geneva, CH
- Ina Wechsung
TU Berlin, DE

