Proc. Thirteenth International Workshop on Quality of Service, Passau, Germany. 298-312.

Case Study in Assessing Subjective QoS of a Mobile Multimedia Web Service in a Real Multi-access Network

Tiia Sutinen[1] and Timo Ojala[2]

[1]VTT Electronics,
Kaitoväylä 1, 90571 Oulu, Finland
Tiia.Sutinen@vtt.fi
[2]MediaTeam Oulu, University of Oulu,
P.O.Box 4500, FI-90014 University of Oulu, Finland
Timo.Ojala@ee.oulu.fi

Abstract. This paper presents an empirical task-based user evaluation, which was carried out for the purpose of assessing the subjective QoS of a mobile multimedia web service in a real multi-access network environment comprising of WLAN and GPRS networks and automatic mobility management with Mobile IP. Subjective quality ratings were collected from 20 test users to obtain a distribution representing the service quality experienced by the subjects. The obtained results show that even though the service usability in GPRS domain was barely satisfactory, this type of automatic utilization of multiple access networks with some enhancements can be considered as a potential access method for mobile services.

1 Introduction

Increasing availability of multiple access technologies and multi-mode user terminals makes it possible for mobile Internet users to utilize the advantages of the different technologies by dynamically routing traffic always via the best available access network. This is referred to as multi-access. Depending on the characteristics of the available access networks, however, vertical handovers can introduce substantial changes in the QoS. Even if the operation of user applications and ongoing communication sessions could be preserved during the handover with different mobility and handover management mechanisms, it still might cause the subjective application quality to degrade to an unacceptable level in the user's opinion.

This work has a two-fold motivation. First, the aforementioned technology push has created environments with multiple access networks (e.g. GPRS, EDGE, UMTS and WLAN), and mobile terminals with multiple radio interfaces (e.g. Nokia 9500, Motorola CN620, and Qtek 9090). Thus, in terms of technology the road is paved for introducing new service scenarios relying on utilization of multi-access networks.

The second motivation is the correlation between market pull and user experience. A positive/negative user experience of a service or an application has high impact on its commercial success. Therefore, it is imperative to evaluate the usability of a new service to expose any factors that might hamper the user experience. Since most mobile services are interactive systems, their true evaluation can be carried out only empirically [10].

The novel contribution of this work is an empirical task-based user evaluation of the end user QoS of a mobile multimedia service in the real environment of use involving a real multi-access network comprising of WLAN and GPRS networks and automatic mobility management by Mobile IP.

Related work is scarce, as published studies on QoS in multi-access networks (e.g. [22]) do not include a concrete empirical evaluation of end user QoS in a real-world realization of the proposed QoS architecture, but report performance characterizations by simulations if any. The Moby Dick project [8] developed an IP-based QoS architecture for multi-access networks, but the results of the six months real environment field trial advertised on the project's web site are not yet available in form of scientific publications. The VHO project [19] has developed prototype services based on vertical handover in a multi-access network, but the available publications do not include any user evaluations of the services.

This paper is organized as follows. In section 2 we describe the ABC (Always Best Connected) concept and its realization with a heterogeneous network comprising of multiple access networks, together with relevant mechanisms for mobility and handover management. In section 3 we discuss various methods for assessing subjective application layer quality of service. Section 4 reports our case study on assessing end user QoS in a heterogeneous network, and section 5 concludes the paper.

2 Inter-technology Mobility

2.1 ABC – Always Best Connected

The different access technologies available today may differ even significantly from each other in terms of bandwidth, coverage, cost of the connectivity, etc. Also the set of networks accessible to a mobile user varies from area to area as depicted in Figure 1. Further, it is acknowledged that no single technology, either today or most likely in the future, can provide optimally all the qualities demanded in a network connection [9], including for instance tetherless connectivity, ubiquitous access, and above all sufficient bandwidth for future applications. Thus, multi-access, which enables the mobile user to utilize the different access networks in a dynamic fashion by routing his/her connections through the best available network in any location, is a widely accepted scenario for the near future [9], [18].

A new concept of staying always best connected (ABC) has thus emerged. The goodness of a network connection can be defined based on various criteria, including QoS parameters, personal preferences, device capabilities, application requirements, operator or corporate policies, and available network resources. The basic idea is to optimally combine the advantages of the different technologies (e.g. the wide coverage of cellular networks and the large bandwidth of occasional WLAN hot spots) and the chosen set of criteria can be used to build algorithms that automatically maintain an optimal connection to the used services. The required functionalities

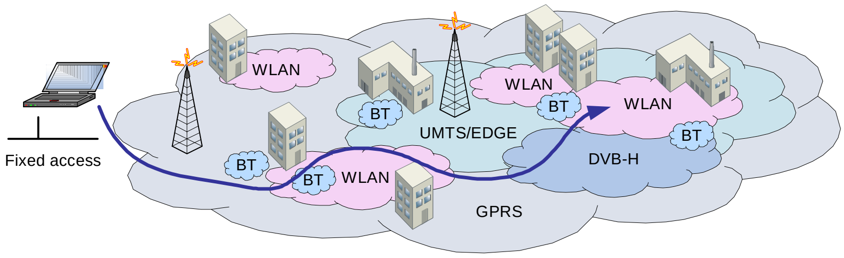


Fig. 1. Multiple access networks in different places

recognized for an ABC service include access discovery, access selection, authentication, authorization and accounting (AAA) support, mobility management, profile handling, and content adaptation [4].

