Evaluation of Perceived QoS with Multimedia Applications in a Heterogeneous Wireless Network

Nemanja Vučević, Francisco Bernardo, Anna Umbert, Łukasz Budzisz

Department of Signal Theory and Communications (TSC), Universitat Politècnica de Catalunya (UPC)
Campus Nord, D4, Jordi Girona 1-3, 08034-Barcelona, Spain
`[vucevic, fbernardo, annau, lukasz]@tsc.upc.edu`

Abstract— **This paper presents a study addressed to evaluate the variation in perceived QoS experienced by a user running multimedia applications in a heterogeneous wireless network. To perform this evaluation, a real-time testbed emulating an all-IP B3G network that includes UMTS Terrestrial Radio Access Network (UTRAN), GSM/EDGE Radio Access Network (GERAN), and Wireless Local Area Network (WLAN); and the corresponding common core network (CN) based on DiffServ technology and Multiprotocol Label Switching (MPLS) has been used. Results for videostreaming and video conference applications when running under interruptions due to a vertical handover (VHO) are included in this paper, pointing out the variations in behaviour of different applications.**

I. INTRODUCTION AND MOTIVATION

Uprising trends in modern communications suppose heterogeneity of communications technologies. The third generation (3G) and its evolution, Beyond 3G (B3G) mobile systems, are taking central role in mobile/wireless access networks and their main objective is to guarantee a negotiated Quality of Service (QoS) during the entire session duration. To affront this challenge, Common Radio Resource Management (CRRM) algorithms have been adopted to coordinate the operation in the radio part seamlessly [1]. Moreover, different QoS policies that coordinate control between core and radio network parts must be defined to provide the end-to-end (e2e) QoS.

At the same time, IP technology is becoming the cornerstone around which wireless networks are converging. The concept of all-IP in wireless networks is commonly used to refer to those systems which include IP-based multimedia services and IP-based transport. In addition, the mobile backhaul is directed towards IP networks and inherits QoS solutions developed for the Internet [2].

The provision of new multimedia services in next generation networks seems to be a highly important issue for both operators and manufacturers and a hot research topic. Future mobile users will access the multimedia contents through B3G networks. Thus, in order to learn about the tuning of these complex networks and the usage of suitable multimedia contents, it is very important to know the impact of such networks on the QoS perceived by the users when using multimedia services.

Our work is devoted to evaluate the variation in perceived QoS experienced by a user running multimedia applications in a heterogeneous wireless network. To perform this evaluation, a real-time testbed [3] emulating an all-IP B3G network that includes UMTS Terrestrial Radio Access Network (UTRAN), GSM/EDGE Radio Access Network (GERAN), and Wireless Local Area Network (WLAN); and the corresponding common core network (CN) based on DiffServ [4] technology and Multiprotocol Label Switching (MPLS) [5] has been used. The QoS in CN is supported by a Bandwidth Broker (BB) entity. Wireless QoS Broker (WQB) with CRRM support is responsible for QoS in the radio part; and manages the e2e QoS negotiation with BB entity. In addition, during session the QoS may be dynamically renegotiated. The real-time testbed is implemented with off-the-shell Personal Computers (PCs) and allows having realistic results (based on real applications) that will give insight into the tuning and developing of QoS-aware B3G networks.

The possibility to evaluate how different network procedures impact over the user perceived QoS of real multimedia applications, like NetMeeting® and QuickTime® (QT), is something really interesting of publication because the real-time testbed allows capturing effects that are hardly captured with analytical models and simulations.

The paper is organized as follows. Section II describes the testbed used to do this study. In Section III, the methodology applied to evaluate the perceived QoS experienced by a user is detailed. Next, Section IV presents some case studies and results obtained. Finally, section V provides the conclusions to the taken study.

II. HETEROGENOUS WIRELESS ACCESS NETWORK TESTBED

A real-time testbed currently emulating a B3G network has been used to study the perceived QoS experienced by a user running multimedia applications in a heterogeneous wireless network.

This testbed has been developed in AROMA project [3] and is an enhanced version of EVEREST testbed [6] since it includes recent technological solutions. The B3G network implemented in the testbed is composed of several RANs emulators: the UMTS Terrestrial Radio Access Network (UTRAN), GSM/EDGE Radio Access Network (GERAN), and Wireless Local Area Network (WLAN). The UTRAN emulator includes both release 99 (R99) channels and High Speed Packet Access (HSPA) in both uplink and downlink. This heterogeneous RAN is connected to IP Core Network (CN) that is based on DiffServ technology and Multiprotocol Label Switching (MPLS).

