International Journal of Computer and Information Technology (ISSN: 2279 – 0764)
Volume 02– Issue 05, September 2013

An Enhanced Analytical Framework Integrating aspects of Network and Application performance, Psychology, and Content in QoE Assessment for Web browsing

Le Thu Nguyen*, Richard Harris
School of Engineering and Advanced Technology
Massey University
Palmerston North, New Zealand
Email: *L.T.Nguyen {at} massey.ac.nz

Amal Punchihewa
Asia-Pacific Broadcasting Union (ABU)
Kuala Lumpur
Malaysia

Abstract—**Evaluation of Quality of Experience (QoE) for web browsing is a challenging topic as it involves many different factors related to human as well as technical issues. To understand such a large number of human and technical metrics is both difficult and tricky. Two main objectives of the research presented in this paper are to propose and develop a new QoE assessment approach based on a comprehensive set of key metrics for web browsing. The paper starts by evaluating each metric related to users' score before integrating them all into the assessment. Then, the proposed QoE assessment of Web based services uses a comprehensive set of metrics in order to construct a full and detailed understanding of users' experiences in a web based service environment. Finally, the paper presents an investigation into the applicability of a mixed effect model for prediction of QoE based on the proposed set of metrics. The analysis of objective factors and Psychological Support of Content is presented with an emphasis on fixed and random effects. The analysis and validation of QoE assessment of Web based traffic using both objective parameters from networking, application, and content perspectives have shown the efficacy of the proposed assessment. The paper aims at getting better understanding of how users experience web browsing by applying a comprehensive set of metrics which has been computed to determine the overall web user QoE.**

Keywords-Quality of Experience; QoE for web based services; user's perception;

I. Introduction

User Experience terminology was created by Dr. Donald Norman, who was the first person to describe the importance of user-centered service design [1]. Recently, this term has become popular for researchers and practitioners in order to identify the means for providing better service to users. Over recent decades, the term Quality of Service (QoS) has been used as the principal descriptor for specifying the performance quality of both circuit and packet switched networks and, in particular, Internet Protocol (IP) based networks. A multitude of QoS characterisations have been studied in various contexts. Recently, a new study area has been proposed with the aim of interpreting end-to-end quality in the proper sense of including human users as being at the start and at the end of a communication chain. Thus, the notion of Quality of Experience (QoE) was born.

There are many different sources for a general definition of QoE, but all definitions express the fact that QoE is subjective in nature and based on human opinions. According to Kalevi Kilkki [2], "QoE, also known as "Quality of Experience" is a subjective measure of a customer's experiences with a vendor. It is related to, but differs from, QoS, which attempts to objectively measure the service delivered by the vendor." According to the International Telecommunication Union (ITU) [3], "QoE is defined as the user's perception of the acceptability of an application or service". Thus, an assessment of QoE may be influenced by a user's expectations and pre-conceived concepts. Subjective assessment for measuring the QoE 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.

Recent studies involve measurements both objectively and subjectively for a user's perception. For example, a view from the European Telecommunications Standards Institute (ETSI) [4], "QoE is a measure of user performance based on both objective and subjective psychological measures of using an ICT service or product."

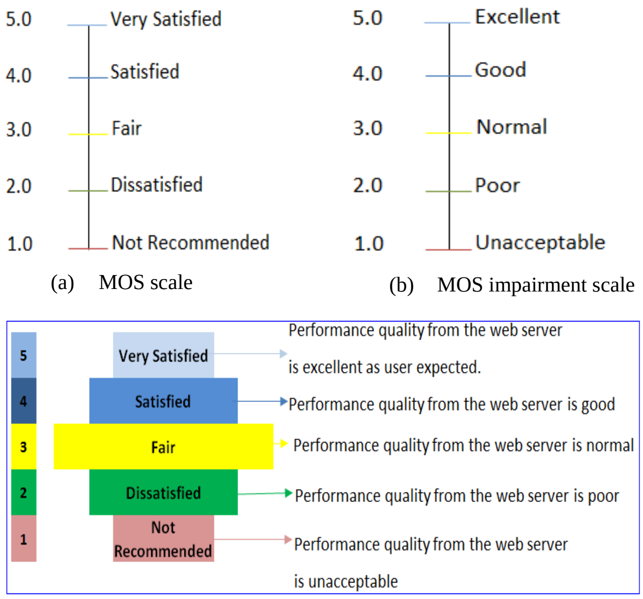
Since QoE relates to a user's experience, it partly involves a form of psychological measurement (subjective); however, it is important to telecommunication service providers to express QoE in relation to the networks and equipment delivering the service which are objective in nature. By combining both the experience of users (subjective) and measurements (objective), QoE may be more reliably measured and estimated. QoE assessment of web browsing has been studied and published by many authors including [5] [6] [7]. The first recommendation for QoE assessment of web browsing was published in the ITU-T standard known as G. 1030 [8].

