# Mapping service quality – comparing quality of experience and quality of service for Internet-based map services

Sara Eriksén, Charlott Eliasson, Markus Fiedler, Stefan Chevul and Annelie Ekelin

Blekinge Institute of Technology, Sweden

*{first name.last name}@bth.se*

**Abstract.** In this paper, we present an on-going research project in which we are focusing on examining how users of map-based services on-line experience the quality of these services when the traffic load is high, and how the users' experiences of acceptable or not acceptable quality can be related to measurable parameters which can be used to manage network traffic and improve technical solutions. The project is a multi- and interdisciplinary project in co-operation between researchers within human work science and informatics, and researchers within telecommunication systems. Additionally, there are two external partners in the project: a provider of Internet-based map services, and a municipality which uses this provider's map services regularly. One of the main methodological issues addressed in the project is how laboratory based, quantitative research methods from research on Quality of Service in the telecommunication systems area can be related to qualitative research methods focusing on workplace- or other live-world based use situations and Quality of Experience as defined by users of the services. How can experiments and studies be designed, and results shared, such that both network traffic measuring and evaluation of user experiences retain their own paradigmatic validity and relevance, while fruitfully informing service design?

**Keywords.** Quality of Service, Quality of Experience, Key Performance Indicators, Use-oriented Design, Service Design

# Introduction

During the past few years, the tendency towards convergence between different forms of information and communication technologies has become increasingly evident. Today, most new mobile phones are designed not only for voice communication but also for sending and receiving text messages and multimedia, and for accessing interactive services offered via the Internet. The Internet has changed the way people, governments and businesses work and operate. Internet-based services are becoming more and more interactive and more and more elaborate. However, different network solutions, variations in network traffic load, as well as the design of different network services and various security measures can make huge differences in individual instances of service performance. Thus, *Quality of Service* (QoS) has become an important issue for service providers as well as for users. Service design, security aspects, network performance issues, as well as the users' perspective, such as user requirements, are all factors of the QoS model, as pointed out by Gustafsson and Fiedler (2004). However, the focus is gradually shifting from what used to be mainly the provider's perspective, primarily highlighting network traffic management issues, to more deliberately including a user's perspective of delivered services, as evidenced by the now frequent use of the broader concept *Quality of Experience* (QoE). This puts new demands on multi- and interdisciplinary research and development in order to cover the broad spectrum of issues relating to QoE and in order to develop new models for understanding how these different aspects interrelate, and what this may imply for service design and service management as well as for research methods and models.

In this paper, we present an on-going research project, *Mapping Service Quality*, in which we are focusing on examining how users of map-based services on-line experience the quality of these services when the traffic load is high, and how the users' experiences of acceptable or not acceptable quality can be related to measurable parameters which can be used to manage network traffic and improve technical solutions. Security measures and how they affect response time are one of several aspects which are being considered in this context. We chose to study map-based services because they are becoming increasingly common and popular. At the same time, they are usually complex, are often used in time- and safety-critical situations and thus put high demands on service performance.

The *Mapping Service Quality* project is a multi- and interdisciplinary pilot project which is being funded during 2007 by the Internet Infrastructure Foundation (.se). It is being carried out in co-operation between researchers within human work science and informatics, and researchers within telecom-munication systems. Besides the researchers from different research groups within the same university, there are two external partners in the project: a provider of Internet-based map services (hereafter referred to as the Map Service

Provider), and a municipality which uses this provider's Internet-based map services regularly (hereafter referred to as the Municipality).

One of the main methodological issues addressed in the project is how laboratory based, quantitative research methods from research on QoS in the telecommunication systems area can be related to qualitative research methods focusing on workplace- or other live-world based use situations and QoE as defined by users of the services. How can experiments and studies be designed, and results shared, such that both network traffic measuring and evaluation of user experiences retain their own paradigmatic validity and relevance, while fruitfully informing service design?

In the following, the telecommunication researchers present the suggested design of a test bed for laboratory based experiments focusing on user experience of varying response time. This test bed design has been jointly agreed upon by all the involved researchers. However, it is mainly focused on QoS and a rather narrow perspective of QoE, even if it allows for a certain amount of "situatedness" of the user, mainly perceived as different levels of stress which can be artificially induced during the experiments.

After this, the human work science and informatics researchers present how they propose to enhance these laboratory based experiments with ethnographic studies focusing on how professional users of map-based services within the Municipality work with these services today and how they perceive the experimental environment set up at the University and the tasks they are assigned to carry out in the laboratory tasks in relation to their actual everyday use of map-based services. The challenge here is currently perceived to be how to bring this rich and mainly qualitative empirical material from "the field" to bear on the quantitative results from the laboratory experiments. Might it have implications for the future design of the next series of laboratory experiments? Might the results from the laboratory experiments have implications for the future design of a follow-up field study? How might our different methods and models merge, if at all?

