Enhancing Quality of Experience (QoE) Assessment Models for Web Traffic

Amanpreet Singh[1], Ahmed Mahmoud[1], Andreas Koensgen[1], Xi Li[1], Carmelita Göerg[1], Mehmet Kus[2], Muhsin Kayralci[2], and Jasmin Grigutsch[2]

[1] University of Bremen
Otto-Hahn-Allee 1, 28359 Bremen Germany
{aps,ahmed,ajk,xili,cg}@comnets.uni-bremen.de
[2] OTARIS Interactive Services GmbH
Fahrenheitstr. 7, 28359 Bremen
Germany
{kus,kayralci,grigutsch}@otaris.de

Abstract. Web applications are becoming the key services in today's networks (both fixed networks and mobile networks). Consideration of web service quality has become essential to provide the end users with satisfying Quality of Experience (QoE). In order to evaluate and manage the web quality, methods for QoE assessment are desired to estimate the service quality perceived by the end users. In this paper, we study a number of existing objective quality assessment models for assessing the QoE of web applications, and compare their performance with simulations to find out their individual advantages and limitations to use in practice. Simulation results show that the proposed QoE model can be applied for evaluating the quality of different web sources in the Long Term Evolution (LTE) networks, considering the lossy property of mobile networks. A fitting model is presented to describe the correlation between network Quality of Service and User QoE obtained in subjective lab test. To overcome the shortcomings of the existing models, this paper also proposes an enhanced QoE model which considers the effects of parameters such as page download size and content, browser cache setting as well as the packet losses and connection throughput in quality assessment e.g., page response time. To study other user related aspects in the evaluation of QoE, subjective tests in real systems and environments are planned as the next step.

Keywords: Web, Quality of Experience, Quality of Service, Modeling.

1 Introduction

Web applications have always been the key service with the onset of the Internet in both fixed and mobile networks. To provide a high-quality web service for the users, evaluating and managing the quality properly is extremely important. The normal Quality of Service (QoS) measurements that reflect the technical parameters of a service, however, do not reflect the user's perception of the obtained performance. Therefore, in the recent years attention has widely been paid on the concept of Quality of experience (QoE).

D. Pesch et al. (Eds.): MONAMI 2013, LNICST 125, pp. 202–215, 2013.

QoE is defined by ITU as "the overall acceptability of an application or service, as perceived subjectively by the end-user" [1]. QoE is typically represented using the Mean Opinion Score (MOS) [2], which is an empirical quality scale which ranges from 5 (excellent) to 1 (bad), indicating the quality from the user's perspective of the received service. To measure the perceived service quality, QoE assessment methodologies are indispensable, which are important for service quality management as well as for network planning, monitoring and optimization. The QoE assessment methods can be based on subjective tests or objective QoE models. Subjective evaluation of quality is usually carried out by a test panel of real users. Objective evaluation of quality is performed by applying objective QoE assessment models on behalf of a real user, trying to imitate or predict user perceptions by mapping network level QoS parameters into user level QoE. This provides the operators and service providers with a metric of the user satisfaction with the service.

Web traffic contributes a major part to the overall Internet traffic. According to real time monitoring using Akamai platform that handles about 20% of the world's total web traffic, there are 13,639,235 hits per second, 50,794,152 global page views per minute [3].

The key parameter that governs the user web QoE is the page response time i.e., the time it takes for the web page to download completely after the web link was clicked [4]. In other words, the total response time is defined as the time between issuing a web request to the system until the end result is visible to the user. The response time can be affected by the sum of time it takes to transfer the request to the remote server, the time the remote server needs for satisfying the request and the time it takes to transfer the response to the end user [5].

There are three main limits for the subjective response time as mentioned in (Nielsen 1993) [6]:

- 0.1s is the limit when a user feels that the system responds instantaneously.
- 1.0s is the limit until which a user perceives the page to be uninterrupted.
- 10s are the limit to keep the user's attention focused on the web page.

According to a study by Akamai in 2006 [7]:

- 75% of people would not return to websites that take longer than 4 seconds to load.
- Most of the Internet users rank page-loading time as a priority.

Another study [8] suggests that most users are willing to wait only about 2 seconds for simple information retrieval tasks on the web. A similar conclusion was released by Akamai in September 2009 [9]:

- 40% will leave off the web page if it takes more than 3 seconds to load.
- 47% expect 2 second or less to load a web page.

Even fractions of seconds of response time can have significant effects as Google found that by moving from a 10-result search web page which loaded in 0.4 seconds to a 30 result page loading in 0.9 seconds decreased the ad revenues by 20% [10].