In 2002, Khirman and Henriksen [5] discussed a relationship for objective networking service conditions and an objective user perception to measure QoE. Human satisfaction of an HTTP service or web browsing is affected by two main network QoS parameters those are bandwidth and latency. The results of their analysis concluded that those factors play a crucial role in end-user satisfaction. Fielder et al., mentioned a generic quantitative relationship between QoS and QoE called the IQX hypothesis [6]. The IQX hypothesis shows a QoE-QoS relationship whereby measured QoS values are inserted into a corresponding exponential QoE formula to manage and control QoE. However, the authors have not considered the practicality of combining the two QoS factors therefore the relationship is an example of singular separate impact factors.

The G.1030 recommendation published in 2005 [8] provides a way to estimate end to end performance in IP based networks for services including web browsing. This recommendation is a starting point for early QoE studies of web-based services. With limited network information, a user's behaviour is not matched fully in G.1030 and this points to a shortcoming of the approach [9]. Human factors have not been taken into account in this standard.

In recent literature from ETSI [4], "QoE is user-centered, expressed in technical QoS measure, and based on both subjective and objective psychological measures". They concluded that there are other effects on QoE assessment from human factors that should be further studied. However, recent studies led to a more complicated QoE definition, in which the definition is related to more areas such as content, network, device, individual personality, etc. For example, a view from the Qualinet Group [10] is, "QoE is the degree of delight or annoyance of the user of an application or service. It results from the fulfilment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the user's personality and current state. In the context of communication services, QoE is influenced by service, content, network, device, application, and context of use".

In this paper, firstly we evaluate the effect of both key metrics from networking and application performance and a key metric from Content that contribute to QoE measurement, then we evaluate all the metrics contributing to user's experience of web browsing. The paper concludes with an investigation of the application of a mixed effect model and its evaluation for the QoE assessment.



(a) MOS scale (b) MOS impairment scale

(c) MOS and its score meaning

Figure 1. MOS scale and its score meaning in the experiment

II. EXPERIMENTAL DESIGN

A. User score-Mean Opinion Score

An experiment was undertaken in which the scores were collected from web browsing users in the form of Mean Opinion Score (MOS) recommendations [11]. MOS are based on the traditional five level scale {1; 2; 3; 4; and 5} that links these numerical values to a user's perception as defined by {Not recommended, Dissatisfied, Fair, Satisfied, and Very Satisfied} as described in Figures 1a and 1b.

The meaning of each MOS score used in the experiment is shown in detail as Figure 1c; when users indicate a score of 5 or "Very Satisfied", it means that performance quality from the web server was excellent as the user expected. Similarly, when users enter a score of "4", "3", "2" or "1", it means that performance quality from the web server was good, normal, poor, or unacceptable respectively.

B. Experimental Setup

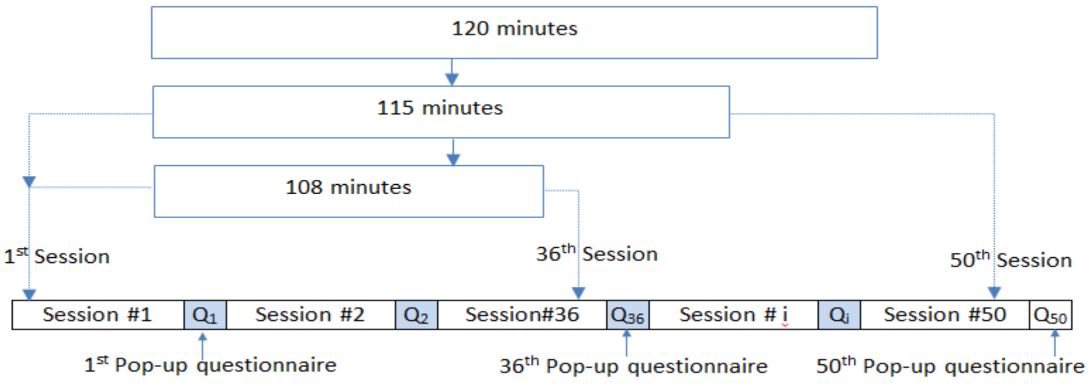


Figure 2. Controlling Session

The experiments for this paper were carried out using the networking infrastructure of a test-bed located at Massey University in Palmerston North. The network topology of the test-bed consists of seven routers and seven switches. The network performance can be changed to different controlled environments by using Netem [12]. The limited number of requests per second and conditional rules were controlled by using mod qos [13] which is the quality of service module for the web server. The experiment was designed and carried out involving human subjects and subsequent analysis. The design of this experiment was based on the Taguchi approach [14] [15]. It is noted that Taguchi refers to experimental design as "off-line quality control" because it is a method of ensuring good performance in the design quality into products and processes [16]. As it is not possible to find a large sample size in order to investigate the impact on parameters thought to be relevant to QoE, our experiments were designed around the Taguchi robust method [15] [16]; which aims to make a process less variable in the face of variation over which we have little or no control. That ensures the number of controlling experiments will be satisfied for the case of limited resources such as a minimum timeframe and number of subjects.