Mobility in a heterogeneous network environment is referred to as inter-technology roaming, where the roaming user's network connection can be transferred either horizontally between points of access belonging to the same technology (horizontal handover) or vertically between points of access belonging to different network technologies (vertical handover) [12]. The realization of inter-technology roaming comprises of mobility management and vertical handover support, which we discuss next from the viewpoints relevant to the case study presented in section 4.

2.2 Mobility Management

To enable host mobility in IP networks calls for a mechanism capable of hiding the changing of the host's IP address from the upper layers whenever it roams from one addressing domain to another (i.e. location transparency). For this purpose several mobility management protocols and other mechanisms have been developed [1]. Our work focuses on network layer mobility and specifically on Mobile IP.

Mobile IP is currently the dominant macro-mobility protocol. That is it enables a mobile terminal to keep its IP address constant while roaming from one administrative domain to another, but it relies on the subnetwork to provide the micro-mobility support (e.g. access technology specific scheme, possibly combined with some micro-mobility management protocol to enhance mobility management efficiency and QoS [1]). There are two versions of Mobile IP: Mobile IPv4 (MIPv4) and Mobile IPv6 (MIPv6). Since the mobility management solution used in the case study presented in this paper is based on MIPv4, we describe only its operation here.

Mobile IPv4 (defined in RFC 3344) [14] extends the IPv4 protocol. It enables constant delivery and reception of data packets regardless of the changing location (i.e. IP address) of the user terminal (named mobile node or MN). This is achieved by associating the MN with two addresses: a home address, which is the MN's statically allocated IP address in its home network, and a care-of address (CoA), which is the node's temporary IP address while in a foreign network. In standard MIPv4, the correspondent node (CN) knows only the MN's home address, and to route packets accordingly between the two entities, Mobile IP uses home and foreign agents. A home agent (HA) is a router in the MN's home network and it is responsible for maintaining the relation (binding) between the MN's two addresses. The HA intercepts all traffic headed to the MN's home address and whenever the MN is outside its home network, the HA forwards the traffic by tunneling to the MN's valid CoA. Depending on the Mobile IP implementation, the CoA can be either the address of a local router (i.e. foreign agent or FA) or a local address obtained by the MN itself (e.g. though DHCP or PPP). In the latter case, FA functionality is included in the MN and the MN is said to have a co-located CoA. Supporting MIPv4 operation in today's Internet however requires additional mechanisms such as Reverse Tunneling (RFC 2344) and NAT Traversal for Mobile IP (RFC 3519).

2.3 Vertical Handover Support

In multi-access, handovers are not triggered just to maintain an ongoing connection but also to ensure that the user always receives the best available service and that the user preferences are met. Thus, in order to achieve an optimal handover management solution for heterogeneous networks requires not only assessing the capabilities of the available networks but also taking the user preferences and application requirements into account. This requires additional mechanisms, referred to as vertical handover support here.

In multi-access handovers can be classified as imperative and alternative based on their urgency [21]. Imperative handovers are triggered whenever the current link becomes unusable, i.e. for example when the link quality (measured e.g. in received signal strength or RSS) drops below a certain level. Alternative handovers on the other hand are triggered to get better access to the used services and therefore do not possess such urgency. The criteria for alternative handovers include QoS (i.e. the new link provides more bandwidth, smaller delay, etc.) or AAA related reasons (e.g. using the new network is cheaper).

Practical handover control solution for multi-access is a mobile-controlled handover (MCHO) scheme with automatic handover management [21]. The automatic handover triggering may be based on policies derived from user specified criteria (e.g. use WLAN if link quality > 10%, else use GPRS). Manual control is regarded as a way for the user to intercept the automatic handover process if needed.

Finally the applications place requirements on the handover performance: and they can be classified according to the type of handover management they require (i.e. fast, smooth, and normal handovers) [18], [21].

3 Application Layer Quality of Service

3.1 Defining Quality of Service

Quality of Service (QoS) is an overarching term covering different parts of end-to-end service quality. The general definition of QoS provided by the International Telecommunication Union (ITU) [5] is that QoS is *"the collective effect of service performance, which determines the degree of satisfaction of a user of the service"*. Different people and communities nevertheless interpret QoS differently, and at least the following viewpoints of QoS can be distinguished: QoS requirements of a user, QoS perceived by the user, QoS offered or planned by a provider, and QoS delivered or achieved by the provider [15]. We are discussing QoS in the user's point of view.

