User-aware Adaptive Applications for Enhanced Multimedia Quality in Heterogeneous Networking Environments

Pedro M. Ruiz[1], Juan Botia[1], and Antonio F. Gomez-Skarmeta[1]

University of Murcia, Facultad de Informatica, Dept. of Information and
Communication Engineering, Campus de Espinardo,
E-30071 Espinardo (Murcia), Spain
{pedrom,juanbot,skarmeta}@dif.um.es

Abstract. A large amount of research is currently focusing on the issue of using real-time adaptive applications being able to self-adapt their settings, in future heterogeneous mobile networks. Traditional adaptive multimedia applications have neglected to consider the users' subjective perception of multimedia content when changing their configuration. We present a novel approach, which we call user-aware adaptive applications, in which self-adaptation decisions are based on the instantaneous assessment of the users' perceptual quality of the multimedia content. We use the SLIPPER algorithm to produce a rule-based model of the user's perceptual QoS, which is then used by the applications. Our experiments show that the proposed approach provides a good user-perceived QoS even in extremely variable network conditions.

1 Introduction

Adaptive applications and services are a key asset for application and service developers to fulfill the new requirements of future heterogeneous networking environments, which are envisioned as consisting of a variety of heterogeneous wireless technologies attached to a high-speed core IP-based network. The dynamic selection of the settings used by multimedia applications during a multimedia session can help at providing a richer user experience and a higher user enjoyment. For instance, a negligible reduction on the coding quality or the frame rate can avoid the perceptual discontent that packet losses produce to the user. We believe that the assessment of the effectiveness of a mobile multimedia session would be hindered if it did not include the quality with which the users perceive the multimedia content. This view is also supported by the definition of Quality of Service (QoS) given by the ITU-T in the recommendation ITU-E.800 [13].

A large body of literature has been devoted to the use of adaptive applications both in wired and wireless networks. Most of it is to do with the dynamic adaptation of the applications' data rates to those supported by the underlying network. For instance, Bolot[1] and Sisalem[2] use adaptive applications to avoid congestion in fixed networks. However, provided that in wireless and mobile networks packet losses are not only due to congestion, new proposals ([3], [4], [5],

[6], [7]) came up particularly addressing wireless and mobile networks. However, these papers have usually limited their studies to a reduced number of application settings, without performing a broad analysis of multidimensional adaptations (i.e. those considering the tuning of many settings such us frame rate, video sizes, audio codecs, etc. simultaneously). Moreover, they have almost completely neglected the subjective user-perceived QoS when designing the adaptation of the application-layer data-rates to those instantaneously supported by the network.

We introduce a new category of adaptive applications, which we call user-aware adaptive applications, in which the adaptation logic is driven by the user-perceived QoS. That is, given concrete network conditions, the application will select when needed the a new combination of settings (e.g. frame rate, video size, etc.) producing the highest user satisfaction. This approach goes beyond existing works by means of not using low-level QoS metrics such as bandwidth estimations, and packet losses, which do not convey information about the real quality of perception or the user satisfaction regarding the multimedia content. In order to support this approach adaptive multimedia applications need to be challenged with an enhanced adaptation middleware being able to to assess the instantaneous user-perceived QoS. Key to this is the ability for the adaptation function to model the subjective components involved in the human perception of multimedia content.

Little work has been reported on the modelling of satisfaction and perception of the user for multimedia content mainly due to the difficulty of modelling the subjective components. Previous work has focused on the evaluation of the individual impact of some of the application settings in the overall quality perception. Accordingly, Apteker[8] showed a non-linear relation between varying frame rates and the user-perceived QoS. Wijesekera[9] analysed the the effect that random packet losses have on the perceived quality, illustrating the high differences in aceptable loss rates for different media. Further work has been undertaken by Steinmetz[10] who explored the bounds up to which lip synchronisation does not produce annoyance. Finally, Ghinea[11] evaluated the impact of cognitive styles on the QoS perception. However, these works have not successfully produced a model for the user-perceived QoS. We have used the SLIPPER[12] rule induction algorithm to generate a set of rules which represents the quality that the user perceives. We show how incorporating the obtained rules into the adaptation mechanism used by the user-aware multimedia application delivers a higher multimedia quality.