# Key Performance Indicator "Response Time"

Map services belong to the category of interactive services, *cf.* Fiedler *et al.* (2005). From the perspective of a user, an interactive service should display the desired result as quickly as possible after the user has issued the request. Fast reponse is an important part of what the user considers as good QoE. We define the *Response Time* (RT) as the elapsed time between when a user requests information and when the information is received and displayed by the application. The probability that the response time of the application is less than or equal to a *threshold t*, $\Pr\{RT \leq t\}$, is the main performance parameter, also called *Key Performance Indicator* (KPI). For instance, a service might be

considered to work sufficiently well if Pr{RT ≤ 4 s} = 95%, i.e. if the chance that the response time does not exceed four seconds is 95%. This type of specified service level objective is usually part of a *Service Level Agreement* (SLA), such as telecommunication providers have been negotiating with their customers since the 1980:s. Measuring service quality within telecommunication systems research has traditionally mainly been focused on measuring QoS parameters such as the blocking ratio of calls, independently of actual user perceptions and use situations. In the Internet context, performance measurements were for a long time not considered to be important. They were mostly left to the end users, providers etc., who have had to perform round-trip time measurements and their interpretations on their own account. Hwever, during recent years, the focus has gradually shifted to include user experience of central parameters such as response time.

Several studies have investigated the willingness of users to wait for web-based information, *cf*. Bhatti (2000), Bouch (2000), Rajamony (2001) and Zona (1999). Based on this and Nielsen (1994), Fiedler (2004, 2005) summarises reported thresholds in user perception as follows:

- $t$ > 100 ms: the user notices that the system is not reacting instantaneously;
- $t$ > 1 s: the user's flow of thought is interrupted;
- $t$ > 4 s: the user gets bored;
- $t$ > 10 s: the user's attention is lost.

A study by Hewlett-Packard, *cf.* Bouch (2000), revealed user reactions to artificial delays in a realistic web shop scenario. Upon perceiving latencies, the users quickly became uncertain and reluctant when dealing with monetary transactions ("If it's slow I won't give my credit card number."). A gradual appearance of a delayed web page was preferred, rather than having to wait for the content to be displayed all at once ("As long as you see things coming up it's not nearly as bad as just sitting there waiting and again you don't know whether you're stuck"). Users particularly welcomed being informed about problems ("I think it's great...saying we are unusually busy, there may be some delays, you might want to visit later. You've told me now. If I decide to go ahead, that's my choice.").

Users perceive response times through the situation in which they find themselves. A "relaxed" user usually has more patience than someone in a stressful situation (e.g. finishing work, knowing you will have to run to catch the bus; making an important decision when being forced to do so quickly, etc.). The more "stressed" the user, the lower the threshold $t$ becomes, i.e. the more critical a high QoE becomes. Thus, due to the shifting focus from measuring data transfer within a technological network to measuring user experience, use conditions and situations, i.e. *user stress*, need to be taken in to account. This reduces the KPI and affects the fullfillment of the SLA term Pr{RT≤ $t$}.

The response time itself is influenced by the ICT system, consisting of applications (clients; servers; peers) and networks (devices; links). Any kind of disturbance in the ICT system may cause increases of the response times. Examples of such disturbances are

- competing processes or performance limitations at the client and server side (e.g. in the terminal hardware and software), yielding delays in sending and receiving data
- network-level perturbations such as situations of overload and/or resource limitations on network links and within network equipment, yielding loss and delay when transfering data Data loss entails extra delays because of necessary data retransmissions. Such perturbations are typically considered to be QoS problems

The end user perceives the response time at the end of a supply chain, i.e. each entity along the path between client and server application might add additional delays. The more and worse the QoS problems, the longer the user-perceived response time becomes. It gets closer to, or even crosses, the threshold *t* more frequently, which means a reduction of the user's QoE perception. Such *network stress*, in other words QoS problems, stress the demands on the SLA term $Pr\{RT \leq t\}$ as well. Obviously, the combination of user and network stress can easily lead to a service level which is much too low compared to the SLA. The goal of this work, from the telecommunication systems researchers' perspective, is to quantify these effects and to establish context-dependent relationships between QoS and QoE.

# Experiments

In order to quantify service levels and QoE as functions of response times (reflecting network conditions and QoS) and related thresholds (reflecting user conditions), laboratory-based experiments involving real users will be carried out. These users will be asked to carry out tasks, for which the reponse time is measured at different entitites along the service supply chain. Also, the user reactions will be measured and subsequently correlated with the response time.