There are two main aspects of QoS: subjective and objective [16]. Subjective QoS essentially is the user's overall perception of service quality, that is, it is the user's opinion whether a service is working satisfactorily or not. Subjective QoS is often difficult to be specified with objective measures, at least in a way meaningful for users, and thus user-perceived quality is often expressed also non-technically [2]. Objective QoS then refers to the technical aspects of QoS, and can be specified with quantitative measures. Figure 2 illustrates different scopes of QoS (i.e. application QoS and network QoS) in a client-server communication scenario [16].

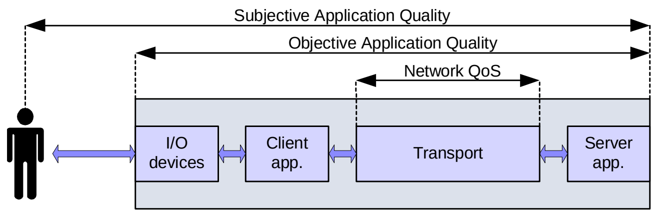


Fig. 2. Scopes of Quality of Service

3.2 Subjective Application Quality

User's Perception. The network QoS parameters, such as bandwidth, delay, jitter and packet loss, are not necessarily applicable to express subjective QoS, since a user has a high-level perspective over application performance, rather than an in-depth conception of details of the underlying implementation and operation of the network service [6]. Therefore, application quality and its variation need to be expressed in terms that describe user-perceivable effects, instead of their causes in the end-to-end transmission path. It should be noted also that subjective application quality deterioration is not solely caused by network QoS fluctuations, but is attributable to numerous other factors, including characteristics of the ongoing task (e.g. urgency), application's incompatibility with the operating system, application or protocol malfunction, disturbing factors in usage environment (e.g. faulty equipment), and so forth [7].

User-Level QoS Requirements. In the user's point of view, QoS requirements are those that are visible to him/her [6]. Thus, it is the user application that leads the main role in determining the user's QoS requirements, although neither does it dictate them solely. The salience of different quality criteria is influenced also by the goal of the interaction, and the ideal QoS profile of an application consequently varies with the task performed [3]. Three general categories for user-level QoS requirements (defined in [6]) are discussed here: criticality, cost, and security.

The criticality comprises of user-level QoS requirements characterized by the application type and data transmission (e.g. QoS metrics for a telephony service are communications continuity and voice quality). Depending on the application type, usage context, previous experience, and personal preferences, a user may consider several factors in order to come into a positive or a negative judgment of the used application or service [6], [7], [15], including service availability, session continuity, response time, throughput, reliability, media quality (video rate, video smoothness, picture detail, etc.), operability (e.g. an easy-to-use UI), etc.

Cost then represents the money value the user is charged when using a service. It is very important for the user to be able to distinguish whether the service generates costs to him/her and on what is the charging based, e.g. transmitted data (bits) or connection duration (seconds).

Finally, security requirement has several types such as confidentiality, integrity, digital signature capability, and authentication [6], and their necessity depends on the nature of the communications. Also where in the end-to-end delivery path (i.e. network, application or service) security is implemented is relevant: for instance, unless security is implemented in the application or in the service a vertical handover may jeopardize the security of the communications (e.g. when roaming to a WLAN).

It is clear that all the user-level QoS requirements can be compromised in a vertical handover as the characteristics of the underlying network connection may change even drastically. How vertical handovers affect the user experience of a service is discussed in section 4.

Methods for Measuring Subjective Application Quality. Two principal approaches for subjective application quality assessment exist: user study methods and objective measurements. The user study methods include, e.g. Mean Opinion Scores (MOS), continuous assessment, Task Performance Measures (TPMs), and qualitative methods [3]. Objective measurements, on the other hand, rely on measurement of some application quality metric(s) (e.g. Peak-Signal-to-Noise-Ratio (PSNR) for video) [20]. We chose to use MOS in our user study to collect the subjects' opinions of the experienced service quality. In short, MOS enables performing controlled assessment of subjective QoS with untrained subjects and controlled levels of quality [3]. The method employs a 5-point scale, according to which subjects judge the experienced quality after conducting a task. The given ratings are then averaged across the subjects to get the final MOS.

4 Case Study in Assessing Subjective QoS in a Heterogeneous Network

The purpose of the presented case study was to evaluate how host mobility in a heterogeneous network environment affects usability and subjective quality of different web-based services. The case study was organized as a task-based user evaluation, where test users conducted a set of predefined tasks in the real environment of use involving a multi-access network with automatic mobility management realized with a Mobile IP implementation.

4.1 Multi-access Network

The main access networks available to a mobile user in Oulu region are GPRS/EDGE, UMTS, and WLAN. In this study, the Octopus GSM/GPRS network [11] and panOULU WLAN network [13] were used. panOULU is a public access network based on IEEE 802.11a/b technology and provides free-of-charge wireless Internet access in different parts of the city. Host mobility between the two networks was enabled with Secgo Mobile IP, a commercial product based on the Mobile IPv4 standard. Secgo's solution has two system components: Secgo Mobile IP server, which in our setup is used to provide the home agent functionality, and Secgo Mobile IP client, which implements the mobile node functionality.