It should be noted that this is a controlled environment in which the network is isolated from the true Internet and network performance is carefully manipulated in order to enable systematic observation of user behaviour without problems caused by miscellaneous external events that might distort the results of our measurements.

To reduce boredom and memory effects, a session known as a controlling session was introduced into our experiment as described by Figure 2. The experiment was carried out for a period of approximately 108 to 115 minutes with 36 to 50 different sessions in which each session was changed at roughly two to three minute intervals. Its structure is shown in Figure 2. There were 21 users in the experiments involving networking and application performance assessments of the users' perception. 12 different users took part in the experiment with a four-level Content factor integrated. All of the users were students at Massey University covering range of demographics.

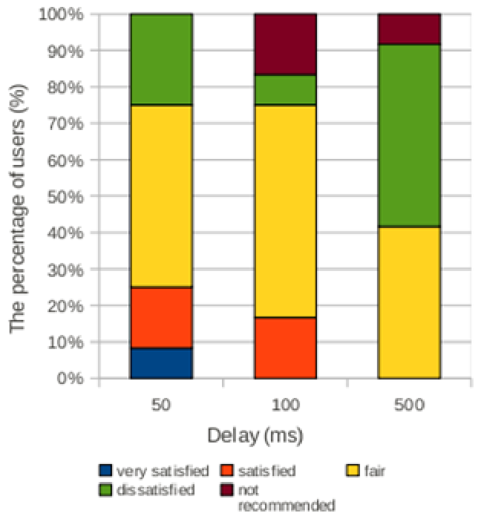


Figure 3. User's Experience for different Delay Scenarios

While users accessed the prepared website, the users' actions were captured to see the behaviour that they exhibited through their selection of hyperlinks on the displayed web page, together with the frequency and nature of these selection activities.

Download time of each web page was estimated using a total combined file size of all items (text, images, graphics, flash objects, flash movies, etc.) which involve embedded information on the web page and the web page itself (HTML/PHP layout code, text, meta tags, etc.).

The following four major tasks have been achieved:

- Task I: To determine a key metric of network performance on users' scores.
- Task II: To determine key metrics of network and application performance on users' scores.
- Task III: To determine a key metric of psychological effect on users' scores.

- Task IV: To evaluate an integration of all key metrics on users' scores.

III. Task I: To Dertermine A Key Metric Of Network Performance On Users' Scores

A. Aim of Delay testing

A key metric of network performance has been chosen as delay [17]. The purpose of this section is to see whether delay can affect a user's perception of quality.

B. Numerical Results

In the existing literature, there are some important contributions on the relationship between networking performance and actual mean opinion scores as mentioned in section I.

The numerical results in our experiment clearly show that users can recognize changes in networking performance involving delay via their scores as illustrated in Figure 3. It depicts the percentages of users' scores for a range of different delay values. It can be observed that, when the network is operating appropriately (delay is less than 50ms as controlled in the local networking environment), it can been seen that 75 percent of users feel "very satisfied", "satisfied", and "fair" in their evaluation of system performance.

C. Discussion

It is observed that the users' scores in the experiment follow our intuition or obey a "rule of nature". For example, the users' scores tend to be "Dissatisfied" when we increase the delay beyond a certain threshold. However, some abnormalities are observed within the specific trends of "Fair" and "Not Recommended" levels, showing that users cannot accurately distinguish between performance delay levels as they deteriorate beyond normally acceptable limits.

Our concept of "Rule of Nature" as shown in Figure 5a is related to users' scores for values of delay itself or for both increasing values of delay and packet loss simultaneously [18]. In the case of both increasing values of delay and packet loss, it can be seen that the users' scores directly reflect the controlled network parameter changes. An increase in the proportion of "Dissatisfied" and "Not Recommended" categories is noticed when both delay and packet loss values are increased. On the other hand, under the same network conditions, the figure also shows the corresponding reduced proportions for the "Very Satisfied", "Satisfied", and "Fair" categories.