## User tasks and conditions

In each test case, the user will be given the same tasks to complete. The same tasks will be assigned in order to make it possible to compare the result, RT, in all the test cases. If the tasks were different, for example involving varying size of the requested data, this would result in varying RT for the test cases even if the test case conditions in all other aspects were identical.

The first task given to the users will be to retrieve a specific map. The second task will be to retrieve further information in the map by clicking a specific symbol on the map. These tasks will be performed both in sequence and separately. Each task will result in a RT. When performing the tasks in sequence, there will be two different RT:s to measure, one for each task. This sequential-tasks case will also include the time it takes the user to analyze the map between the two tasks, resulting in a total time for the test case, consisting of the sum of two RT:s and one AT (Analysis Time).

The participants in these test cases will be professional map service users, employed in the Technical department of the Municipality. They will perform predefined tasks in the laboratory-based test application, and the situation and context will vary for different test cases.

- There will be a basic test case where the user situation is relaxed and there is no specific context. These user conditions will be considered as good.
- Subsequently, the user might be stressed by the context in which the tasks have to be performed, e.g. by urging a decision to be made based on the results from the map handling.

Additionally, the varying network conditions described in the following are intended to contribute to simulating problematic conditions and stressful situations in real life use of map services.

## Network conditions

During the experiments, the conditions of the network will be varied by deliberately influencing the data transmission capabilities. In the first test case, the network QoS will not be affected by any disturbances, i.e. the network conditions will be good. The subsequent test-cases will involve imposing different disturbances on the stream of packets carrying the application data of interest. Such QoS disturbances will include

- Packet loss according to given ratios;
- Packet delays, following different distributions;
- Throughput constraints.

The larger these impacts, the worse the network conditions, the worse the QoS and the worse the QoE. The bad network conditions will be created using a *shaper*, having the possibility to take a controlled impact on the QoS, e.g. by throwing packets according a given loss ratio or by increasing the spacing between packets order to emulate a certain limitation of transmission speed.

As described by Fiedler *et al.* (2005), security affects the QoS through increased processing times and additional overhead. The impact of security on the

QoS must be studied from a user's point of view in order to see how security affects the QoE, although this will be subject of future work.

## Measurement setup

The end-to-end response time will be measured at the user end, at the client. The request and response will be logged and compared to get the response time. This includes delays from the whole system.

There will be several taps in the network. These taps eavesdrop on the traffic and logs when a packet passes it. Figure 1 shows a possible setup. The user issues the request on her/his terminal (arrow 0), which is captured by the measurement point M1. The terminal passes on the request via the network towards the server, thereby passing the network-level measurement point M2. The request is processed by the server and the response is passed on to the client (arrow 2) at the user's site, which then displays the result (arrow 3). It is worth noting that the response time at the server is in general smaller than the response time at the client, i.e. RT(M2) < RT(M1). The user, sitting at the end of the supply chain, perceives the worst response time throughout the whole chain. The difference between RT(M1) and RT(M2) illustrates the impact of the network in-between and its QoS.

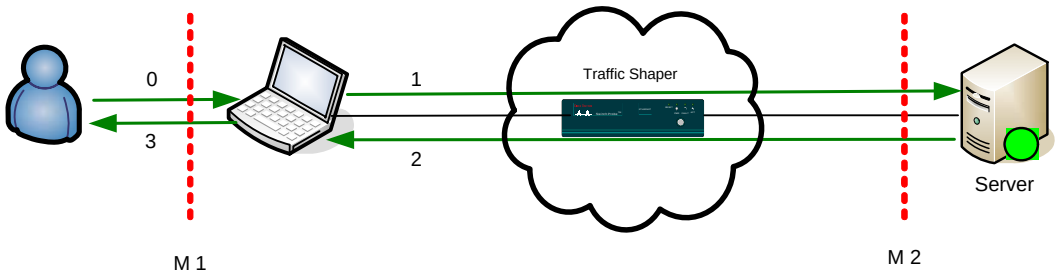


Figure 1. Illustration of measurement points (M1 and M2) for response times.

## Test cases and setup

During the experiment period of this project, several test scenarios will be designed and performed. Each test scenario will have a specific hoped-for result parameter that will be considered during design to make sure that we actually measure what we have set out to measure.

The test cases will be divided into smaller areas such as User and Network, where the conditions can be of different levels of good or bad. There will be scenarios where all conditions are good, where all conditions are bad and also

where the conditions for User and Network are good and bad respectively and the other way around.