The remainder of this paper is structured as follows. A number of selected objective QoE models for web applications are described in section 2. Section 3 discusses the important aspects that should be considered for precise modeling of the web QoE. The subjective tests that were conducted in a lab environment, their results and corresponding fitting models are presented in section 4. Section 5 describes the simulation scenario that is used to compare the different objective QoE models against the proposed fitting models. Finally, section 6 gives the conclusion and also an outlook.

2 Existing Web QoE Objective Models

Since, the page response time is what a user experiences while opening a web page, it is normally assumed to be the parameter that one should map to the web QoE values. Therefore, most of the web QoE models described in the literature which are mainly based on the page response time [11] [12] are explained in sections 2.1 and 2.2. The page response time is dependent on the design of the respective web page and also on the quality of the connection. One of the simplest forms used to identify the quality of a connection is the data rate that the user gets while opening the web page. Therefore in [13], a QoE mapping model is given based on the different user data rates. All these models are based on their respective subjective test results and we also follow the same norm to propose our new model in section 4.2 and 5.

2.1 Fitting Functional Models

In [11], subjective test results were used to map the impact of the web page response time (T) to web QoE (MOS) by different curve fitting function models such as based on the linear, logarithmic or exponential functions (refer Table 1). To measure the closeness of the fitting functions with that of the subjective QoE results, the coefficient of correlation was calculated for each fitting function. The exponential function as well as the logarithmic functions fit the subjective results best. The logarithmic function also has a high correlation as supported by ITU-T Rec. G.1030 [14].

Table 1. Web QoE functional fitting models [11]

Relation	QoE Model	Coefficient of correlation
Logarithmic	$MOS = -1.426 \cdot \ln(T) + 4.469$	0.994
Linear	$MOS = -0.318 \cdot T + 4.158$	0.983
Exponential	$MOS = 4.836 \cdot \exp(-0.15 \cdot T)$	0.995

The QoE values for the different models are depicted in Fig.1 for varying page response times. For the subjective tests to which the different models were mapped, the QoE values were considered in the normal range of 1 (unacceptable) to 5 (excellent) with an additional level of 0 which identified when the users simply disconnected the session. Therefore in Fig.1, the models map the page response time to the web QoE values ranging between 0 and 5.

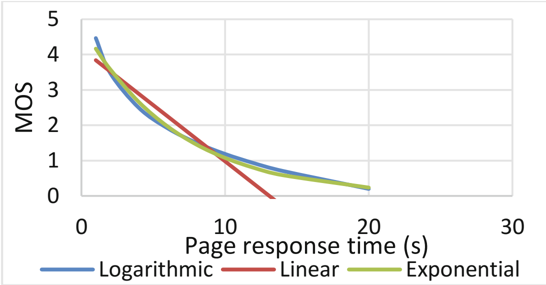

Fig. 1. Mean opinion score (MOS) curve fitting functional models based on page response time

2.2 Fitting Models Based on the Lorentzian Function

Like most of the Quality of Experience models, in [12] an experimental survey of the subjective quality as perceived by the end user was used to develop a mapping function based on the Lorentzian function [13] between the web page response time, T (measured in seconds)) and the user's quality experience (*MOS*). This model will be referred as the PRT non-linear model in this paper; it is defined as given in Eq. (1).

$$MOS = 5 - \frac{578}{1 + \left(11.77 + \frac{22.61}{T}\right)^2} \tag{1}$$

One of the important factors that influences the page response time is the user data rate for the web connection. The authors in [15] formulated a mapping function between the web QoE and the user data rate, r (measured in kbps) with which the web page was downloaded. This model will be referred as UDR non-linear model in this paper; it is defined as specified in Eq. (2).

$$MOS = 5 - \frac{578}{1 + \left(\frac{r + 541.1}{45.98}\right)^2} \tag{2}$$

The QoE values for these two models based on the Lorentzian function are depicted in Fig. 2. According to these results, a 10s limit exists for a user to fall into the category of being completely dissatisfied i.e., a mean opinion score of less than 2. A similar QoE value is obtained for a user data rate of around 100kbps. The user data rate can also be seen as the goodput of the web connection.