However, we claim that assessment based on networking parameters and MOS *only* may be not adequate for QoE assessment. MOS is subjective and depends upon objective information. As seen in Figure 3, the difference between "Satisfied" and "Dissatisfied" can easily be seen by changing network conditions. However, other adjacent scores such as "Dissatisfied" or "Not Recommended" are difficult to

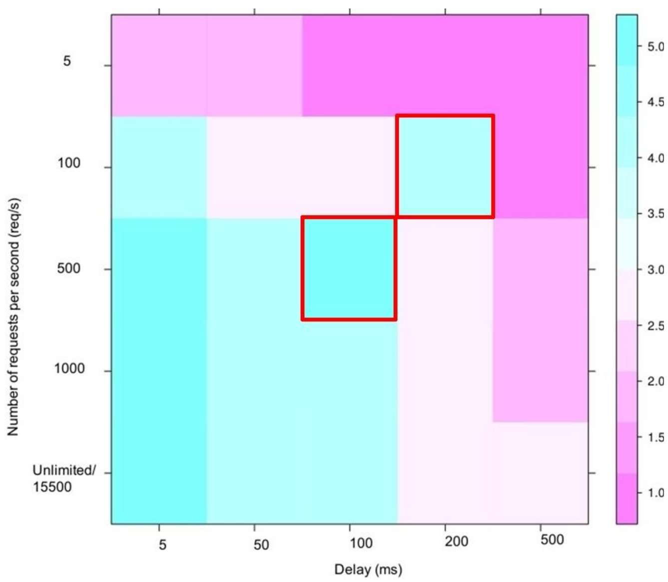


Figure 4. Mean Opinion Score vs. Delay and Requests per second

differentiate based on network parameters only. From our experiment, we shall show that the users' perception, represented by the MOS scores, can be mapped into from both the networking and application parameters.

IV. Task II: To Dertermine Key Metrics Of Network And Application Performance On Users' Scores

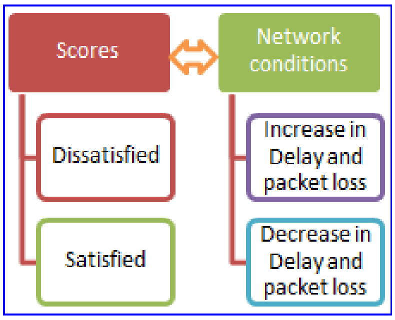
A. Aim of Delay and Requests per second testing

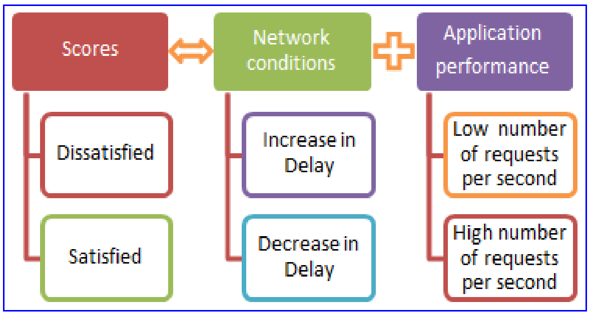
The aim of this experiment is to recognize the impact of these various metrics and their two- way interactions both from a networking and an application performance perspective in assessing QoE.

B. Results

Figure 4 shows the data collected from this experiment using the experiment design based on the Taguchi approach mentioned in section II-B.

It shows a plot of the different values of delay (D) and requests per second (RPS) with their associated mean opinion scores, which users recorded, based on the settings for D and RPS. The density of MOS using D and RPS indicates that the MOS of users depends on controllably fixed factors of D and RPS. For example, when both D and RPS are set at {500; 5}, the highest value of D, and lowest level of RPS, the density of MOS shown in the purple colour, means that it is at MOS equal to 1 or {Not Recommended}.





(a) A simple test related to Network performance (b) a complex test related to both Network and Application performance

Figure 5. Concept of "Rule of Nature"

(a) Type 1

(b) Type 2

(c) Type 3

(d) Type 4

Figure 6. Four different types of Content

In conclusion, in this more complex test, the result recognizes the impact of these various metrics and their two-way interactions both from a networking and application performance perspectives in assessing QoE as shown in Figure 4. In this environment, the rule of nature is shown as in Figure 5b in which users' scores directly reflect the controlled network and application parameter changes.

The "rule of nature" in this complex test shown by Figure 5b is an agreement between users' perception and increased values for both objective metrics of network and application performance. As shown in Figure 4, the density of MOS using delay and requests per second indicates that the MOS of users depends on uncontrollably fixed factors of delay and requests per second {D;RPS}. For example, when both delay and requests per second {D;RPS} are set at {5;Unlimited}, the lowest value of delay, and the highest level of number of requests per second, the density of MOS shown in dark blue colour, means that it is at MOS equal to 5 or {Very Satisfied}.

However, since both delay and requests per second are taken into account, there are two remaining abnormalities which are shown highlighted by red-line rectangles of Figure 4 in detail as:

- The light blue block between 100 {RPS} and 200 milliseconds in {D} which should be in an extremely light purple linked to "Fair" of users' perception instead of a light blue block linked to "Satisfied or 4".
- The dark blue block between 500 {RPS} and 100 milliseconds in {D} which should be in a light blue referred to "Satisfied or 4" of users' perception instead of a dark blue block linked to "Very satisfied or 5".

These show that there are some abnormalities in the scores of "Satisfied or 4" and "Very satisfied or 5" and that "the rule of nature" (in our terminology) is not followed if based on both networking and application performance. This suggests that we should develop further experiments to test *Content* and that may or may not have affected users' scores in the current experiment.