The conditions for User and Network have direct impact on the RT and also indirect impact, in the form of perception of the RT. If the User conditions are bad, e.g. in form of stress, then the user will have less patience with variations of the RT. A value that in a relaxed situation is considered to be a normal and accepted RT will, in a stressed situation, perhaps not be acceptable for this specific user. The Network conditions on the other hand, that have a direct impact on the RT, will increase with the worsening of the conditions. This can then lead to values of the RT that will not be acceptable for this specific user.

If a further discussion of these two points of view is held, it can be seen that they are almost each others inverse. When the user has bad conditions his /her threshold for RT values is low, and when the network has bad conditions it results in a high value of the RT. In the case where both network and user conditions are bad, the network will provide a higher RT and the user has lower thresholds than in normal conditions, which will not "converge".

For testing the system used by the Map Service Provider, a mirror will be made in an experimental environment at the University. The mirrored system will then be expanded with a traffic shaper and measurements point.

When the equipment is running properly, it will be moved to a real-life environment at the Municipality. All the experiments will then be performed there, with real-life conditions and professional users.

# The field workers' perspective

If the above is seen as the telecommunication systems researchers' approach, which is mainly quantitative and experiment based, then what follows below may be seen as an explorative and qualitative approach brought to the *Mapping Service Quality* project by the researchers from human work science and informatics. This approach involves starting out with ethnographic field studies among professional users of map-based services, who are employed within the Technical department of the Municipality. During early autumn 2007, we will be carrying out open-ended interviews and workplace observations with users in order to gain insight in to how, and in what situations, these services are used today, and how they are perceived by the users. The interviewees will also be invited to join a user group for future reference and for taking part in laboratory based tests.

# Exploring QoE in situated use of map-based services

In the research area of informatics, there is an on-going shift of focus from systems to services (Dahlbom 2002, Hultgren 2007). QoS is regarded as a complex and multi-layered concept with shifting definition depending on the context. It can be understood as both a support system requirement and a high level requirement (Östlund 2007), though it is often treated as either one or the other, and the two are seldom focused simultaneously and in relation to each other. Similarly, but even more so, QoE in e-service research within the area of informatics is usually understood as complex, multi-dimensional and situated. In a recent review of research on e-service, Rowely summarizes her findings by arguing that it is necessary to go beyond studies of e-service quality dimensions and also take into account the inherent characteristics of e-service delivery and the factors that differentiate one service experience from another, in order to understand e-service experiences (Rowely 2006).

In this article, we do not have the ambition to go further in presenting current theories within the area of e-service, or even to attempt to define more precisely the concepts of QoS and QoE. Ours is a social constructionist and pragmatic approach, in which our primary aim is to explore the meanings of these concepts as they are construed and perceived in everyday work practice and in telecommunication systems engineering research practice, respectively, and thereafter to relate these understandings to each other and to research and methods concerning e-service design and management. Our ethnomethodologically informed, ethnographic approach to studying, in this case, QoE as it is perceived in situated use of map-based services by municipal employees (who are professional and experienced users of these services), is a research approach that we have developed over a number of years of studying use and design of e-services in the public sector (Dittrich *et al*. 2002, Eriksén 2002a, Eriksén 2002b, Dittrich *et al.* 2003). It is strongly influenced and informed by the Scandinavian tradition of participatory design as well as by ethnomethodological understandings of design-in-use of IT (Suchman 1994, Dittrich *et al.* 2002).

So what is new in this case? In the project *Mapping Service Quality*, we are attempting to unpack what is, from our perspective, the "black box" of telecommunication systems. Our ambition is to seriously and extensively relate engineering understandings of QoS and QoE to user experiences of e-service quality in use. In order to challenge our own assumptions, and stretch further across disciplinary boundaries, we are, in this project, tentatively exploring the concept of *contextual engineering* as a way of talking about our aim to merge and rework methods and models for measuring and exploring QoS and QoE from telecommunication systems research on the one hand and informatics and human work science research on the other hand. We are not sure how far, or quite where, this will get us, but we feel, as action researchers involved in a number of on-

going, technology-driven R&D projects, that we need to seriously address the fact that we are located in, and taking action from within, the School of Engineering, at a technical university, which is explicitly profiled towards designing, developing, managing and understanding applied and emerging technologies.