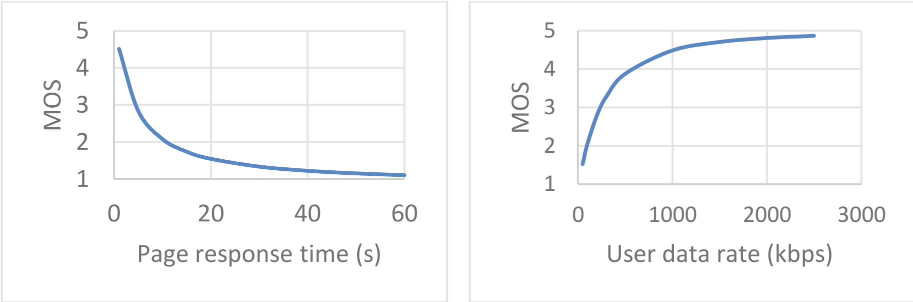

Fig. 2. Mean opinion score (MOS) curve fitting Lorentzian models w.r.t. a) Page response time (PRT non-linear model) and b) User data rate (UDR non-linear model)

3 Important Aspects for Web Browsing Experience

When a user visits a webpage, the browser first performs a DNS lookup to obtain the IP address of the web server. It then establishes a TCP connection with the server before it starts to download the main webpage. The main page may embed many web objects, including CSS, Java scripts, and images, which sometimes are hosted by servers in multiple domains. In that case, the browser has to perform more DNS lookups, establish multiple connections to different servers, and download objects in parallel. This process continues recursively until all the objects are downloaded. Thus the page response time depends on factors such as DNS lookup time, TCP setup and transfer (including the slow start phase and other congestion control aspects), and the client browser itself (the browser's response time and rendering speed of the displayed page). In addition, contents of most of the popular web pages are dynamic i.e., the structure and embedded objects, e.g. advertisements, may change over time.

Some of the important aspects to evaluate web browsing performance (or page response time) are listed here:

- **Total page size:** the page response time is monotonically increasing with the downloaded page size. A web page usually contains several embedded objects and all these add to the total page size.
- **DNS lookup time:** the web page's embedded objects may be hosted under different domains leading to a significant delay caused due to the DNS lookup process. Fig. 3 shows a snapshot of a web page (www.radiobremen.de) download, where the DNS lookup time is depicted for two embedded objects that are hosted at two different domains [16]. Fig.3 also shows other types of delays such as the connection setup delay, the time to first byte and finally the download time.
- **Browser concurrency:** Browsers support concurrent TCP connections within the same domain to improve download efficiency. If no network bottleneck exists, a higher concurrency means better utilization of bandwidth and shorter page response time. The maximum number of concurrent TCP connections within a domain varies for different browsers [17].

- **Cache strategy:** To speed up the web page response, browsers and web page designers allow for caching the web page content for a certain time. Thus if the same page is re-visited then the page can be reloaded from the cache or if only certain parts of the page are updated in the interactive browsing session then only the incremental information (embedded objects) are retrieved from the server.
- **Effective page size:** Due to the cache effect, during an interactive browsing session, a user may click on several embedded links and observe a favorable performance as in most cases the whole page is not required to be retrieved from the web server. The amount of downloaded data is referred to as effective page size. This reduces the load on the network as well as makes the response time shorter and hence should be considered whenever browser cache is used instead of the total page size.

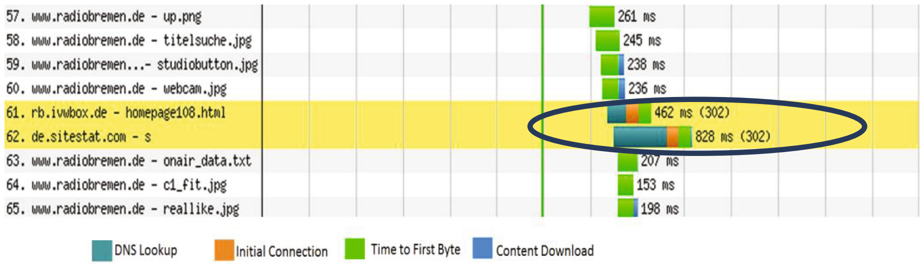

Fig. 3. Waterfall view shows DNS lookup, initial connection and object download time for different domains

4 Subjective Test for Web QoE

The concept of Quality of Experience (QoE) [1] is an approach to measure the network performance from the user's point of view, which means that degradations or disruptions of the service are being considered. The challenge of QoE measurements is that the perception of the user is a subjective phenomenon which varies between different users. Hence, in order to design a model to map QoS to QoE, tests with multiple users have to be performed. This section discusses QoE measurement experiments which were performed in the scope of the investigations discussed in this paper.

4.1 Experimental Setup

The hardware setup is shown in Fig. 4. Six probands can be tested at a time; each of them is provided a netbook (Lenovo S10-2) with a screen size of approx. 26 cm (10 inches) and a screen resolution of 1024×600 pixels. The netbooks are connected via an Ethernet switch to each other as well as to the Internet. The switch is manageable which allows configuring the uplink and downlink speed separately for each port. The switch is also connected to the Internet by a router which is required for the web browsing test. A controlling computer inside the network allows the automatic configuration of the switch and the start of required software on the netbooks.