The CN is composed of seven routers, three of them are edge routers (2 ingress routers and 1 egress router), and the

remaining four are core routers which interconnect all the edge routers and form a fish model topology slightly unbalanced. The different RANs enter the CN through one of the ingress routers depending on the configuration. The scenario selected for the testbed considers that UTRAN is connected to one ingress router whereas WLAN and GERAN are connected to the other.

As it has been said, WQB handles QoS management in the radio part as well as CRRM functions, BB manages QoS in the CN and, MPDP acts as a master broker taking the final decision on the acceptance of a new user flow. The initial negotiation of the QoS during session establishment as well as QoS re-negotiation procedures have been implemented in the testbed.

The developed platform constitutes a realistic framework to test different e2e QoS strategies and to evaluate the QoS level experienced by a user that is immersed in a heterogeneous mobile environment with IP connectivity. Real client-server IP based applications are executed in the edges of the testbed and the perceived QoS will be measured once the real IP packets have passed through the testbed. This framework allows, at the same time, the testing of particular implementation of the QoS entities which may be important for operators before putting these implementations in their real networks.

The presented B3G real-time testbed for an all-IP heterogeneous network is a powerful tool for carrying out realistic trials, usually not achievable by means of non-real-time simulations. However it is worth mentioning here, that it is not the aim of this paper to describe the testbed but the e2e QoS perceived using it. The power of testbed is in that it permits projection of emulated network behaviour on real applications.

III. E2E QoS Measurement Framework

Metrics for evaluation of the perceived QoS spin mostly around technical quantitative results. These methods are usually referred to as *objective* perceived QoS and a metric of the experienced QoS is given without human influence. On the other hand, *subjective* methods take into account the human perception of the multimedia content.

In our study we use a full-reference model-based objective metric [7] for the QoS evaluation that is based on ITU recommendations [8][9]. This kind of methods compare a reference sample of the media under test with a degraded sample obtained at the output of the system (e.g., our testbed emulating a B3G network). As a result, a degradation level measurement is given by the QoS evaluation method. This degradation level tries to express the subjective evaluation that human beings would have given to the experiment. Then, degradation level is expressed as a number between 0 and 1. The degradation level of 0 corresponds to a perfect quality perception (e.g. the original video and a not degraded copy are compared), while degradation level of 1 means complete loss of information. Nevertheless, these methods rarely give a degradation level equal to 0 since human perception is reluctant to give the maximum score (i.e. perfect perceived

quality) even if the compared videos are almost equal. The same is applicable to the worst score.

The process followed to make the QoS measurements takes three major stages. First, specification of the media under test is done, selecting suitable contents (regarding movement, color, image definition) and suitable codecs and bit-rates. Table I specifies the codecs and bit-rates used in the experiments.

Second, the testbed is configured to emulate the desired behavior of the B3G network (e.g., number of users, core network load, CRRM policies, vertical handovers (VHO), etc.). As is described in section IV, special attention has been paid to Vertical Handovers (VHOs). These are usual events that can appear during a session in a heterogeneous wireless network; and present a critical point in e2e QoS provisioning. Different mechanisms included in e2e QoS support during the VHO determine the loss and delay. For example packets may be dropped during the change of radio access technology (UTRAN, GERAN, WLAN) execution; or may be tunneled to a new attachment point causing an extra delay. Another point that may contribute is a bandwidth limitation that can cause packet drops in the moments of burs (bitrate peaks).

Finally, the media is captured and analyzed to obtain the objective QoS metric stated above. An illustration of the described process is shown in Fig. 1. It can be seen that a client-server application is run in the edges and the data packets flow through our testbed. Then, the testbed impairs the data packets accordingly to the emulated RAN and the CN load (i.e. packets delays or drops depending on the load, coverage, handovers, etc). The modified videos are captured in the Client side with Camtasia® software [10]. Then, they are compared with a reference video and the degradation level is obtained.
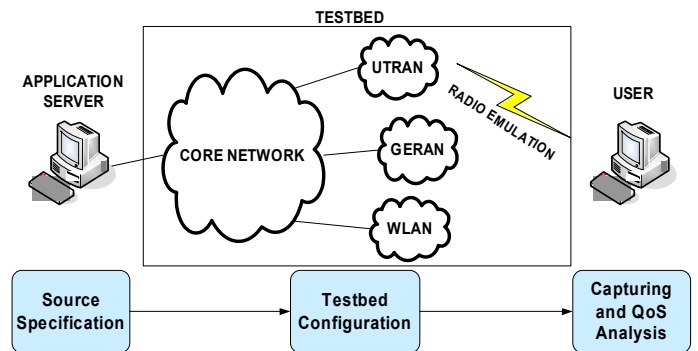