TABLE I. MOS VS. CONTENT

MOS	Content			
	Type 1	***Type 2***	***Type 3***	***Type 4***
1	4	10	19	21
2	25	26	29	32
3	51	48	36	26
4	18	17	15	23
5	10	7	9	6

V. TASK III: TO DERTERMINE KEY METRICS OF PSYCHOLOGICAL EFFECTS AS CONTENT ON USERS' SCORES

A. Natural Concept of Content

Content is the embedded information on a website. Alternatively, content is broadly described as the stuff in a website, this may include documents, data, applications, services, images, audio, video, or email messages [19].

In our experiments, the contents of any particular web page typically may include text, images, videos, or other multimedia content such as flash objects, online games, and Java script objects. The nature and content of the web pages dictate the elements that are presented on the web as described in Figure

6. Four different types of *Content*, type 1, type 2, type 3, and type 4, are described by Figures 6a, 6b, 6c, and 6d respectively.

B. Aim of Content testing

From the evaluation mentioned in [17] and [20], it is concluded that there is general agreement between the users' perception and increased values for both objective metrics of network and application performance and we called it "the rule of nature" as shown in Figure 5b. However to understand further changes in specific cases of users' perception and for un-matching cases of our "rule of nature" as described by the red line rectangles in Figure 4, we proposed a further metric that is related to content into our customer QoE for Web browsing model.

In this experiment, *Content* has been tested on a group of twelve users. The Content of websites consisted of four different configurations, which were categorized based on their content. The *Content* has been divided into four categories of {type1, type2, type 3, and type 4} following an increase in the complexity of the content. In the scope of this paper, the complexity of the content is objectively measured by increasing the download time of the Web page. This is because more resources and functionalities have been integrated into the higher level of content type to reduce boredom in terms of content on the web site.

According to Callan in "Content is King"[21], it is concluded that the more useful and interesting the content is, then the more successful a website will be in practice. However, based on Odlyzko in [22], "Content is not King", he argues that "connectivity is more important than content". Therefore, the aim of this experiment is to see whether Content influences users' scores under different controlled network and application performance conditions.

C. Assumption

It is assumed that if users are not affected by the content itself, then they can differentiate between an increase/decrease of download time for different types of content.

D. Analysis and Results of Content on Mean Opinion Score

Table I lists a number of different Mean Opinion Scores (MOS) from web browsing users for different types of content under the same settings of networking delay and available requests per second responded from the web server for each type of content. The Content has been divided into four categories {type1, type2, type3, and type4}.

As shown in Table I, it suggests that, the number of users' scores of 1, and 2 is increasing and that for scores of 3 is decreasing when increasing the complexity of the content. Those for scores of 4, and 5 show no trend when increasing the complexity of the content.

For example, if it is assumed that downloading of web pages is performed using an ADSL 512 K connection infrastructure, then the estimated download time is assigned for different types of content {type1, type2, type3, type4} as shown by the circle-symbol marked lines (red lines) of Figures 7a and 7b. Other lines are MOS at {1; 2; 3; 4; 5} for different types of content.

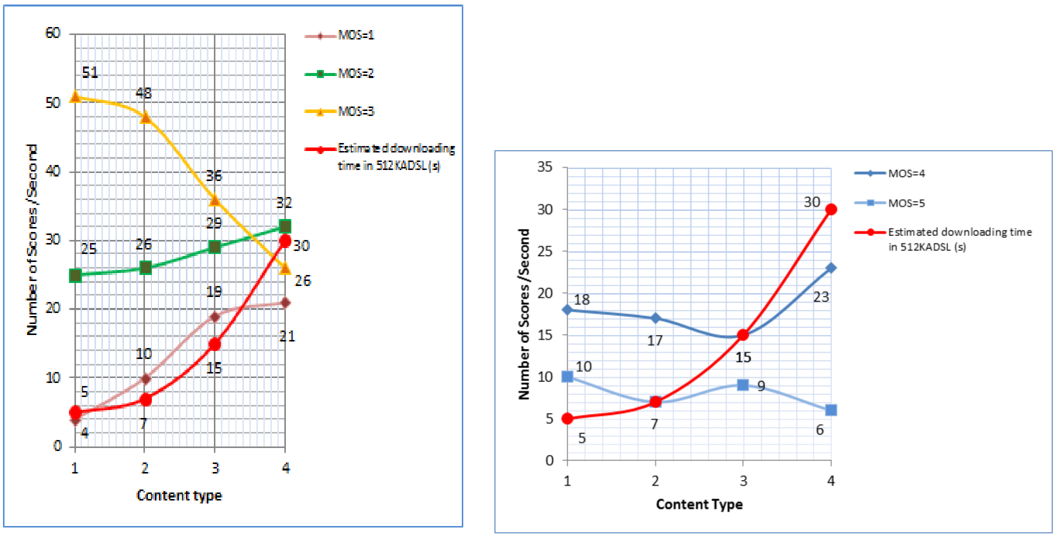
As depicted in Figure 7a, the line-trends of MOS at {1; 2; 3} fits with our assumption. When the download time is increased, with more complicated content respectively, the number of users who chose either {1; 2} illustrated by the brown and green lines is increased, while the number of "fair" users (yellow line) decreases. Thus, we see no effect of Content on users at MOS scores of {1; 2; 3}.