In our setup presented in Figure 3, the Secgo MIP server application was running on a Linux computer located in a local area network (denoted OPOY's network in the figure), of which public IP addresses are provided by Oulun Puhelin Plc. (OPOY). During the testing, the MN never roams to its home network. The panOULU WLAN and the Octopus GPRS network are the visited domains, in which the MN uses a co-located CoA, that is, there are no FAs deployed in either network. As shown in the figure, the roaming user's Internet connection was always routed via the home network and thus the location of the HA is optimal in terms of routing delay.

The MN has a public home address in the OPOY's network (i.e. 212.50.147.109) but the used care-of addresses are always private addresses since both panOULU WLAN and Octopus GPRS are located behind a NAT. Due to the presence of firewalls and NAT devices both reverse tunneling and NAT traversal are used. The MIP client is configured so that WLAN had a higher priority than GPRS and a WLAN-to-GPRS handover occurred if the link's signal quality dropped below 20%. To avoid the ping-pong effect the link reselection quality was set higher, i.e. the MN did not roam back to WLAN unless the quality of the link was higher than 40%.

4.2 Device Setup

Since Secgo MIP client is only available for Windows 2000/XP and Linux (≥2.2), the choice of the user terminal was limited to a laptop computer. The devices used during the testing included a laptop (IBM ThinkPad R40) and a PDA (Fujitsu Siemens PocketLOOX 610 BT/WLAN), which was connected to the laptop with an USB cable. The laptop's network interfaces included an integrated IEEE 802.11 WLAN radio (Intel PRO Wireless LAN) and Nokia D211 multi-mode card, which was used for GPRS access. The communication between the two devices was enabled with ActiveSync v3.7.1 application that was set to a pass-through mode, i.e. the traffic emanating from the PDA was sent by using the laptop's IP address. Applications running on the laptop computer included Secgo MIP client v3.0.6 and Ethereal 0.10.6, configured to monitor HTTP traffic passing through the Mobile IP virtual network adapter. The PDA ran PocketPC2003 OS and Pocket Internet Explorer web browser.

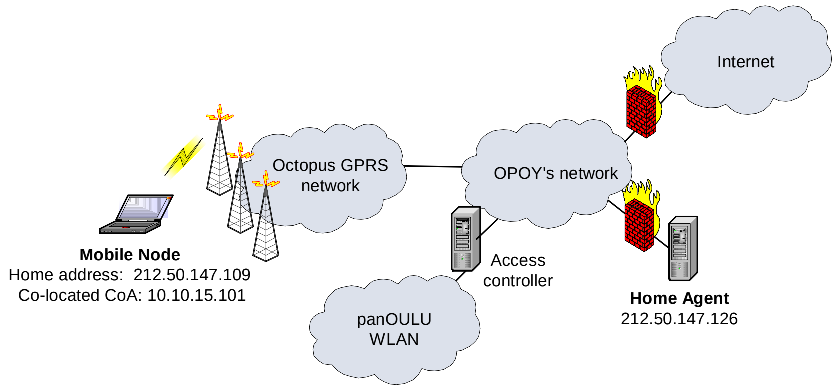


Fig. 3. Test network topology

4.3 Tasks

The user study was organized into tasks that were based on using the Digital Oulu Cultural Database web service (abbreviated DOK). DOK offers different types of services, including browsing of XHTML pages, searching for cultural objects, downloading multimedia presentations of the objects (text, images, and audio and video files), and streaming audio and video. The image and video content in DOK are adapted separately for different device types: desktop/laptop computers, PDAs, and Nokia S60 and S40 mobile phones. The user device is identified at the server-side from the received HTTP request message. Since access-based adaptation is not supported and the same user device will be used throughout the test, the test users will receive the same sized content regardless of the underlying access network.

There were two types of tasks (listed in Table 1). A Type 1 task comprised of browsing XHTML pages consisting of 2-4 small objects (approx. 1-20 KB in size). The pages contained text and/or image(s). A Type 2 task involved downloading of a video file (approx. 2 MB in size).

Table 1. Tasks performed during the test

Task ID	Access Network	Task Type
1	WLAN	1 (text+image)
2	WLAN	1 (text+image)
3	GPRS	1 (text+image)
4	GPRS	1 (text+image)
5	WLAN	2 (video)
6	GPRS	2 (video)
7	GPRS/WLAN	1+2 (text+image+video)

The users were asked to walk a given route around the city of Oulu and perform the different tasks. The route was carefully defined so that in the first six tasks only one access is used, whereas in the last task the users could move around freely in a specified area where vertical handover(s) were likely to occur. This way it was ensured that the users were exposed to as controlled levels of QoS as possible.

