

A Pilot Study to Assess Quality of Experience Based on Varying Network Parameters and User Behaviour

Le Thu Nguyen ¹, Richard Harris ², and Jusak Jusak ³

School of Engineering and Advanced Technology
Massey University, Private Bag 11 222, Palmerston North 4442, New Zealand

¹ Email: L.T.Nguyen@massey.ac.nz ² Email: R.Harris@massey.ac.nz ³ Email: J.Jusak@massey.ac.nz

Abstract. In this paper, we describe a pilot study to assess Quality of Experience (QoE) for users who are web surfing in a controlled environment. The assessment is based on parameters that are derived from a network performance metric and user actions. This study forms part of a contribution to the development of tools for determining users' QoE, based on our proposed analytical model. We set out to investigate whether the proposed model is able to predict user QoE as we modify network parameters and observe user behaviour.

Keywords: Quality of Experience (QoE), user's perception, Quality of Service.

1. Introduction

The User Experience terminology was created by Dr. Donald Norman, who was the first person to describe the importance of a user-centred service design [1]. Recently, this term has become popular for researchers and practitioners in order to identify the means for providing better services to users. A user's experience of a telecommunication service has been defined by the ITU-T as the users' perception of the acceptability of an application or service [2].

A quick survey of the current literature reveals that the concept of QoE is receiving a great deal of attention at the present time, with many local and international conferences being held and almost an entire edition of the IEEE Networks Magazine (March/April 2010 Vol 24. No .2) being devoted to this topic [3]. It can be said that the study of QoE is a rapidly evolving issue for a number of reasons, including: Objective QoE assessment is needed in order to reduce time and cost since it was previously performed using time consuming subjective assessments in order to determine Mean Opinion Scores (MOS). However such assessments are prone to various errors due to their subjective nature; hence it is important to develop reliable but objective methods for determining an accurate measure of a user's experience of a telecommunications service. It is obvious that QoE should attract a great deal of attention in the Telecommunications Industry where a user's experience is now considered to be even more important than the traditional Quality of Service (QoS) perspective.

Modelling of QoE has been studied and applied to various kinds of services such as voice, video and multimedia services. The latter have involved services such as video conferencing, storage, and retrieval applications. In these services, users are acting as listeners, viewers, and audiences respectively.

QoE assessment of web browsing has been studied and published by many authors including [4], [5], [6], [7], and [8]. The first recommendation for QoE assessment of web applications was published in the ITU-T standard known as G.1030 [4] in 2005. This study mapped the MOS to objective network measurements. The current studies [5], [6], [7] have contributed enormously to user-centred designs where a QoS-QoE correlation model for objective QoE was proposed. Even in the case of MOS measurements, current research

is attempting to enhance its reliability through the use of a probability distribution function instead of the traditional discrete valued MOS scores [8].

In this paper, a QoE assessment is applied to the special case of web browsing in which users in a pilot experiment are web surfers. This paper presents a novel view on QoE assessment which is based on a combination of network performance and user’s actions. Hence, these results are targeted as helping us with calibrating our model [9] for future work.

2. Methodology

Our pilot study was carried out using the networking infrastructure using a test-bed located at Massey University in Palmerston North. The network topology of the test-bed consists of seven routers and seven switches as shown in Figure 1. 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.

The network performance has been changed in different controlled environments by using Netem [10] which is readily available for the Linux Kernel [11]. Our experiment runs over a period of approximately 25 minutes during which time the networking conditions are changed at approximately five minute intervals. The delay and packet loss of the network can be changed every five minutes to different values as summarised in Table I under different networking conditions as shown in Figure 2.

There were 12 users involved in our pilot study. All of them were post-graduate students studying at our university. They accessed the same website sources simultaneously. In this experiment, the impairments that were applied were the same for all participants and experienced simultaneously.

While users accessed the website, the users’ actions were captured to see the behaviour that they exhibited through their selection of hyperlinks on the displayed page, together with the frequency and nature of these selection activities. This information is captured by using the well-known packet capture software known as Wireshark [12]. Only http information is filtered from the captured data. Users’ actions are captured on each user’s computer.

The contents of the particular web pages include text, images, videos or other multimedia content such as flash objects, online games, and Java script objects.