However, Table I suggests that the users' scores at levels 4, and 5 are not matching with our assumption as mentioned above. As depicted in Figure 7b, the estimated download time for different types of content (shown by the red line) does not match with the line trends of MOS at {4 or 5} illustrated by the blue and light blue lines. The expected trend should be the same as for the yellow line trend of MOS of 3 ("Fair") as indicated in Figure 7a. Thus, it is concluded that Content has its own effect on users' scores at levels {4; 5}.

In conclusion, it is shown that Content influences users' scores under different controlled networking and application performance. Although Content is not the most important factor to be taken into account in the QoE assessment; however it is necessary to include it in order to have a full assessment of QoE Web based services and to get a better understanding of how users experience web browsing. We found that scores at levels {4; 5} were not matching with our assumptions related to download time. It is concluded that Content influences users' score at levels {4; 5}. This also can be explained by abnormalities that occurred in our previous evaluation of the networking and application performance mentioned in Figure 4 where two blocks of users' scores at {4; 5} did not follow the setting of networking and application performance. It is concluded that this has affected users' scores in this case.

VI. TASK IV: TO EVALUATE AN INTEGRATION OF ALL KEY METRICS ON USERS' SCORES

A. Analytical modelling framework of QoE assessment for Web browsing



(a) MOS of {1,2,3} matching type of Content

(b) MOS of {4, 5} un-matching type of Content

Figure 7. MOS vs. Four different types of Content

QoE for web browsing is assessed via metrics derived from Network, Application, and *Content* as described in Figure 8a. In the framework, QoE for Web based services is accessed via their four grouped affected factors as described in Figure 8b. As mentioned previously, users can perceive a change of networking performance. Therefore, it can be established that there is a relationship between QoS and users' mean opinion scores. Thus, networking-derived metric must be the first input factors into our analytical framework. All of the networking and application based factors need to be integrated with content and time factors in order to create an appropriate framework for a QoE model involving web traffic. The content and time(CT) derived metric from our perspective is a metric related to the specific content present at the time that the user surfs a website.

Web based services are involved in viewing web content on browsers where the content and delivery of that content have been subjected to parameter modifications that are expected to influence user perception of the web services.

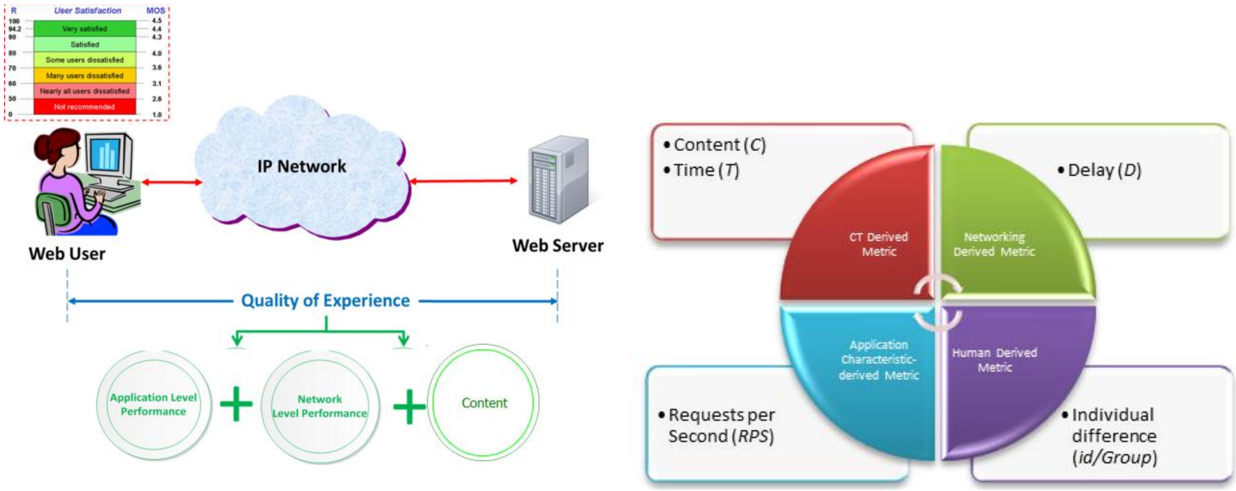
B. Assessment Results

We used a mixed effect model [23] to assess QoE of web browsing via comprehensive metrics as shown in Figure 8b. We applied the applicability of a mixed effects model in predicting QoE in World Wide Web based multi-media services [24]. The mixed effects model has been chosen for our analysis for number of important reason, viz: (1) It is robust against missing data, provided that data is missing at random. (2) Users are treated as random variables; the main effects represent human users in general, not a local group of users. (3) It allows simultaneous consideration of all factors that may potentially influence the nature of the data.