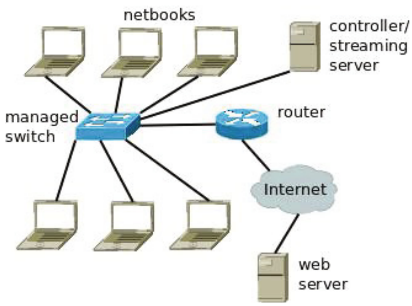

Fig. 4. Experimental setup for the Web QoE subjective tests

For web browsing, the netbooks connect to a preconfigured server on the Internet (www.radiobremen.de/vier) which provides a web page with a mix of text and photos. The photos play an important role in the experiment. They are objects embedded into the web page and have a considerable file size, so the speed by which they appear on the terminal screen provides a good way of ranking the connection quality. The experiment is repeated with different link speeds which are given in the results section. After the end of each experiment, the user has to fill in a questionnaire where the QoE of the service is ranked and some additional comments about the individual impression can be given.

The questionnaires are provided in an electronic way so that the data can be automatically collected inside a database which allows a flexible evaluation according to certain criteria. E.g., for the determination of the QoE/QoS relationship, it is important to consider whether a user is an “expert” with sound knowledge about networking devices or an inexperienced user who might use the Internet only occasionally.

During the preparation phase of the experiment, link speeds suitable to run experiments with the different services were determined by friendly-user tests (FUT). Five speeds have been selected individually for each service, where the number of five is chosen according to the Mean Opinion Score model as described in section 1. For one of the five speeds, the experiment is run twice, once after an experiment with a higher speed and once after a lower speed. The aim is the identification of memory effects, i.e. it is considered that the user has a different perception when coming from an experience worse than the current one or when coming from a better one. During the FUT, the tendency that different users sometimes have different opinions about the same technical service quality already could be observed, which highlights the need to perform the experiment with a large number of users to find the Mean Opinion Score (MOS) for each speed.

4.2 Subjective Test Results

In the subjective tests, 35 users took part which with equal gender distribution (18 male, 17 female). Most of the test users were either university students (21) or employees (10) with sufficient knowledge about the Internet and related terminology to web related applications. Fig. 5 depicts the average number of hours spent per day by the test users for different applications on the Internet. It is clear that web surfing and e-mail are two most often used applications.

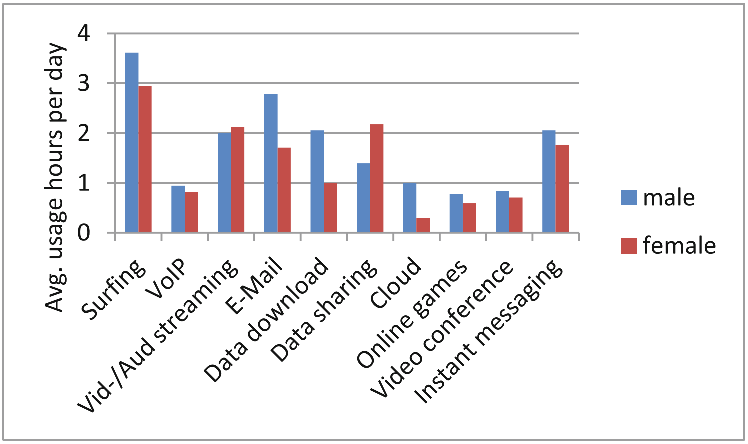

Fig. 5. Average number of hours spent per day on the Internet for different activities

Fig. 6 depicts the QoE results obtained for different link speeds. The test for a link speed of 512kbps was done twice, once after the best speed of 2Mbps and then again after the worst speed of 256 kbps. Both the set of results are not too much apart though the one that follows the speed of 2 Mbps gives a slightly better MOS value.

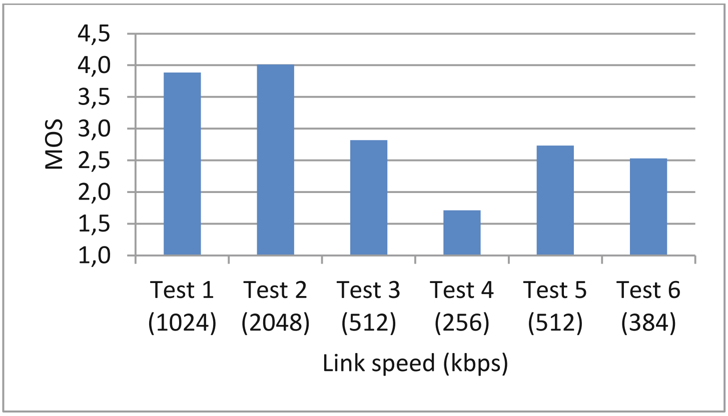

Fig. 6. Web QoE values for the different link speeds in the order of performed tests

Fig. 7 indicates the frequency of the different Individual Opinion Score (IOS) values that the test users gave for different link speeds. As expected, different users rate the quality differently but the trend shows a shift towards higher MOS values for larger link speeds.