The reminder of the paper is organized as follows: section 2 explains the process of modeling of the user perception using the rule induction algorithm. Section 3 describes the architecture for user-aware adaptive applications we have developed. Finally, section 4 shows some empirical results derived from the use of our adaptive application approach over an ad hoc network, which is one of the most challenging scenarios regarding the variability of the network conditions.

2 Modeling User-Perceived QoS by rule induction

There are many combinations of settings which can result in similar data rates, but very different user-perceptual QoS even when the network conditions are optimal. For instance, a user could prefer a bigger video transmitted at 6 fps and encoded using less quantization than an smaller one at 12 fps using a high quantization even when the latter requires more bandwidth.

Taking advantage of the perceptual preferences when adapting multimedia applications requires the characterization of the subjective components which define a user perception of quality and its relation with the different parameters which define the behaviour of the real-time multimedia application (e.g. codecs, rates, etc.). Given the difficulty to model the user-perceived QoS analytically, we have used the SLIPPER[12] rule induction algorithm to generate a set of rules which models the user-perceived QoS. These rules have been extracted applying the algorithm to a large number of learning examples which have been evaluated and scored by real users. SLIPPER was mainly selected because it is well-known to outperform other rule extraction algorithms.

The SLIPPER algorithm takes as an input a set of tuples with evaluations that the users have made about a concrete combination of settings in a concrete network scenario, and generates a set of rules which models the QoS perception of the user. Each tuple contains the network bandwidth, the audio and video codecs, the quantization of the video codec, the frame rate used, the video size, the loss-rate and the user's score in the interval $[1, 5]$ (as recommended by ITU recomendation ITU-P.800 [15]). In order to produce all this input tuples, we have setup a scenario using our adaptive video conference application, over a link emulator which emulates different network conditions. A total of 864 different combinations of applications settings (in the ranges shown in Table 1) have been tested, and scored by users.

Parameter	Values	Explanation
BW	33, 56, 88, 128, 384	Limit of network bandwidth
LOSS	0..100	% loss packets
AUDCOD	PCM, G711-u, G722, GSM	Audio codec
VIDCOD	MJPEG, H.263	Video codec
FSIZE	CIF, QCIF, 160x128	Size of video frames
QFVIDEO	5, 10, 15, 30, 60	Quantization factor of video codec
FPS	2, 6, 12	Frames by second sent
QoS	1, 2, 3, 4, 5	User perceived quality

Table 1. Parameters changed to generate user's scores

Our goal is to obtain a model of the users' evaluation of QoS, let it be denoted with ξ, representing how they perceive the quality of the multimedia service. SLIPPER produces an approximation $\hat{\xi}$ for that user perception. It can be defined, as shown in equation (1), as a function producing a MOS-like score

taking as inputs the specific network conditions and the particular settings that the multimedia application is using.

$$\hat{\xi} : BW \times AudCod \times VidCod \\ \times Size \times Qual \times FPS \times Loss \to QoS \tag{1}$$

The SLIPPER algorithm makes use of a configuration parameter, noted as $t \in N$, which corresponds to successive refinements of the rule set. The higher the value, the more accurate the model and the bigger its complexity (in number of resulting rules). We applied SLIPPER to the aforementioned data set, using values of $t \in [1, 15]$. In figure 1 we have represented the evolution of the classification error during learning and during testing and the complexity of the rule set in number of rules. It is clear that there is a tradeoff between classification error and complexity.
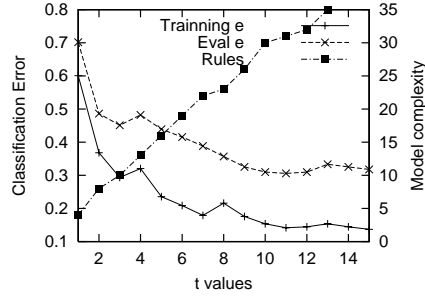