The subjects were told prior to the testing that they will be using different access networks during the test. However, they were not provided with any information of the used access networks while conducting the test. Thus, the users had a complete transparency to roaming in the sense that they did not have to modify any settings or connect to any access networks before they were able to use the services.

4.4 Data Collection

The test includes both subjective and objective evaluation of application layer QoS. Research data was collected with questionnaires and both client-side and server-side logging of user sessions.

Subjective user data was collected with three questionnaires: A pre-test questionnaire gathering background information, a questionnaire filled during the test with task-specific questions of the service's operation in different situations (including quality ratings), and a final questionnaire filled in a debriefing session. The questionnaire filled during the test included a scale for subjective evaluation of the experienced level of QoS. After conducting a task the users were asked to assess the service's quality according to the six-point Likert scale defined in Table 2. A six-point scale was used instead of the traditional five-point to avoid vague results [17].

Table 2. The six-point Likert scale used in evaluating subjective application quality

Score	Description
1	Excellent: "The service worked impeccably."
2	Good: "The service worked well, I noticed only few deficiencies."
3	It was all right: "The service worked sufficiently well considering the purpose of use."
4	Somewhat poor: "The operation of the service was a bit annoying, but I would use the service anyway."
5	Unsatisfactory: "I would use the service only if it was absolutely necessary."
6	Unusable: "I could not use the service at all."

In addition to the MOS scores, users' experiences were collected in the questionnaires with open ended questions and questions associated with 1-7 rating scales. This is simply because MOS scores can convey only a limited amount of information of the experienced service quality, e.g. it does not reveal any information of the reasons behind the given evaluations.

Objective measurement data was collected by capturing user traffic with Ethereal and Secgo MIP client's log at the client side. This allowed clocking actual download times, and to map this them to the subjective evaluations. The available data was sufficient for determining which access network was used at a given point of time, but not for making precise measurements of the performance of the vertical handover.

4.5 Test Users

Altogether 20 test users participated in the study. All of the participants were Finnish and 50% of them were men and 50% women. Majority of the users were quite young: 30% (6) were 20-24, 60% (12) were 25-34, one subject (5%) was 35-44, and one decided not to report his/her age. The professional background of the test users grouped by field is: 40% IT professionals or students, 15% data processing science teaching staff or researchers, 25% economics teaching staff or researchers, and 20% other. All the subjects use at least some kind of an Internet service (basically e-mail and web browsing) more or less daily, and use mainly fixed access technologies.

Test users were asked to familiarize themselves with the UI, content, and operation of DOK beforehand. It was assumed that if the users were familiar with the service to begin with, they would not focus so much on the content or UI related issues that remain more or less static throughout the test. Since 35% of the subjects reported that they had never used a PDA before, they were also given an opportunity to try out the device before the test.

4.6 Main Results

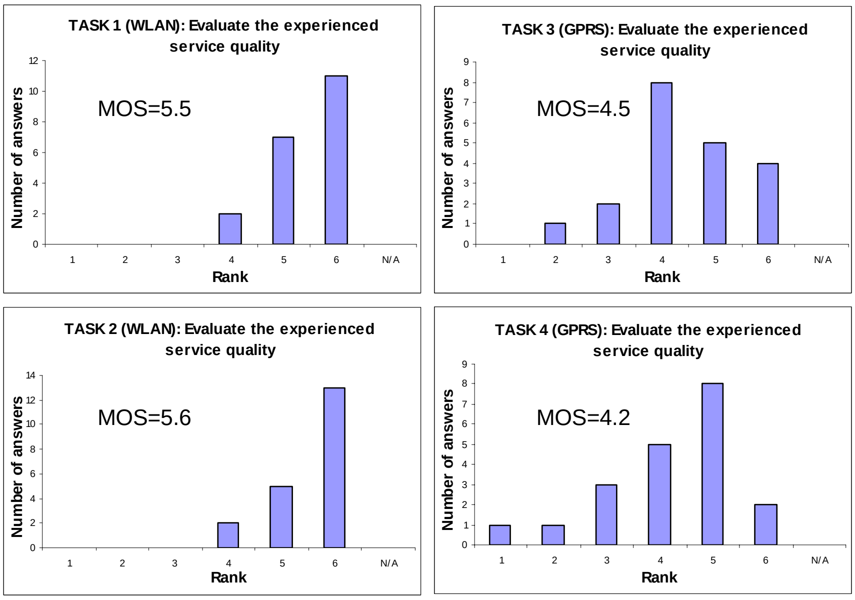
The results consist of MOS ratings representing the experienced levels of QoS in each task, qualitative data providing more information of the user experiences, and objective measurements of the service's operation.

Subjective QoS Evaluations. The distributions of the QoS evaluations on the six-point Likert scale (Table 2) in each task are presented in Figure 4. According to the obtained results in the Type 1 tasks, the service worked very well via WLAN access but with GPRS the service can be considered as barely satisfactory.