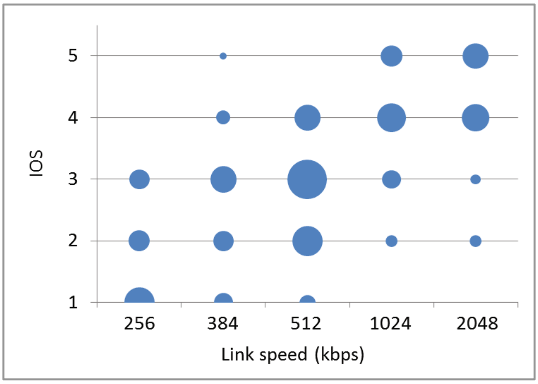

Fig. 7. Number of occurrences (indicated by bubble area) of Individual Opinion Score (IOS) vs. link speed

4.3 Fitting Model for Web QoE

The MATLAB curve fitting toolbox provides graphical tools and functions for fitting curves and surfaces [18]. Regression analysis can be conducted using a library of linear and nonlinear models or by providing custom equations like we did here with the Lorentzian, power and logarithmic (Log_our) functions depicted in Eq. (3), (4) and (5), respectively to fit the subjective web QoE results for varying user data rates.

$$MOS = 5 - \frac{500}{1 + \left(\frac{r + 1271}{123.2}\right)^2} \tag{3}$$

$$MOS = -104 \cdot (r^{-0.6168}) + 5.016 \tag{4}$$

$$MOS = 1.1592 \cdot \ln(r) - 4.6099 \tag{5}$$

Table 2 evaluates the goodness of fit based on the following metrics [19], [20]:

- **Sum of Squares Errors (SSE):** the total deviation of the data values from the fitting values. SSE values closer to 0 indicate better fitting results.
- **R-Square:** the square of the correlation between the data values and the fitting values. R-square values range from 0 to 1, with 1 indicating a perfect fit.
- **Adjusted R-Square:** it is considered to be one of the best goodness statistics for the fitting quality as it considers additional coefficients. Adjusted R-square is particularly useful in the feature selection stage of model building. Unlike R-square, the adjusted R-square increases only if the new term improves the model more than would be expected by chance.
- **Root Mean Squared Error (RMSE):** this statistic is also used to measure the difference between the fitting values and the data values with a better fit having RMSE value closer to 0.

From Table 2, it can be seen that the power fitting model gives the best results, and a comparison between the subjective web QoE results and the fitting model results is depicted in Fig. 8.

Table 2. Web QoE fitting models goodness of fit statistics

Fitting method	SSE	R-Square	Adjusted R-square	RMSE
Power	0.05992	0.9845	0.9689	0.1731
Log_our	0.1971	0.9489	0.9319	0.2563
Lorentzian	0.1708	0.9557	0.941	0.2386

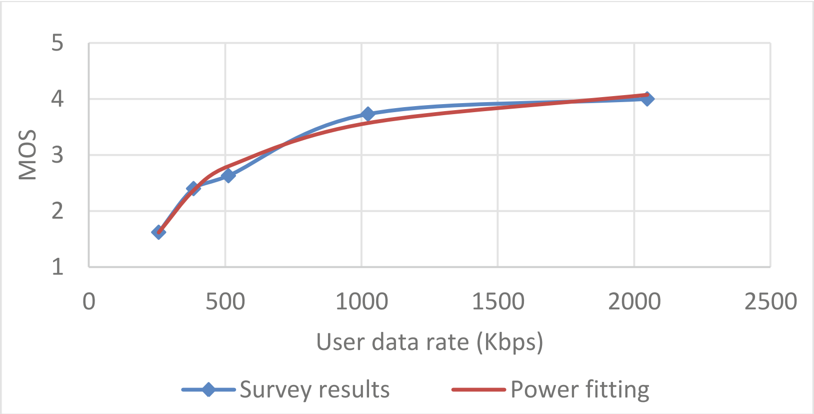

Fig. 8. Power fitting model for the obtained web QoE subjective test results

5 Simulation Scenario and Results

This section discusses the results obtained by the simulation of an LTE Scenario in the OPNET simulator. To realize controlled packet losses and link delay, an impairment object is added on the link between the eNB1 and router R2, as depicted in Fig. 9.