Fig. 1 QoS evaluation process

The degradation evaluation of the application bases on the comparison of the images captured when application has passed with and without distortions. The handover will usually make video application freeze the image during the dropping. After the dropping ends, the application will recuperate, and during certain time the video will be changing but with significantly lower quality (see Fig. 2). The bandwidth limitation will make this recuperation period longer, and it may last until the application ends.

Fig. 2 Distortion of the video.

Regarding the application, suitable and aligned with the state of the art applications have been installed in the testbed to cover the services envisaged for 3G heterogeneous networks (i.e., conversational, streaming, interactive and background). Nevertheless, the study presented in this paper is focused on conversational and streaming services as representative ones. Thus commercial applications for video streaming and video conference have been used.

Table I sums up the applications and multimedia contents used in this work. It can be seen that codecs and bit-rates are aligned with 3GPP specifications [16][17].

In next section, some illustrative results that reflect diversity in behaviour that applications may express are given.

IV. CASE STUDIES

The AROMA testbed functionalities were adjusted for specific tasks under test. For more set of functionalities of the testbed the reader is referred to [18]. The effect we are considering is the influence of the VHO that provokes complete loss of IP packets. For the more precise time adjustments and VHO policy execution, the UUT (User Under Test) has been set as static with full coverage of the UTRAN and WLAN. The user starts a session in UTRAN, then after certain time, a VHO is forced from the network, and it is passed to WLAN. As mentioned those two radio access technologies are connected to different ingress routers, involving the CN in the QoS negotiation and thus making the HO duration longer. In testbed, the VHO duration can also be tuned, so the desired values of handover duration can be obtained with a precision around 0.5s.

Three case studies are considered, one examining influence of VHO duration on video streaming applications and resistance of the applications to losses, other dealing with resistance to the VHO appearance moment, and the third one is the same as the first one just examining video conference tool results, and comparing results obtained for different codecs.

A. Case Study 1

The first case study considers video streaming application analysis under VHO of different duration. For the test QT (QuickTime) and VLC (Video LAN Client) are considered. The users are streaming a 30s long video of 128kbs (average bitrate), however, the video is also accompanied with the 32kbps audio signal (average bitrate) whose degradation is not evaluated in this study. Anyhow, as the intention is to measure pure VHO influence on service degradation, we do not want to have the distortions due to the channel bandwidth limitation in cases of burst peaks of application. Therefore, the user is connected to the both access technologies with 384kbps when executing this application.

The VHO duration according to examples in [19] reaches 10.720s, so we test here influences that VHO may have up to length of 10s. In Fig. 3 the degradation level of video when exposed to a different VHO duration (complete loss) is presented. The handover is triggered 5s after the session initialization.

It can be observed that QT application expresses slower degradation increase than VLC application with increase of VHO duration. In both cases, the degradation is higher for CIF (Common Intermediate Format) than QCIF (Quarter CIF) video.

The QCIF format passed with VLC shows clear tendency of degradation increase with increase in VHO length. At the same time this tendency is expressed in QT only after the VHO duration overcomes 6s. This property depends on the applications immunity to cut-offs and will be explained in Case Study 2. In almost all the cases the VLC expresses higher degradation than QT which will also be something expectable after the next section.

All the analysed tests show that the level of degradation in a video streaming application is lower for CIF video format regardless of the duration of the interruption generated for example during a vertical handover.