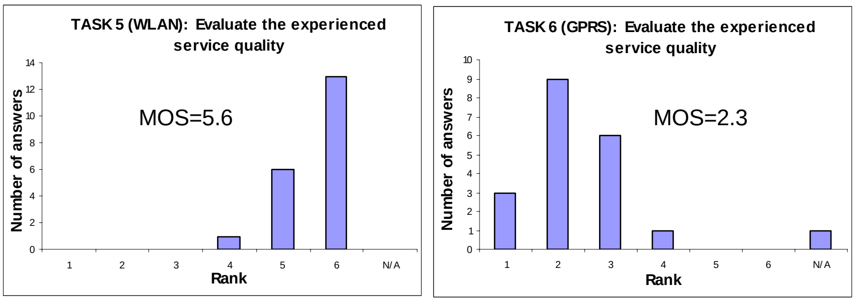
After conducting each task the users were asked to compare the experienced service quality with the previous task. Interestingly, only 65% of the users reported that they had noticed degradation in the service quality relative to the previous task when conducting Task 3, i.e. the first Type 1 task that was performed using GPRS after conducting Tasks 1 and 2 (and a Type 2 task 5) via a WLAN connection.

In the two Type 2 tasks, the users were asked to assess the quality of a video file download. The size of the video files downloaded was 1.9 MB and the first video download was done via a WLAN connection (Task 5) and the second via a GPRS connection (Task 6). The users were in general very pleased with the quality (speed) of the first download, whereas the second received very low ratings: 30% of the subjects considered the service quality to be “Somewhat poor”, 45% “Unacceptable”, and 15% “Unusable”. Also in Task 6, 60% of the subjects had interrupted the video download themselves, thus it can be considered practically unusable.

The users were in general pleased with quality of the video presentation. The obtained results indicate however that offering files of this size to be downloaded over GPRS is not practical in this type of a service. Whether streaming services would perform better in this case will be studied in future work.



(a)



(b)

Fig. 4. Subjective evaluations in (a) Type 1; and (b) Type 2 tasks

In Task 7, the users were asked to move while using the service in an area where vertical handover(s) are likely to occur. In this task 35% of the users reported that they had noticed no changes in the service quality, 40% had noticed the quality to change once, and 25% had noticed 2-3 changes in the service quality. The number of vertical handovers per usage session is presented in Figure 5.

However, the open-ended questions reveal that the reported occurrences of quality changes were not always correlated with the occurrences of vertical handovers. Only 30% of the subjects had written that they had noticed a clear improvement in the service quality – mainly when downloading the video but a couple of users had also noticed some quality variation in other parts of the service, as well. 65% of the users had not noticed that their location had any effect on the service quality and one subject (5%) gave no answer.

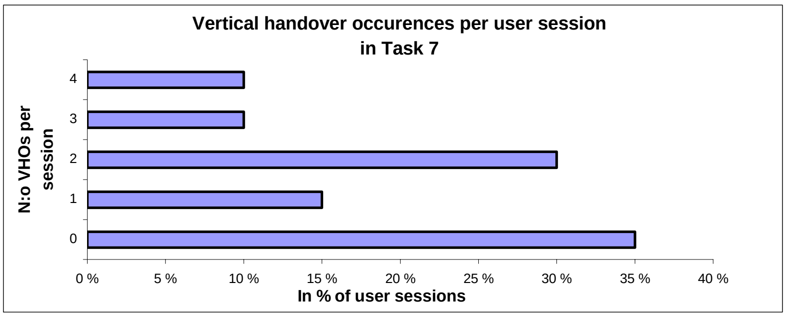


Fig. 5. Vertical handovers per user session

The fact that the users missed most of the vertical handover instants is due to the properties of the used service and the task realization. Using a web-browsing service is not continuous and thus small changes in the quality are difficult to perceive. Also the period of time, during which the other access network was used after a handover was in some cases so short that the user did not even download anything before a handover in the reverse direction had already occurred. Also using two types of services (i.e. basic browsing and video downloads) in a same task was a bit confusing; many of the users understood poor video downloading performance as a change in quality since the basic browsing had worked well or least satisfactorily in their opinion, even if GPRS was used all the time. To obtain proper results in this scenario would require the usage of a service that generates a continuous traffic pattern (e.g. streaming). Due to these problems, MOS ratings could not be used for the analysis of the experienced service quality in this task, and thus the following comments are solely based on the gathered qualitative data.

Changes that increased the service's operational speed were considered very positive but changes in the opposite direction were irritating to the subjects. The subjects who had noticed a clear improvement in speed at some point of the video download in Task 7 were positively surprised. However, in these cases it is highly probable that the subjects let the downloading proceed slowly at first due to the test situation and this way they were able to experience the effect of the handover. In general, after experiencing the slow video download speed in Task 6 most subjects were very eager to interrupt the download in this task as they noticed that it was so slow again. Thus, to be of any use the downward vertical handover has to take place in the beginning of the download of a large file. But without getting any information of anticipated handovers the user does not have any means to predict the situation and will most likely discard the download as unusable as soon as he/she becomes frustrated with it.