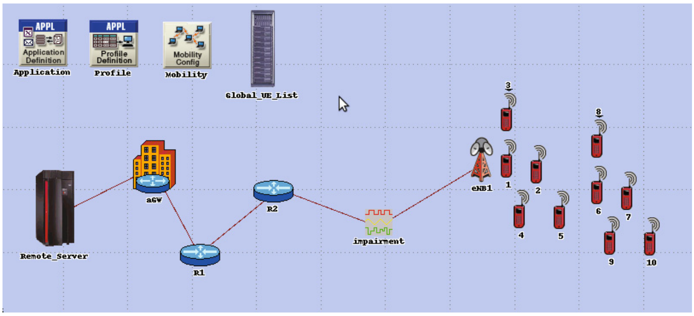

Fig. 9. Simulation scenario for a LTE network in OPNET simulator

Fig. 10 depicts the variation of the user data rate and the corresponding page response time for different packet loss rate configurations in the LTE scenario described in Fig. 9. The web QoE results for different fitting models with respect to varying packet loss rates and link delays are shown in Fig. 11 and 12, respectively. The logarithmic curve is the fitting model taken from Table 1, while the PRT non-linear and UDR non-linear models are described by equation (1) and (2), respectively. The Lorentzian and power fitting models are the ones obtained to fit our subjective web QoE results and are described by Eq. (3) and (4), respectively.

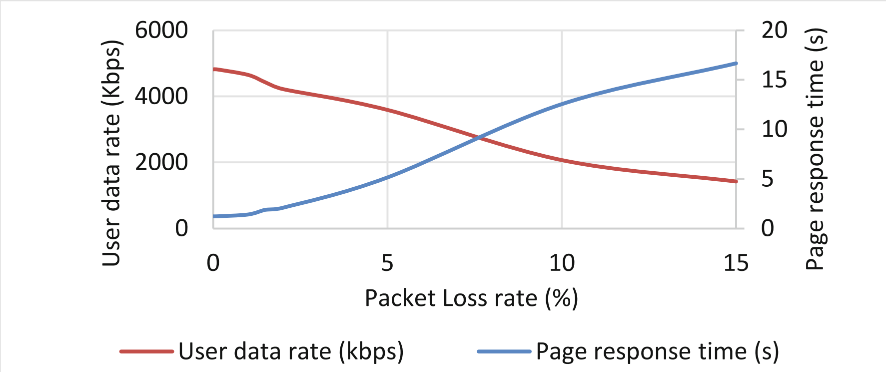

Fig. 10. Effect of the packet loss rate on the user data rate and the web page response time

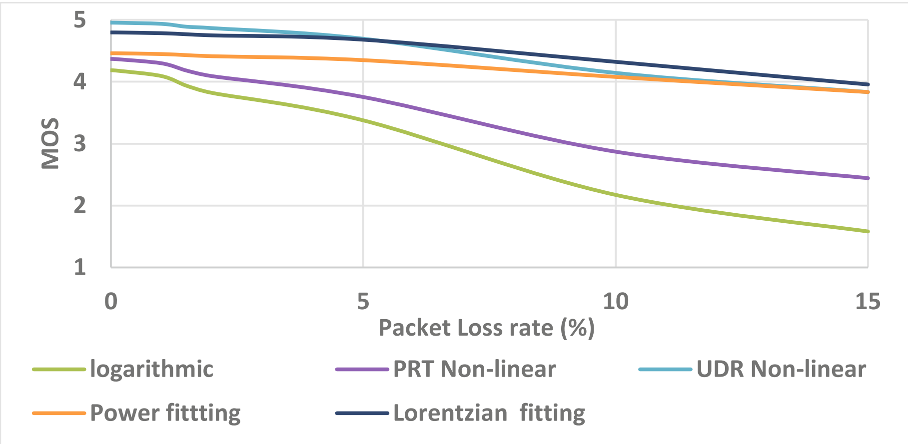

Fig. 11. Effect of the packet loss rate on the web QoE for various models

From Fig. 11 and 12, it can be seen that the UDR non-linear model is too optimistic while evaluating the web QoE whereas the PRT non-linear model is too conservative with the logarithmic curves (from Table 1) being the most conservative.