TABLE I
APPLICATIONS AND MEDIA SPECIFICATION

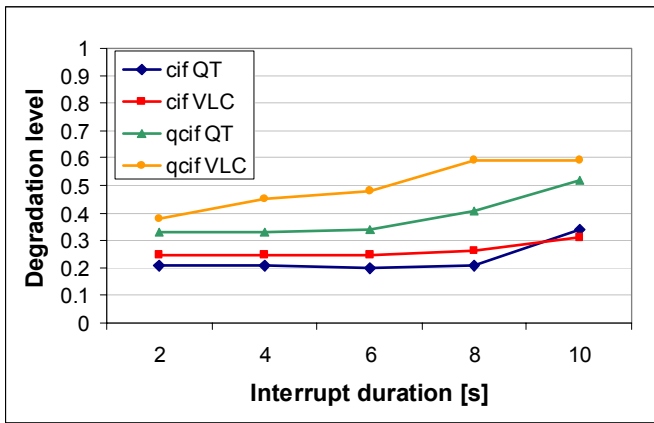| Service	Applications	Contents				
		Description	Size	Length	Video Codec	Bit rate
Video Streaming	Darwin Streaming Server [11] + QuickTime® (QT) [12]	Fragment of movie	CIF / QCIF	30 s	H.264	64 /128 kbps
	Video LAN Client (VLC) [13]					
Video Conference	Microsoft NetMeeting® [14]	Person speaking with static background	QVGA	30 s	H.263	128 kbps
	Video Conference Tool (VIC) [15]					

Fig. 3 Degradation level of streaming video when VHO duration changes. The tested video is of 128kbps.

B. Case Study 2

In the second case the study of QoS dependence on the VHO start time during the session has been observed. For this reason, the triggering moment when VHO initiates has been changed. The difference in time between this triggering of VHO and session initialization will be referred to as interrupt initialization in further text.

Again, the 30s long videos were streamed. The videos are of 64kbps (average bitrate) with audio of 24kbps (average bitrate), streamed over the 256kbps connection. The VHO duration was 6s in all the tests.

Fig. 4 shows that VLC's behaviour is independent of the beginning of the interruption. However, QT presents lower degradation when VHO appears later.
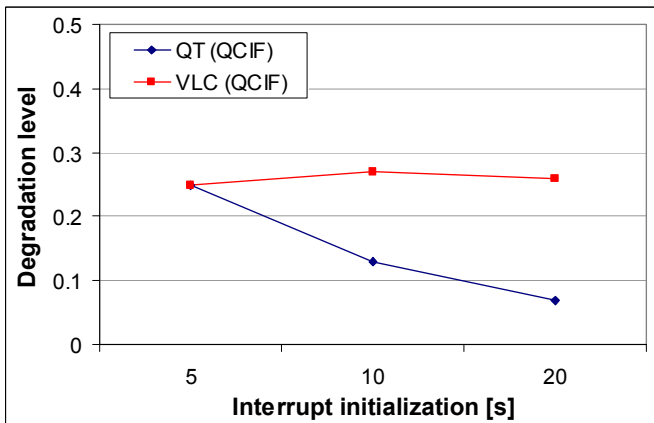

Fig. 4 Degradation level of streaming video when VHO initialization time changes. The VHO duration is 6s. The tested video is of 64kbps.

This property of "immunity" to packet loss in QT is due to the application implementation that is enabling it to use all the available bandwidth when the application rate is under the channel's bandwidth limit. Thus the application is sent faster than it is seen and the user does not observe a cut-off of 6s if during the previous lifetime of the session it has accumulated

packets on the user's side buffer. That is why with later interrupt initialization the UUT is experiencing less degradation. This is also the reason why in Fig. 3 the increase in the QoS degradation for QT is noticeable only for lengths of the VHO higher than 6s.

C. Case Study 3

This case study is testing the same VHO conditions with video conference as application under test. In order to have comparable results, the video has been captured with web cam, and then as a capturing device Camtasia® recorder has been selected on the peer sending the video over VIC (Video Conference Tool) or NetMeeting®. The Camtasia® then captures the recoded video played in another window, and passes it to the videoconference applications as if it was captured with a camera device