Users were asked to give an evaluation of their own experiences at five minute intervals through the use of a scoring system for the particular web pages that they were surfing under different networking situations. The scores were based on MOS values [13] as described in Table II below; via a feedback page accessed via the pre-prepared web pages. The feedback page was designed as a radio button selection; therefore, it is expected that this task would not involve much time for the users in sending their user feedback experience. The feedback from users of web surfing during their use of the service was gathered automatically and stored in a standard MySQL database [14].

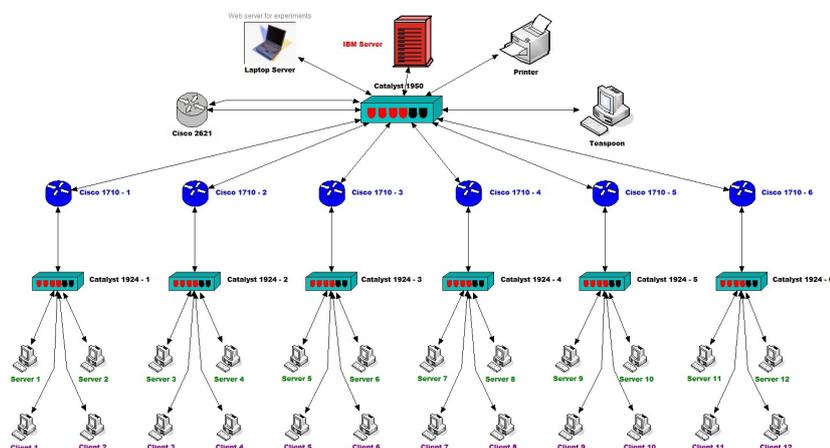


Fig. 1: Network Topology for the Test Bed

Figure 3 shows the percentage of user scores for different delay values. It can be observed that when the network is operating appropriately (delay is less than 50 milliseconds), we see that 75% of users feel “Satisfied” and “Fair” in their evaluation of the system. Meanwhile, the percentage of users who vote for “Dissatisfied” and “Not Recommended” increases when the network performance becomes worse. It can be seen in the poorest networking situation depicted in Figure 3, there is a 35% higher percentage of users who voted for “Dissatisfied” and “Not Recommended” than compared to the first case. In the nearly perfect networking conditions (where the delay is less than or equal 50 milliseconds), it is shown that less than 10% of users are “Very Satisfied”. This means that the users’ expectations about the system response are higher than that of the current condition.

Figure 4 shows users’ scores in both increasing values of delay and packet loss simultaneously. 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 conditions. 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 reduced proportions for the “Very Satisfied”, “Satisfied”, and “Fair”. From this figure, it can be concluded that there is an agreement between users’ perception and increased values for both delay and packet loss.

In general, from Figure 3 and Figure 4, we can see that the users’ score obeys the “rule of nature”. For example, the users’ scores tend to be “Dissatisfied” when the delay and packet loss of the network are increased, on the other hand the users’ scores move towards “Very Satisfied” and “Satisfied” when the delay and packet loss are decreased. In conclusion, it is shown that the general trend of the experiment for the different network conditions explained above is as expected.

However, there are some abnormalities that are noticed in Figure 3. The users’ scores do not show any specific trend in the case of “Fair” level. Furthermore, we also notice abnormal behaviour when the network condition is set to become worse where the percentage of users that vote for “Not Recommended” is decreasing. These two cases show that the users’ perception seems to be the opposite of our previous argument. The reasons for this might be because (1) Users’ perception is subjective in nature. (2) There is a limitation in the number of users involved in the experiment.

To get a more concise view of the users’ perception, we suggest looking at users’ actions which might be able to capture a users’ perception via their activities. This can be done by counting the number of requests per user for different networking conditions. This is motivated by the fact that the users’ perception assessment based on network performance only is not adequate for comprehensive analysis of the QoE.

Figure 5 shows the number of requests against different kinds of delay values. It can be seen that the number of requests is decreasing when we increase the delay across the network.

For example, when we set the delay to less than or equal to 50 milliseconds, the number of requests for web services is more than 61%, on the other hand it is just around 56% when the delay is increased to 500 milliseconds.