The missing link between the user data rate (UDR) and the page response time (PRT) based model is the effective page size, *ps* (in kB) of the downloaded page. Therefore, the Lorentzian fitting model described in Eq. (3) has been extended as

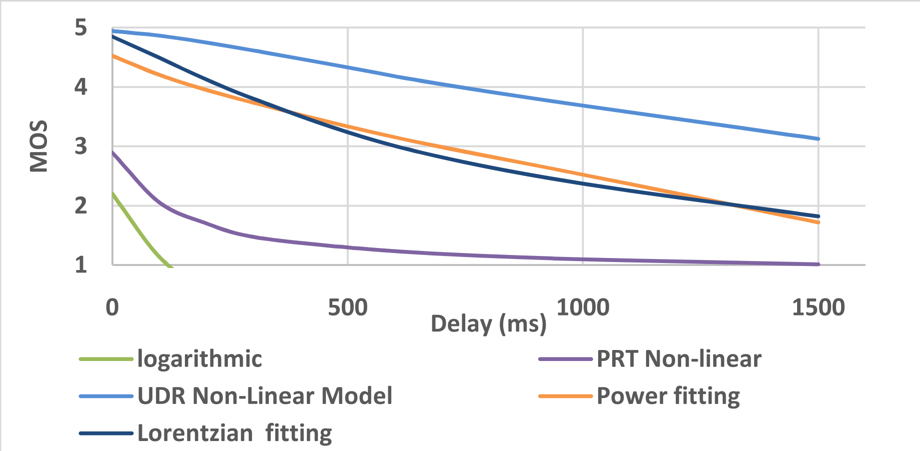

Fig. 12. Effect of the link delay rate on the web QoE for various models

specified in Eq. (6) and will be referred to as the "Final Proposed Model based on the Lorentzian function" (FPML):

$$MOS = 5 - \frac{500}{1 + \left(10.316 + \frac{2.841 \cdot r}{ps}\right)^2} \tag{6}$$

The UDR non-linear model considers an average page size of 125 kB while the page sizes have increased over the years, being 312 kB in 2008 and 1114 kB in 2012 [21]. The web page (www.radiobremen.de/bremenvier) has on average an effective page size (downloaded over the network) of about 350 kB due to the Firefox browser cache settings. Therefore the subjective test results curve is much closer to the web QoE results for year 2008 obtained by the page size and the user data rate dependent Lorentzian fitting curve in Fig. 13. The power fitting model is also a good estimate but

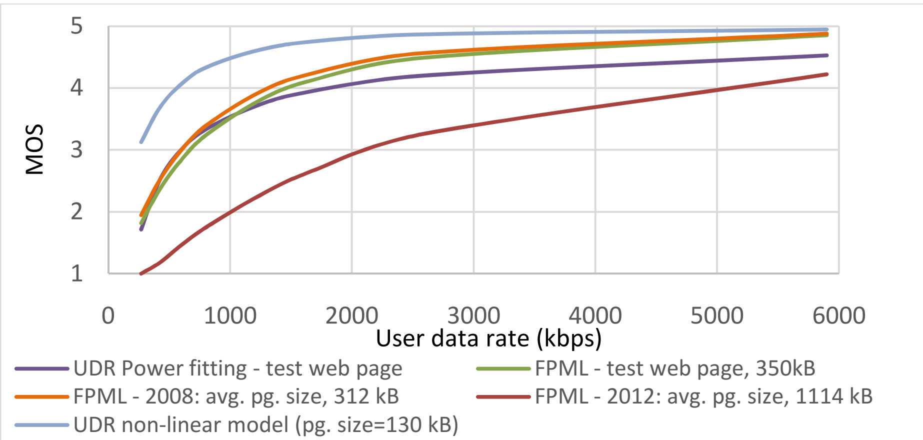

Fig. 13. Effect of user data rate and web page size on web QoE

like the UDR non-linear model, it cannot be used for different web pages as they differ in their size. As the next step, we will extend the power model to also consider the effective page size in addition to the user data rate.

6 Conclusion and Outlook

A number of existing objective models for assessing the QoE of web applications was studied and compared for different QoS parameters such as packet loss rate and link delay. To improve the existing objective QoE models or develop new models, different aspects related to web applications were identified and for validation subjective tests were carried out.

The aim of the subjective tests was to understand the human perception of quality with respect to the Key Performance Indicators (KPI) defined by ETSI such as the user data rate. After each controlled test, the user was asked to answer a questionnaire to ascertain the QoE as well as other related aspects. Based on the obtained results, different fitting (objective) QoE models were evaluated with an LTE scenario in an OPNET simulation. The model given in Eq. (6) that considers both user data rate and effective page size is proposed as the most suitable one to model web QoE objectively.