The aim of the test is to see if the two applications experiencing VHOs of different lengths during the session of 30s give different results. The streamed video was of 128kbps with audio of 64kbps. In case of VIC, audio was passed through RAT (Robust Audio Tool). The applications use H263 codec with the QVGA video size.
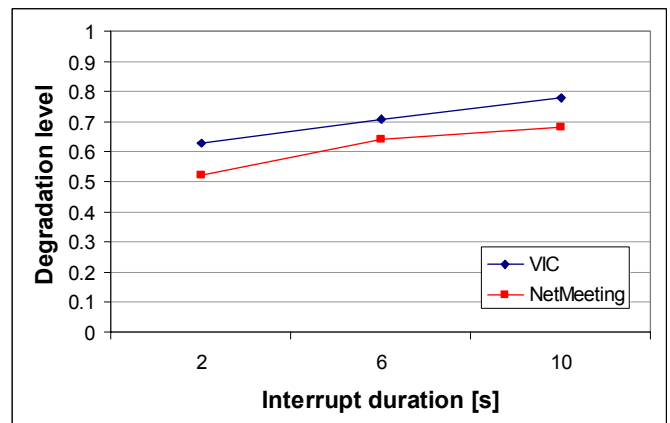

Fig. 5 Degradation level of video conference video when VHO duration changes. The tested video is of 128kbps.

From Fig. 5 it can be seen that the behaviour of the NetMeeting® gave a bit better performances in all the cases. However, all the degradation seems to be quite high, which leads us to conclude that the codecs used in video conference made worse impression in QoS evaluation. It is important to notice that both applications still show clear tendency of increasing the degradation level with increase of the VHO duration, as expected.

V. CONCLUSIONS

The difference in robustness that applications have expressed in this study may serve as motivation to include the knowledge acquired (application-awareness) in the configuration of B3G networks. In that sense, QoS class mapping, QoS policies, CRRM mechanisms, etc. may be adapted to take into account the application that users are running. For some of the applications, previous session

duration may also be valuable information when deciding on VHO execution. In this paper the evaluation of the e2e QoS experienced by a user in a heterogeneous mobile environment with IP connectivity under realistic conditions is presented. To perform this study a real-time testbed currently emulating a B3G network has been used to project the effects of the wireless all-IP environment on the real applications.

REFERENCES

[1] A. Tolli, P. Hakalin, H. Holma, "Performance evaluation of common radio resource management (CRRM)". *In Proc IEEE* ICC 2002. Vol 5, pp: 3429- 3433. 2002

[2] 3GPP TR 22.978: "All-IP Network (AIPN) feasibility study", Available: http://www.3gpp.org/

[3] A. Umbert, et al, "Testbed Specification", Public Deliverable D07, AROMA IST-4-027567. Available: http://www.aroma-ist.upc.edu/

[4] S. Balke et al, *An Architecture for Differentiated Services*, IETF RFC 2475, Dec. 1998

[5] E. Rosen, et al, *Multiprotocol Label Switching Architecture*, IETF RFC 3031, Jan. 2001

[6] R. Ferrús, et al, "EVEREST Testbed: QoS Management Evaluation in B3G Networks", TRIDENCOM, March 2006, 6pp

[7] G.-M. Muntean, P. Perry, L. Murphy, "Objective and Subjective Evaluation of QOAS Video Streaming over Broadband Networks", eTransactions on Network and Service Management, 2005

[8] ITU-T Recommendation J.144 Revised, "Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference".

[9] ITU-R Recommendation BT.1683, "Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference".

[10] Camtasia Studio, web site: http://www.techsmith.com/camtasia.asp

[11] Apple Darwin Streaming Server, web site: http://developer.apple.com/opensource/server/streaming/index.html

[12] QuickTime Pro 7, web site: http://www.apple.com/quicktime/win.html

[13] VideoLan Streaming Server and Player, web site: http://www.videolan.org/vlc/

[14] Microsoft Windows NetMeeting, web site: http://www.microsoft.com/windows/netmeeting/

[15] Video Conferencing Tool, web site: http://www-nrg.ee.lbl.gov/vic/

[16] 3GPP TS 26.234 v6.6.0, "Streaming Service (PSS): Protocols and codecs", Release 6, 2005-12

[17] 3GPP, TS 25.306 v6.8.0 *"UE Radio Access Capabilities"*, Release 6, 2004-12

[18] F. Bernardo, N. Vucevic, L. Budzisz, A. Umbert, "A Beyond 3G Real-Time Testbed for an all-IP Heterogeneous Network", to be published in the 5th ACM International Workshop MobiWAC 2007, October 2007

[19] R. Ljung Ed., "Target Scenarios specification: vision at project stage 1", Public Deliverable D05, AROMA IST-4-027567. Available: http://www.aroma-ist.upc.edu/