In all of the tasks, there were no significant differences in the given ratings between technically oriented and novice users. Reasons for this may be well-defined test scenarios or the familiarity of the used service types (i.e. browsing and file download) to the users.

Acceptability of QoS Variations. After completing the tasks the users were asked to assess the degree of the experienced QoS variation in the two types of task on scale 1 (imperceptible) – 7 (high). The chart in Figure 6 shows that video download introduced high subjective QoS variations, which is expected. The chart in Figure 7 depicts the users' assessment on how acceptable the QoS variation experienced during the test was.

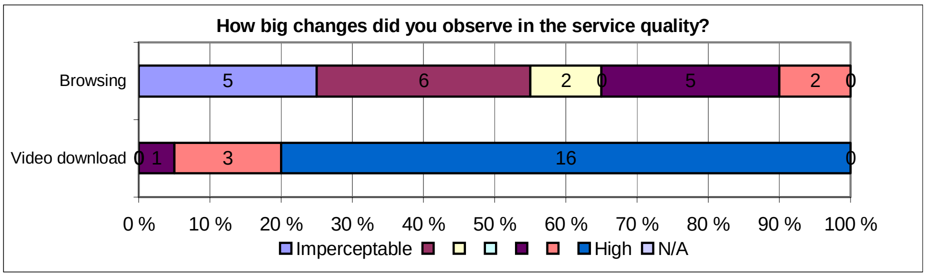


Fig. 6. Degree of the experienced QoS variations in different task types

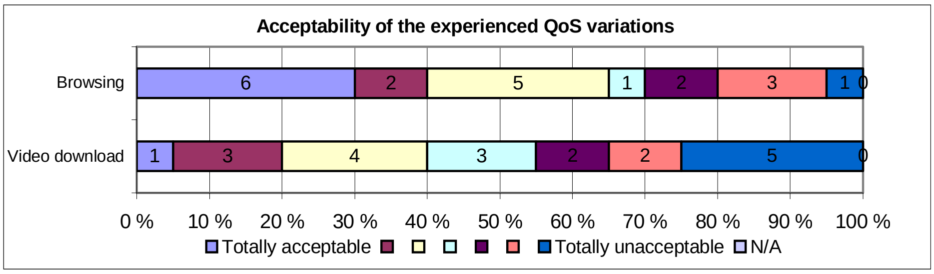


Fig. 7. Acceptability of the experienced QoS variations

There is a high variation in the obtained results. To summarize the opinions obtained from the open ended question related to this evaluation, the acceptability of fluctuating QoS was not only dependent on the degree of the variation but also on the type of the used service: basic browsing was expected to work impeccably all the time whereas some variation can be tolerated in the video download due to the novelty of the technology in the mobile domain of usage, and how usable the service was under the poorest QoS: browsing over GPRS was still operable whereas downloading the video over GPRS was not reasonable.

User Viewpoint on Inter-technology Roaming. After completing the test the users were given a short description of the characteristics of WLAN and GPRS networks followed by few questions related to inter-technology roaming.

Based on the chart shown in Figure 8 the users clearly understood the advantages of the multi-access capability. As expected, WLAN connection (fast and free-of-charge but small coverage) was preferred to GPRS (slow, chargeable, and ubiquitous coverage), but it was acknowledged that service availability is more important than the characteristics of the underlying access network.

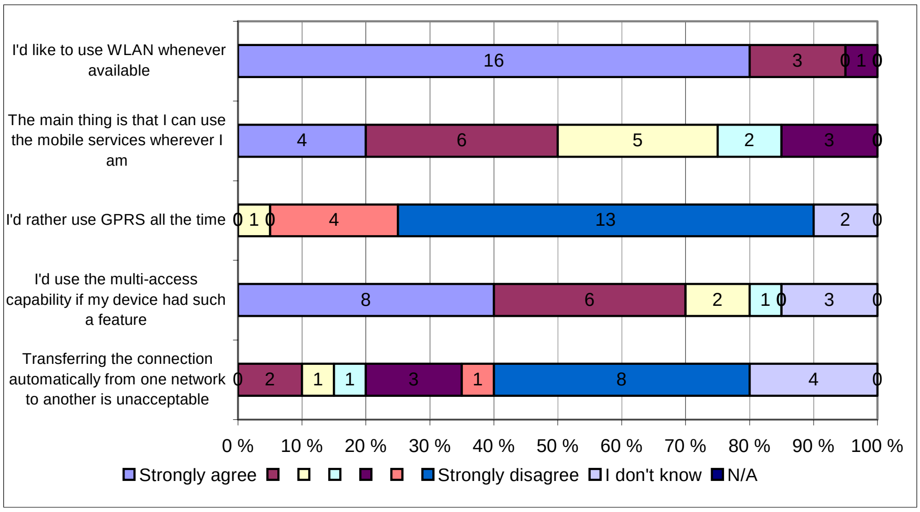


Fig. 8. Users' opinions on inter-technology roaming

A majority of the users would utilize inter-technology roaming if their mobile device had such a capability. The fact that using GPRS is not free-of-charge was the main argument against transferring the user's connection automatically from one access network to another. However, notifying the user or asking for his/her consent when roaming to a chargeable connection was considered sufficient in this case.