The knowledge gained from the lab tests will be used to roll out a larger campaign in the form of field tests that will be conducted using real networks based on various technologies e.g. xDSL, UMTS, HSPA, LTE, etc. The persons under field test install an app on their smartphone which measures different technical (QoS) parameters i.e., the connection quality as well as at which time different services are used and where the person is located while using the service. In this way, it can be identified whether the environment and context of the usage (e.g. at home, in the stadium, on the way, at the department store or at work) will affect which services are used and how they are rated in terms of QoE. After a completed service, the app presents a small questionnaire to the user to enquire the QoE rating. Thus the information about how the service quality is perceived by the user in the respective context and technical link quality. The results obtained from the field test will be used to further refine the proposed objective QoE model for web and also consider other aspects such as DNS lookup time and other related delays.

References

1. ITU-T definition: Quality of Experience (QoE) Study Group 12 Geneva, January 16-25 (2007)
2. ITU-T recommendation P.800: Methods for subjective determination of transmission quality (August 1996)
3. Akamai.com (2013), Akamai Homepage, `http://www.akamai.com/` (last accessed: June 6, 2013)
4. Egger, S., Hossfeld, T., Schatz, R., Fiedler, M.: Waiting times in quality of experience for web based services. In: 2012 Fourth International Workshop on Quality of Multimedia Experience (QoMEX), July 5-7, pp. 86–96 (2012)

5. Fiedler, M.: 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), http://eurongi.enst.fr/archive/127/JRA611.pdf (last accessed: June 6, 2013)
6. Nielsen, J., Hackos, J.: Usability engineering, vol. 12518406. Academic press, Boston (1993)
7. Fiona Fui-Hoon, N.: A study on tolerable waiting time: how long are web users willing to wait? Behaviour & Information Technology 23(3), 153–163 (2004)
8. Jupiter Research Report: Retail Web Site Performance: Consumer Reaction to a Poor Online Shopping Experience (2006), http://www.akamai.com/dl/reports/Site_Abandonment_Final_Report.pdf (last accessed: June 6, 2013)
9. Akamai Reveals 2 Seconds as the New Threshold of Acceptability for eCommerce Web Page Response Times, Press Release (2009), http://www.akamai.com/html/about/press/releases/2009/press_091409.html (last accessed: June 6, 2013)
10. Battelle, J.: The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture. Portfolio (2005), http://perspectives.mvdirona.com/2009/10/31TheCostOfLatency.aspx (last accessed: May 17, 2013)
11. Shaikh, J., Fiedler, M., Collange, D.: Quality of Experience from user and network perspectives. Annals of Telecommunications 65(1-2), 47–57 (2010)
12. Ameigeiras, P., Ramos-Munoz, J.J., Navarro-Ortiz, J., Mogensen, P., Lopez-Soler, J.M.: QoE oriented cross-layer design of a resource allocation algorithm in beyond 3G systems. Computer Communications 33(5), 571–582 (2010)
13. Johnson, N.L., Kotz, S., Balakrishnan, N.: Continuous Univariate Distributions, ch. 16, vol. 1. Wiley, New York (1994)
14. ITU-T, G.1030: Estimating end-to-end performance in IP networks for data applications (2005)
15. Ameigeiras, P., Ramos-Munoz, J.J., Navarro-Ortiz, J., Ramiro-Moreno, J., Lopez-Soler, J.M.: QoE Evaluation of Scheduling Algorithms for NRT Services in LTE. In: Proc. ICT-MobileSummit 2009, Santander, Spain (2009)
16. Meenan, P.: WebPagetest – Website Performance and Optimization Test, http://www.webpagetest.org (last accessed: June 6, 2013)
17. Huang, J., Xu, Q., Tiwana, Q., Mao, Z.M., Zhang, M., Bahl, P.: Anatomizing application performance differences on smartphones. In: Proc. ACM/USENIX Int. Conf. Mobile Systems, Applications, and Services (MobiSys). ACM, San Francisco (2010)
18. Mathworks: Curve Fitting Toolbox, http://www.mathworks.de/products/datasheets/pdf/curve-fitting-toolbox.pdf (last accessed: June 6, 2013)
19. Rawlings, J.O., Pantula, S.G., Dickey, D.A.: Applied regression analysis: a research tool. Springer (1998)
20. Hyndman, R.J., Koehler, A.B.: Another look at measures of forecast accuracy. International Journal of Forecasting 22(4), 679–688 (2006)
21. Average Web Page Size Triples Since (2008), http://www.websiteoptimization.com/speed/tweak/average-web-page/ (last accessed: May 22, 2013)