An analysis is presented on objective factors such as delay, requests per second and human factors like *Content* and the uniqueness of individuals themselves that may impact on outcomes of observations as fixed effects and random effects. The fixed effects model takes into account population mean as well as the response itself. Random effects take into account group specifics and account for the correlation structure of variations amongst users. The objective metric for QoE evaluation has been based on networking perspectives from WWW and metrics based on the application layer and human perception. The subjective metric is related to variance of each individual amongst the group. An assessment is developed to account for these factors and other potential covariates to QoE assessment during the course of our experiments. There is a random effect representing an uncontrollable variation amongst groups that is required in the model. Thus, we have prepared an assessment that takes random effects into account as a necessary feature for a potential source of random effects.

The model is developed to account for these factors and other potential covariates in QoE assessment during the course of the experiment. The significance of the model is tested using the Spearman's rank correlation [25]. The Spearman test is applied since MOS is ordinal data. Each variable is ranked separately in order to put the value of the variable in order and numbering them. The lowest gets rank 1 and the highest gets rank 5, for example, in the MOS. The test of significance of the model is shown in detail as Table II. The S value is 3229472. The Spearman value of S is calculated using Equation 1.

$$S = \sum_{i=1}^{882} d_i^2 = \sum_{i=1}^{882} (x_i - y_i)^2 \qquad (1)$$



(a) QoE Assessment Approach

(b) Analytical Framework

Figure 8. Analytical modelling Approach and Framework of QoE assessement for Web based services

TABLE II. TEST OF SIGNIFICANCE OF THE MODEL

Spearman's rank correlation rho		
Model	***S***	***p-value***
Value	3229472	< 2.2e-16 "***"
Alternative hypothesis	True rho is not equal to 0	
Sample estimate	Rho: 0.9716	

The correlation between fitted values and observed values (rho or R) is calculated following Spearman value using Equation 2, giving a value of R = 0.9716 and R^2 = 0:941.

$$R = 1 - \frac{6\sum_{i=1}^{882} d_i^2}{n(n^2-1)} = 1 - \frac{6S}{n(n^2-1)} \quad (2)$$

P value is < 2.2e-16 "***" and assessed via t-distribution by Equation 3.

$$t = \frac{R}{\sqrt{\frac{1-R^2}{(n-2)}}} \quad p = 2(1 - pnorm(t)) \quad (3)$$

C. Limitation

The effect of *Content* proposed in this paper shows a clear impact on web user QoE. However, in this work, the *Content* is limited to the concept where different kinds of content have been represented by different download times. Further context and the concepts of *Content* have not been tested in this work. Thus, an extension of *Content* concept will be considered for our future work.

VII. CONCLUSION

Four new metrics from four different categories of Networking (delay), Application (requests per second), psychology (as grouping user identity and time and limiting boredom in controlling session) and Content are proposed for the assessment of QoE for Web based services.

The approach determines the requests per second and delay metrics of web users to calculate the percentage of user satisfaction, and then the assessment is integrated with further psychological effects of Content.

The analysis and validation of integrated metrics of QoE models of Web based services based on both objective parameters from networking, application and content perspectives have shown the efficacy of the proposed model.

A comprehensive set of metrics has been evaluated on the overall web browsing customer QoE. Generally, we see that the users' scores obey the rule of nature for the case of a distorted networking environment. However, there were some abnormalities when the users' scores are at the levels of "very satisfied", and "satisfied" for the networking and application performance.

To get a more concise view of a user's expectations, a further experiment related to Content has been performed. The objective metric of *Content* has been investigated showing its effects on web browsing customer QoE. This metric has been evaluated objectively by using the download time of each type of *Content*. Specifically, in our scope of the research, we found that in some cases of (very) satisfied users, the users pay more attention and enjoy Content more, rather than the actual performance of networking and application. Alternatively, *Content* is more important than connectivity in this case. Otherwise, *Content* is less important than connectivity.

The integration of *Content* explained our existing abnormalities with users' scores of {4; 5} when the evaluation was based only on networking and application layers using delay and requests per second. Evidently, at those levels of users' scores, we found that Content has affected a users' perception. In this case, users pay more attention to content rather than connectivity.

The assessment shows an acceptable correlation between values of the fitted analytical model and observed user scores. The QoE estimates are aimed at providing service and planning, understanding user opinion, and it also targets actual customer opinion prediction.