Fig. 1. Model accuracy vs. complexity of the model

From our point of view, the model with $t = 5$, represents an example of a good tradeoff consisting of a total of 12 rules. This configuration produces following rule set, in which the resulting QoS score is obtained as the sum of the output of all the rules whose left-hand antecedent is met:

```
if matchConfidence {
[QFVIDEO >= 60, VIDCOD = MJPEG,
 FSIZE = QCIF, LOSS <= 10, FPS >= 6] -> 2.8792
[AUDCOD = GSM, BW >= 80,
 QFVIDEO >= 30, FSIZE = QCIF, FPS <= 6] -> 1.4357
[AUDCOD = GSM, BW >= 128, LOSS = 0,
 QFVIDEO >= 30, FPS >= 3, VIDCOD = MJPEG] -> 1.7013
[] -> -2.4188
} > 0 then 5 else if matchConfidence {
[BW >= 384, QFVIDEO >= 40, FSIZE <= 2] -> 2.7121
[QFVIDEO >= 30, VIDCOD = MJPEG,
 LOSS <= 3, AUDCOD = G722, BW >=80] -> 1.1756
[FSIZE = CIF, QFVIDEO >= 30,
 LOSS <= 3, AUDCOD = G722, BW >= 80] -> 1.4437
[] -> -1.5044
} > 0 then 4 else if matchConfidence {
[LOSS >= 30] -> 2.1188
```

```
[QFVIDEO <= 5] -> 1.4142
[LOSS >= 16, FPS <= 3] -> 1.5438
[] -> -1.0984207275826066
} > 0 then 1 else if matchConfidence {
[LOSS >= 16] -> 1.9109
[QFVIDEO <= 10, FSIZE = QCIF] -> 1.5861
[FSIZE = 160X128, QFVIDEO <= 40, VIDCOD = H.263] -> 1.2546
[] -> -0.3953
} > 0 then 2 else 3
```

The rule set extracted by SLIPPER allows us to assess how different combinations of settings impact on the user-perception of QoS. In order to get the evaluation of the current settings the application just needs to sum the value of all the consequents of all the rules whose antecedent is matched. For instance, the rules imply that the higher the frame rate the better the quality, but the user prefers changing from a higher frame rate to a lower one (never below 3fps), provided that the video size is increased. Of course, these are just examples of QoS preferences which can be extracted from this rule-set. However, the important point, is that these rules can be used by any multimedia application to assess the user's satisfaction for a particular combination of network and application parameters. In fact, these rules are used by the user-aware adaptation logic presented in the next section.
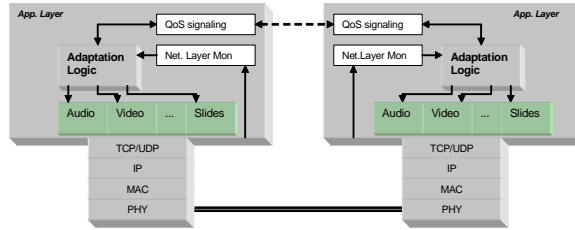
3 Architecture for User-Aware Adaptive Applications

A lot of research has been devoted during the last years to make traditional real-time multimedia applications adaptive. The proposed adaptation mechanisms have neglected the use of perceptual QoS components to drive the adaptation process. In order to support user-aware adaptive applications we add some components to the architecture of a traditional real-time multimedia application. These new components are in charge of the signalling of QoS information, and the provision of the intelligence to keep the user-perceived QoS at an acceptable level. So, the overall architecture, as shown in Fig. 2, consists of a set of multimedia components (e.g. audio, video, etc.), a QoS signalling mechanism and an adaptation logic which is in charge of taking the adaptation decisions. We describe these components in the following subsections.

The QoS signaling mechanism is the protocol in charge of sending and receiving reports describing the network conditions from the other end. When such a report is received, it is passed to the Adaptation Logic so that it can decide which internal settings the application has to use to adapt to the current network conditions while providing a good user-perceived QoS.

3.1 Multimedia Application Components

Real-time multimedia applications usually consist of a number of multimedia components dealing with different types of data such as audio, video, white boards, etc. In our case, we will focus our work on the videoconference service consisting of audio and video components. