User quality of service perception in 3G mobile networks

Citation for published version (APA):

Papamiltiadis, K., Zisimopoulos, H., Gasparroni, M., & Liotta, A. (2004). User quality of service perception in 3G mobile networks. In *Proceedings of IEE Telecommunications Quality of Services : the Business of Success (QoS 2004), March 2-3, 2004, London, UK* (pp. 64-69). Institute of Electrical and Electronics Engineers. <https://doi.org/10.1049/ic:20040018>

DOI:

[10.1049/ic:20040018](https://doi.org/10.1049/ic:20040018)

Document status and date:

Published: 01/01/2004

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

USER QUALITY OF SERVICE PERCEPTION IN 3G MOBILE NETWORKS

K. Papamiltiadis*, H. Zisimopoulos, M. Gasparroni, A. Liotta

*Vodafone House, Newbury, Berkshire, RG14 2FN, UK,
Email: konstantinos.papamiltiadis@vodafone.com, Facsimile: +44 (0) 1635 631127

Keywords: QoS perception, GPRS, UMTS, mobile web browsing, instant messaging.

Abstract

UMTS reflects the latest technical and standardisation efforts to realise and deploy a QoS-enabled mobile network infrastructure that can support QoS-dependent services and applications. The next challenge is therefore how to efficiently use and configure the network in a way which meets user expectation while conserving network resources. In order to achieve that, it is important to have qualitative information on how different QoS levels are perceived by the user. The network operator/service provider should be aware of the point at which the user perceives the effect of QoS fluctuation, performing a reasoned and accurate dimensioning of the network. Many subjective studies have already been undertaken for conversational and streaming services, whereas in this paper we focus on the analysis of user QoS perception in the context of mobile, interactive services, specifically Internet access, web browsing, and messaging.

1 Introduction

The project objective is to determine the level of quality of service perceived by end users when they access mobile services, focussing in particular on mobile Internet access including Web browsing [5] and Instant Messaging [3]. These two applications, classified as "Interactive" in 3G networks, were chosen because many subjective studies have already been undertaken for conversational and streaming services. Mobile users expect to access such services through UMTS or WLAN hotspots with similar levels of performance enjoyed through wireline networks. While this is theoretically possible thanks to the increased capabilities of mobile and wireless networks, it is also important to have a precise understanding of how many resources should be allocated in face of a given QoS perception target.

The project has also concentrated in developing an effective methodology for choosing the series of tests aimed at assessing how and to what extent different network bearer QoS parameters are perceived by the mobile user. The tests include a sequence of identical service delivery patterns in which different end-to-end QoS parameters (i.e. bandwidth, delay, jitter, and packet loss) are modified without the user being aware of it.

In our study, a sample of 72 mobile users has been subjected to individual 30 minute tests in order to compile a QoS perception matrix that has been, finally, used for the statistical analysis of the results.

The test methodology defined during the project can also be applied to determine user QoS perception for a number of more advanced mobile services that have been envisioned for several years now; examples are adaptable services, dynamically composable services and context-aware services (including QoS- and location-aware services), which require QoS support both towards the network (e.g. QoS provisioning and configuration) and towards the application (e.g. adaptation to changing network QoS conditions).

2 Related Work

In order to calibrate effectively the test matrix, existing projects on this subject were analysed. Specifically, Bouch et al [2] conducted a study aiming to analyse the efficiency of web-site design by identifying how long users would be prepared to wait for Web pages to load. Users were presented with Web pages that had predetermined delays ranging from 2 to 73 seconds. While performing the task, users rated the latency (delay) for each page they accessed as high, average or poor. Latency was defined as the delay between the request for a Web page and its complete download. They reported the following ratings:

Rating	Download Time
High (good)	Less than 5 sec
Average	From 6 to 10 sec
Low (poor)	Over 10 sec

Table 1: Rating of download time [2].

Despite the clear classification made in Table 1, the overall conclusion drawn by the study was that "acceptable" response times are difficult to determine because people expectations depend on many factors such as the perceived complexity of the request. For example, "people will wait longer for requests they think are hard or time-consuming for the computer to perform" [2].

The study conducted in Karlstad University, Sweden by Asplund and Brunstrom [1] explored how users trade-off between image quality and latency when downloading Web pages. A reference Web page was used to provide qualitative

judgements of the received service for ten different cases concerning two web sites. With the latency varying from 1 sec to 40 sec and image quality from perfect to very low, the users were asked to answer three questions regarding image quality and service speed. The answers were measured using a continuous scale with two anchor points and nine tick marks and were converted to integers between 0 and 100.

While previous studies only focused on total download time in the context of fixed networks, our research specifically concentrated on the influence of network QoS parameters such as bandwidth, delay, jitter and packet loss on users' perception, in the context of UMTS.

3 Methodology

This section describes the methodology followed in setting up the test bed used in the study, performing the tests and extracting the results, for the two applications considered.

Figure 1: Test bed abstraction from the UMTS QoS architecture

Figure 1 shows the UMTS QoS infrastructure as devised by 3GPP. In order to reproduce this end-to-end bearer service, the commercial network emulator Shunra Cloud™ [12] was used. The basic parameters characterising network behaviour included in Cloud are the following:

Delay – the time it takes for a packet to travel from one endpoint to another.

Delay Variation – The difference in the delay consecutive packets experience when passing through the network. This is configured in Cloud by specifying the distribution of delay.

Bandwidth Limitation – Bandwidth available for the session.

Packet Loss / Errors – The percentage of packet loss can be configured in Cloud.

3.1 Web-browsing

In the first part of work, user perception in downloading web pages under different QoS conditions was investigated. A local web site was set up and consisted of both written text and pictures with a minimum total size of 200 Kbytes, allowing for sufficiently long downloading times, even under the best network conditions. This “frozen” web site had three

other links, which were again locally stored. The procedure of “jumping” from one web page to another captured better the effect that the QoS parameters had on the overall user perception [2].

Tests have been undertaken with an instance of three “frozen” web sites, namely bbc.co.uk, uefa.com and msn.com.

Four different sets of experiments have been identified. Each set focussed on one of the four QoS parameters (bandwidth, delay, delay variation, packet loss) under analysis, while maintaining the other three at a constant value with minimum influence on user's experience.

Seventy-two people from Vodafone, UK and from the University of Surrey, UK participated to the subjective tests. In this way the sample was composed of users having different professions, income, educational level etc.

Input concerning the different conditions of the tests and the formulation of the Test Matrix were obtained from previous Vodafone projects on related issues.

The questionnaires were in the form of multiple-choice, with the first part containing personal questions aimed at classifying the test users according to age and Internet frequency and purpose of use.

The second part of the questionnaire was related to the questions of the tests undertaken by the users. This part consisted of seven or eight questions, depending on the type of test, marked from -5 (slow), to +5 (fast), when compared to a given reference condition (marked as 0).

The duration of the test was set strictly to 30 minutes for each sample user. Each user performed a series of tests using different web sites, for example MSN website for bandwidth-related tests, BBC website for delay-related tests, etc. For each of these web sites the user was previewed with the best and worst conditions available in the experiment, in order to understand the scale used and “calibrate” his or her perception. After the preview session, the test started with the user downloading the specified website with the “reference condition” before testing the specific “experimental condition”. The “reference condition” was shown at each test, making the test sequence like “reference”, “conditions a”, “reference”, “conditions b”, etc.

3.2 Instant-Messaging

Several Instant Messaging (IM) implementations are currently available on the market, but none of them can be considered to be the “common practice”. The mobile industry is investigating all the potential solutions in terms of functionality with the aim to ease their integration into the UMTS architecture [8]. The IM architecture considered in this project was the SIP-based Instant Messaging ([3,4]). Furthermore, to the best of our knowledge, the QoS requirements of IM have not been studied so far.

The Shunra Cloud™ network emulator was used in the test bed to analyse the effects of the abovementioned QoS

parameters (delay, delay variation, bandwidth, and packet loss) on user QoS perception. Microsoft Portrait, which is a prototype application for mobile video communication ([11]), has been used in the tests.

During the experiments, the users exchanged two pre-defined dialogues under different QoS conditions. The first pre-defined dialogue included only plain text (the lightweight test); whereas the second one included the exchange of an image between (the heavyweight test).

The starting point of our experimental procedure was the test matrix used for web browsing. After conducting an internal phase of preliminary tests, three different experiments were identified. In particular, due to the small size and the bursty nature of the data exchanged, it has been observed that the granularity of the QoS parameters was not as fine as the one found in the case of the web browsing experiments. As a consequence, we decided to adopt only three network configurations, representing average conditions for UMTS RACH & FACH shared bearer, GPRS bearer with 4TS and UMTS DCH bearer at 64kbps. The aim was to derive a possible mapping strategy for this application to different mobile bearers.

The reference condition was set to GPRS-like network, as it represents an intermediate conditions. The same data collection method as the one specified for Web browsing was used. The user population selected was from Vodafone R&D, where twelve people in total undertook the tests.

The same scales and measurement procedure as in Web-Browsing were used. The duration of the test was again 30 minutes. For each dialogue, the "reference condition" was presented only once at the beginning of the test sequence, followed by the three network conditions (UMTS RACH & FACH, GPRS bearer with 4TS DL and UMTS DCH bearer at 64kbps).

4 Results

4.1 Web-browsing

The first result represents a sort of sanity check and aims at establishing whether or not the user could recognise the absolute downloading time in comparison with the reference condition.

Figure 2: Percentage of Correct Recognition

The situation is depicted in Figure 2; a total of 71.7% of the sample population has correctly identified the difference between each test and the reference. As expected, the conditions close to the reference were the ones associated with a larger ambiguity in user perception. This was due to the fact that the difference in the absolute downloading time was less than 10 sec, a gradient that Bouch *et al* have already found to be difficult to be identified by the average user [2].

The results were also analysed from another point of view to evaluate whether the users could detect the relative increase/decrease in the page downloading time (delay granularity). 10 possible answers were considered, ranging from -5 to -1 marks for quality decreases and from 1 to 5 marks for quality increases. In order to proceed with this analysis, a parameter called "Expected Rating" has to be defined as follows:

$$\text{Expected Rating} = \frac{\text{Abs}(5 \times (\text{Download time of the test} - \text{Download time of the reference}))}{\text{Download time of fastest/slowest test} - \text{Download time of the reference}} \quad (1)$$

The expected rating, as given in equation 1, is calculated by the combination of two elements, the real download time and the scale used for rating. The expected rating was then calculated for each of the web sites individually. Figure 3 depicts the number of users that could correctly detect the exact relative increase/decrease according to the expected rating formula and, at the same time, the number of users that rated each test as better or worse than the expected rating.

Figure 3: Percentage of Accuracy in rating

The next category of results concerns each of the aforementioned network QoS parameters. The first one was bandwidth, where the results achieved are shown in Figure 4. The graph depicts the average rating of all answers given by the sample population in each specific test. It has to be noted that the test sequences between 20 and 80 kbps and between 100 and 300 kbps were always undertaken by different people, which may have created a "memory effect" to the sample users although every time the reference was presented. This might explain why there is a peak in the area of 80 kbps.

The results of Figure 4 allow drawing some important conclusions. The increase in bandwidth (which implies a

decrease in the total downloading time) is correctly identified by users for scenarios below and around the reference level, but this was not registered for cases above the reference level. The increase in bandwidth beyond 100Kbps did not have an effect on the total download time and the users were able to detect that. This was probably due to the TCP "slow start" mechanism and to the relatively small size of the web pages.

Figure 4: Average and Expected Rating versus bandwidth

The next parameter investigated was round-trip delay. A strong influence of this parameter on the end user perception has been observed; the average rating decreased almost linearly as the delay was increased (Figure 5).

Figure 5: Average and Expected Rating versus delay

The results obtained for delay variation are shown in Figure 6 and Figure 7. Two constant delay values were investigated with the delay variation to be considered as a percentage of the standard delay deviation to the mean value.

The results prove that when the overall delay is low, the effect of delay variation starts becoming detectable only when jitter is higher than 50%. It is important to note that with an average delay of 100ms, the values of delay variation above 50% seriously degrade the user perception although the objective measurement of expected rating (which is based on total download time) is not degraded. This is due to the multi-component nature of web-browsing applications and proves

that the absolute download time is not the absolute measure of the user perception.

The last parameter investigated was packet loss ratio. In mobile networks, considering the facts that for Interactive and Background services reliability is assured by both backward (i.e. ARQ protocols below IP such RLC) and forward (i.e. coding) error correction mechanisms ([3,4,6], the erroneous nature of the air-interface is normally traded with increased delay variation [7].

Figure 6: Average and Expected Rating versus delay jitter with 100ms delay

Figure 7: Average and Expected Rating versus delay jitter with 1000ms delay

Consequently, only packet loss values lower than 5% were investigated. The effect of packet loss was examined for two distinctive values of offered bandwidth, namely 40 and 100 kbps, in order to demonstrate possible correlations between these two parameters in user QoS perception. The results prove that the effect of packet losses becomes more noticeable for higher bandwidth situations. In cases of narrow bandwidth, packet loss does not significantly worsen user QoS perception. This can be explained by the fact that the limited bandwidth already degrades user perception, hence, the "extra" degradation produced by packet loss is not

noticed. Figure 8 shows the results in the case of 100Kbps bandwidth.

Figure 8: Average and Expected Rating versus packet loss ratio

4.2 Instant-Messaging

As mentioned in the methodology section, the study carried out for Instant Messaging was more limited. In this case, only three “network configuration” scenarios were investigated.

Figure 9 depicts the percentage of users that could correctly identify whether the condition under test was better or worse than the reference. This graph proves the significance of the influence that the nature of the dialogue scenarios can have on users perception and on their ability to distinguish between different conditions. For the first, lightweight dialogue (dialogue 1), the users could correctly identify that the UMTS RACH / FACH case was worse than and that UMTS DCH case was better than GPRS (reference case) in percentages around 45% and 63% respectively. On the other hand, for the heavyweight dialogue scenario the users could correctly identify that difference in a percentage slightly higher than 90% for both cases.

Figure 9: Percentage of Correct Recognitions (in a two points scale, better/worse)

As it can be seen from Figure 10, the users could hardly detect any difference when just chatting to each other (Dialogue 1). On the other hand, the exchange of files in the second dialogue made the application more “resource demanding” and therefore made the different QoS levels more evident to the users.

Figure 10: Average rating for the two dialogues (in a -5 to +5 scale)

5. Conclusions and Recommendations

5.1 Web-browsing

It was concluded that the user perception of bandwidth and delay are related. The increase in the offered bandwidth over 60-80Kbps, when the underline end-to-end delay is 500 ms (high delay), cannot be detected by the users. 60Kbps (UMTS like conditions or GPRS with more than 4 timeslots allocated) create better user perception than 40Kbps (GPRS with 4 time slots) under the same delay conditions for web browsing. Consequently, it is ineffective to offer high bandwidth bearers in UMTS (higher than 64kbps) if the delay cannot be decreased.

It is necessary to maintain the delay variation below 50%. Above this value it starts affecting user QoS perception even under relatively low delay conditions (100ms). It is not valuable to spend effort trying to minimise jitter when the total delay is too high, because this will not have any effect on the user's perception. Delay variation creates worse user perception than any other QoS parameter in web browsing.

In lossy environments, the offering of more bandwidth does not improve user's perception. It is also not convenient to “trade” bandwidth with increasing packet loss (e.g. by not applying ARQ below IP or by employing less robust coding schemes).

5.2 Instant-Messaging

Due to the size of the sample population, the results obtained from this series of tests cannot be generalised. Additionally, the fact that we only considered three testing conditions

(emulating three different wireless bearers), prevented us from extracting individual conclusions for each of the QoS parameters, as we did for Web browsing.

However, some consideration can be drawn for the initial study. Under our testing conditions, the nature of the exchanged messages considerably influenced user perception. When the context of the dialogue is just text, regardless of its size, including “emoticons”, there is no need to map the Instant Messaging application to a dedicated bearer. A narrow bandwidth shared bearer such as the one offered by RACH/FACH is sufficient. When the dialogue leads to a file transfer or to anything heavier than just text messages, the user starts realising differences in the performance between different bearers.

Acknowledgements

The authors would like to thank Vodafone Group R&D UK for funding this project. We wish to thank all the people who have kindly accepted to take part in the subjective studies, including the Vodafone employees and the students of the University of Surrey.

References

- [1] K. Asplund, et al. “An Evaluation of User Web Service Requirements”, Karlstad University, Karlstad, Sweden.
- [2] A. Bouch, et al. “Quality is in the eye of the beholder: Meeting users’ requirements for the Internet quality of service”, CHI 2000, (2000).
- [3] G. Camarillo “SIP demystified”, McGraw Hill, (2002).
- [4] B. Campbell, et al. “Instant Message Sessions in SIMPLE”, IETF draft, draft-ietf-simple-message-sessions-01, (2003).
- [5] R. Fielding, et al. “Hypertext Transfer Protocol – HTTP 1.1”, RFC 2616, (1999).
- [6] R. Lloyd-Evans, “QoS in integrated 3G networks”, Mobile Communications Series, Artech House Inc, (2002).
- [7] R. Ludwig “A Case for Flow-Adaptive Wireless Links”, Ericsson Research, (1999).
- [8] K. Markopoulos, et al. “Mobile Instant Messaging and presence”, Vodafone Group R&D internal document, (2002).
- [9] 3GPP TS 23.107 v5.8.0 “QoS Concept and architecture”, Mar.2003
- [10] 3GPP TS 23.207 v5.6.0 “End-to-end Quality of Service (QoS) concept and architecture”, Jan.2002
- [11] <http://research.microsoft.com/~jiangli/portrait/>
- [12] http://www.shunra.com/products/cloud/cloud_1.php