The paper provides an understanding of psychological effect of Content in amending the standards for QoE of Web based services. Therefore, in order to take account of G.1030's shortcomings, the directions for ITU-T for QoE of web based services should be:

- Current network performance for the case of web browsing should be updated in this recommendation.
- Customer's perception should be accessed via key metrics of application and networking for web browsing.
- Human psychological effects such as *Content* should be integrated and combined together with networking and application performance for comprehensive assessment of customer QoE for web browsing, fully understanding, and further matching with customer behaviour when they are web browsing.

REFERENCES

[1] R. D. Pea, "User Centered System Design: New Perspectives on Human-Computer Interaction," Journal educational computing research, vol. 3, pp. 129–134, 1987.
[2] K. Kilkki, "Quality of Experience in Communications Ecosystem," Journal of universal computer science, vol. 14, no. 5, pp. 615–624, 2008.
[3] ITU-T, Definition of Quality of Experience (QoE) Ref.: TD 109rev2 (PLEN/12), 2007.
[4] ETSI, "Human Factors (HF); Quality of Experience (QoE) requirements for real-time communication services," 2009-2012.
[5] S. Khirman and P. Henriksen, "Relationship between Quality-of-Service and Quality-of-Experience for public Internet service," in In proceedings of the 3rd Workshop on Passive and Active Measurement, 2002.
[6] M. Fiedler, T. Hossfeld, and P. Tran-Gia, "A generic quantitative relationship between Quality of Experience and Quality of Service," Network, IEEE, vol. 24, no. 2, pp. 36–41, 2010.
[7] J. Shaikh, M. Fiedler, and D. Collange, "Quality of experience from user and network perspectives," annals of telecommunications-annales des t´el´ecommunications, vol. 65, no. 1-2, pp. 47–57, 2010.
[8] ITU-T, Rec. G1030: Estimating end to end performance in IP networks for data application, 2005.
[9] E. Ibarrola, F. Liberal, I. Taboada, and R. Ortega, "Web QoE evaluation in multi-agent networks: Validation of ITU-T G. 1030," in The fifth International Conference on Autonomic and Autonomous Systems, 2009. ICAS'09. IEEE, 2009, pp. 289–294.
[10] K. Brunnstrm. (2012) Qualinet White paper on Definitions of Quality of Experience and Related Concepts. Available from: http://www.qualinet.eu/images/stories/whitepaper v1.1 dagstuhl output corrected.pdf (last accessed 4/13).
[11] ITU-T, Rec. P800 Methods for subjective determination of transmission quality, 1996.
[12] S. Hemminger et al., "Network emulation with NetEm," in Linux Conf Au. Citeseer, 2005, pp. 18–23.
[13] OpenSource. (2012) Mod qos. Available from: http://opensource.adnovum.ch/mod qos/ (last access 4/13).
[14] R. K. Roy, Design of experiments using the Taguchi approach: 16 steps to product and process improvement. Wiley-Interscience, 2001.
[15] G. Taguchi, S. Chowdhury, and Y. Wu, Taguchi's quality engineering handbook. John Wiley & Sons Hoboken, 2005.
[16] G. Taguchi, Introduction to quality engineering: designing quality into products and processes, 1986.
[17] L. T. Nguyen, R. Harris, and J. Jusak, "Analysis of Networking and Application layer derived metrics for web Quality of Experience," in Consumer Communications and Networking Conference (CCNC). IEEE, 2012, pp. 321–325.
[18] L.-T. Nguyen, R. Harris, and J. Jusak, "A pilot study to assess Quality of Experience based on varying network parameters and user behaviour," International Proceedings of Computer Science and Information Technology, vol. 5, pp. 30–34, 2010.
[19] P. Morville and L. Rosenfeld, Information architecture for the World Wide Web: Designing large-scale Web sites. O'Reilly Media, 2008.
[20] L.-T. Nguyen, R. Harris, J. Jusak, and A. Punchihewa, "Modelling of Quality of Experience for Web Traffic," in 2010 Second International Conference on Network Applications Protocols and Services (NETAPPS). IEEE, 2010, pp. 84–89.
[21] D. Callan. (2006) Content is king. Availabe from: http://www.akamarketing.com/content-is-king.html (last accessed 4/13).
[22] A. Odlyzko, "Content is not king," First Monday, vol. 6, no. 2, 2001.
[23] R. H. Baayen, D. J. Davidson, and D. M. Bates, "Mixed-effects modeling with crossed random effects for subjects and items," Journal of memory and language, vol. 59, no. 4, pp. 390–412, 2008.
[24] L. T. Nguyen, R. Harris, A. Punchihewa, and J. Jusak, "Application of a Mixed Effects Model in Predicting Quality of Experience in World Wide Web services," in Computational Intelligence, 2012 Fourth International Conference on Modelling and Simulation (CIMSiM). IEEE, 2012, pp. 316–321.
[25] C. Spearman, "The proof and measurement of association between two things," The American journal of psychology, vol. 100, no. 3/4, pp. 441–471, 1987.