It is quite interesting when we observe the user’s actions when deeply analysing the http protocol data. We notice that when the network delay becomes large, the number of requests for web pages from the server decreases. The actual reasons for this notion are hard to explain. It could be because users have an ability to react in different ways to response in the change of delay in the network. The two possibilities are (1) the users have a high level of patience or, the opposite, (2) the users may feel bored with the experiment and do not really want to continue.

All in all, the experiment shows that it is possible to measure QoE for web surfing in combination with MOS scores, network parameters, and user’s actions.

4. Conclusions

This study reflects the relationship among MOS scores, users’ actions and network performance for the purpose of QoE assessment as required in our model [9].

In this paper, users' actions were modelled by using only one parameter that is the number of requests sent from each user in a specific time period only. While the network performance was modelled using two parameters; viz: the packet delay and packet loss.

We can see from this study that users' perceptions are affected by delay and packet loss. However, to gain a correct assessment of the QoE, the users' actions should be included in the measurement.

In this experiment, we realized that we only have a few data points because of limitations in available time and personnel for the experiment. Ideally, we would wish to have more data points for both packet delay and packet loss. In addition, we only employed a limited number of subjects in the experiment, i.e. 12 users. That is the reason why we have described this work as a pilot study. However, we seek and expect a larger number of users in future experiments.

5. Future work

More experiments are needed to be performed and more parameters will need to be analysed for establishing the relationship among users' scores, users' actions, and network parameters. More users will be surveyed in future experiments.

In more detail, we shall analyse information about the probability of abort, next, and retry when users browse a web page. Packet loss is going to be further analysed in the next step.

In the near future, the model will also be calibrated by using the two main factors which comprise some of the networking parameters and some parameters involving user actions.

6. References

- [1] D. Norman and S. Draper, *User centered system design: New perspectives on human-computer interaction*. L. Erlbaum Associates, 1986
- [2] ITU-T, "Definition of quality of experience (qoe)," *International Telecommunication Union Liaison Statement*, vol. Ref: TD 109 rev2 (PLEN/12), Jan 2007.
- [3] "IEEE network - the magazine of global internetworking - front cover," *Network, IEEE*, vol. 24, no. 2, pp. 4–48, 2010.
- [4] I. Rec, "G. 1030-Estimating end-to-end performance in IP networks for data applications," *International Telecommunication Union, Geneva, Switzerland*, 2005.
- [5] T. Hoßfeld, P. Tran-Gia, and M. Fiedler, "Quantification of quality of experience for edge-based applications," in *Proceedings of the 20th international teletraffic conference on Managing traffic performance in converged networks*. Springer-Verlag, 2007, pp. 361–373.
- [6] J. Shaikh, M. Fiedler, and D. Collange, "Quality of Experience from user and network perspectives," *Annals of Telecommunications*, vol. 65, no. 1, pp. 47–57, 2010.
- [7] S. Khirman and P. Henriksen, "Relationship Between Quality-of-Service and Quality-of-Experience for Public Internet Service," in *Proc. of the 3rd Workshop on Passive and Active Measurement*.
- [8] L. Janowski and Z. Papir, "Modeling subjective tests of quality of experience with a Generalized Linear Model," in *International Workshop on Quality of Multimedia Experience*, 2009. QoMEX 2009. IEEE, 2009, pp. 35–40.
- [9] L. 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*. IEEE, 2010, pp. 84–89.
- [10] S. Hemminger *et al.*, "Network emulation with NetEm," in *Linux Conf Au*. Citeseer, 2005.
- [11] D. Bovet and M. Cesati, *Understanding the Linux Kernel, Second Edition*, 2nd ed., A. Oram, Ed. Sebastopol, CA, USA: O'Reilly & Associates, Inc., 2002.
- [12] G. Combs *et al.*, "Wireshark," *Web page: http://www.wireshark.org/last_modified*, pp. 12–02, 2007.
- [13] Rec, "P. 800.1, Mean Opinion Score (MOS) Terminology," *International Telecommunication Union, Geneva*, 2006.
- [14] MySQL, "MySQL Administrators Guide," Extracted from MySQL Administrators Guide 2005 Sams Publishing, <http://www.volalaw.com/uploads/OptimizeMySQLwiththesepowerfultips.pdf> last accessed 10th February 2011.