The subjects were also asked how the selection of the access network and handover management should be carried out. The preferred roaming solution was automatic handover management (90% of the answers) complemented with some sort of notification or other solution that enables the user to be aware with which access network he/she is using at any time (required by 60% of the users). Only two subjects preferred manual control of vertical handovers. It should be noted however that in this study the users did not get to try manual network switching.

5 Discussion

We presented an empirical task-based user evaluation of the end user QoS of a web service in a multi-access network utilizing automatic mobility management between WLAN and GPRS networks by Mobile IP. The empirical results and feedback from test users show that this type of a multi-access network combined with automatic mobility management offers a promising setting for provisioning future mobile multimedia services.

Future work will focus on addressing the limitations of the present study. The Mobile IP client needs to be ported to the actual mobile device. Services involving streaming data have to be included in the evaluation, as they are good candidates for future services in multi-access networks. Other access network technologies might be considered, as well. Further, solutions for service adaptation and user feedback need to be studied to obtain better user experience in heterogeneous networks.

References

1. Banerjee N, Wei W & Das SK (2003) Mobility Support in Wireless Networks. IEEE Wireless Communications 10(5):54-61.
2. Bouch A, Kuchinsky A, & Bhatti N (2000) Quality is in the Eye of the Beholder: Meeting Users' Requirements for Internet Quality of Service. CHI Letters 2(1).
3. Bouch A, Sasse MA & DeMeer H (2000) Of Packets and People: A User-centered Approach to Quality of Service. Proc. Eighth International Workshop on Quality of Service, Pittsburgh, PA, 189-197.
4. Gustafsson E & Jonsson A (2003) Always Best Connected. IEEE Wireless Communications 10(1):49-55.
5. ITU-T Recommendation E.800 (08/94) Terms and definitions related to quality of service and network performance including dependability.
6. Jamalipour A. (2003) The Wireless Mobile Internet: Architectures, Protocols, and Services. John Wiley & Sons Ltd., West Sussex, UK.
7. Miras D (2002) A Survey of network QoS needs of advanced Internet applications, http://qos.internet2.edu/wg/apps/fellowship/Docs/Internet2AppsQoSNeeds.pdf
8. Moby Dick Project (2004) http://ist-mobydick.org.
9. Mohr W (2002) Heterogeneous Networks to Support User Needs with Major Challenges for New Wideband Access Systems. Wireless Personal Communications 22:109-137.
10. Newman W & Lamming M (1995) Interactive System Design. Addison-Wesley.
11. Octopus (2004) http://www.mobileforum.org/octopus/.
12. Pahlavan K, Krishnamurthy P, Hatami A, Ylianttila M, Makela, J-P, Pichna R & Vallstron J (2000) Handoff in Hybrid Mobile Data Networks. IEEE Personal Communications 7(2):34-47.
13. panOULU (2004) http://www.panoulu.net.
14. Perkins CE (1998) Mobile IP: Design Principles and Practices. Addison-Wesley, USA.
15. Räisänen V (2003) Implementing Service Quality in IP Networks. John Wiley & Sons, West Sussex, UK.
16. Scaefer C, Enderes T, Ritter H & Zitterbart M (2002) Subjective Quality for Multiplayer Real-Time Games. Proc. First Workshop on Network and System Support for Games, Braunschweig, Germany, 74-78.
17. Serif T, Gulliver SR & Ghinea G (2004) Infotainment Across Access Devices: the Perceptual Impact of Multimedia QoS. Proc. ACM Symposium on Applied Computing, Nicosia, Cyprus, 1580- 1585.
18. Sun J & Sauvola J (2004) Mobile IP Applicability: When Do We Really Need It? Proc. 2004 International Conference on Parallel Processing Workshops, Montreal, Canada, 116-123.
19. VHO Project (2004) http://www.cs.hut.fi/~pmrg/VHO.html.
20. Wang Z, Banerjee S & Jamin S (2003) Studying Streaming Video Quality: From an Application Point of View. Proc. o11th ACM International Conference on Multimedia, Berkeley, CA, 327-330.
21. Zhang W, Jaehnert J & Dolzer K (2003) Design and Evaluation of a Handover Decision Strategy for 4th Generation Mobile Networks. Proc. IEEE Vehicular Technology Conference, Orlando, FL, 3:1969-1973.
22. Zhuang W, Yung-Sze Gan Y-S, Loh K-J & Chua K-C (2003) Policy-based QoS-management architecture in an integrated UMTS and WLAN environment. IEEE Communications Magazine 41(11):118-125.