# Discussion

In the introduction, we posed the question "*How can experiments and studies be designed, and results shared, such that both network traffic measuring and evaluation of user experiences retain their own paradigmatic validity and relevance, while fruitfully informing service design?"* This article is a first step in articulating part of a multi- and interdisciplinary dialogue which we need to continue and get deeper into, in order to achieve what we are aiming for in the project *Mapping Service Quality* – a merging and reworking of methods and models for measuring and exploring QoS, QoE and for service research and design. It is obvious, to the reader as well as to us, that we are writing from two very different research perspectives, and that, as yet, we have not been able to develop and present a shared framework, other than the general framework of the project itself. Problematizing the measuring of *t* in the SLA term $Pr\{RT \leq t\}$, and contextualizing it relative to situated use and user perceptions of QoE, is a challenging beginning. However, a large part of the potential quality of this project, as we see it, lies in the active involvement of the Mapping Service Provider and the Municipality (in this case the customer/users). Together, they provide a practice-based touchstone for our theory-building, which we need to develop methods to make good use of. This practice-based knowledge and experience, we believe, will become more visible as we get further in to the project.

# Conclusion

As we are still in the initial stages of both the multi- and interdisciplinary project *Mapping Service Quality* and the multi- and interdisciplinary dialogue around QoS, QoE and the emerging (at least for us) concept of contextual engineering, we feel it is too early to draw any serious conclusions. Let it suffice that we have raised, and hopefully also provoked, one or two serious questions across disciplinary boundaries.

# References

Bhatti, N., Bouch, A. and Kuchinsky, A. (2000), Integrating user-perceived quality into web server design. In *Proceedings of WWW´00*, Amsterdam, 2000.

Bouch, A., Kuchinsky, A., and Bhatti, N. (2000). Quality is in the eye of the beholder: Meeting user's requirements for internet quality of service. Technical Report HPL-2000-4, HP Laboratories Palo Alto, January 2000.

Dahlbom, B. (2002). From *Systems to Service*s. Internet document available at http://www.viktoria.se/~dahlbom/

Dittrich, Y., Eriksén, S. and Hansson, C. (2002). PD in the Wild; Evolving practices of Design in Use. In *Proceedings of the PDC 2002*.

Dittrich, Y., Ekelin, A., Elovaara, P. Eriksén, S. and Hansson, C. (2003). Making e-government happen. Everyday co-development of services, citizenship and technology. In *Proceedings of the 36st Hawaii International Conference on System Sciences - 2003.*

Eriksén, S. (2002a). Localizing Self on the Internet; Designing for "Genius Loci" in a global context. In Dittrich, Y., Floyd C. and Klischewski, R. [eds.] (2002), *Social Thinking – Software Practice*. MIT Press.

Eriksén, S. Designing for Accountability. In *Proceedings of NordiCHI*, Århus, Denmark 2002.

Gustafsson, K., and Fiedler, M. (2004), "E-government: services, needs and user satisfaction". Report, Blekinge Institute of Technology. February 2004. Available at www.bth.se/fou

Fiedler, M. (ed.) (2004). "EuroNGI Deliverable D.WP.JRA.6.1.1: State-of-the art with regards to user-perceived Quality of Service and quality feedback". May 2004. Available at http://www.eurongi.org.

Fiedler, M. (ed.) (2005). "EuroNGI Deliverable D.JRA.6.1.2: Studies of user-perceived Quality of Service within EuroNGI". Jan. 2005. Available at http://www.eurongi.org.

Fiedler, M., Chevul, S., Isaksson, L., Lindberg, P., and Karlsson, J. (2005). "Generic Communication Requirements of ITS-Related Mobile Services as Basis for Seamless Communications", *Proc. of the First EuroNGI Conference on Traffic Engineering.* Rome, Italy

Gleeson, B., Lin, A., Heinanen, J., Armitage, G., Malis A. (2000). RFC2764 *A Framework for IP Based Virtual Private Networks*

Hultgren, G. (2007). *eTjänster som social interaktion via användning av IT-system – en praktisk teori*.[eServices as social interaction via use of IT systems – a practical theory] Dissertation. Linköping University.

Nielsen, J. (1994). *Usability Engineering*. Morgan Kaufman, San Francisco.

Rajamony, R. and Elnozahy, M. (2001). Measuring client-perceived response times on the www. In *3rd USENIX Symposium on Internet Technologies and Systems* (USITS). San Francisco, March 2001.

Rowely, J. (2006). An Analysis of the e-Service Literature: Towards a Research Agenda. *Internet Research*, Vol. 16, No. 3, 2006, pp.339-359.

Suchman, L. (1994), Working relations of Technology Production and Use. In *Computer Supported Cooperative Work*, Vol.2, p 21-39, Kluwer Academic Publishers, Netherlands

Zona Research Inc. (1999). The economic impacts of unacceptable web-site download speeds. April 1999

Östlund, L. (2007). *Information in use – in- and outsourcing aspects of digital services.* Dissertation. Blekinge Institute of Technology.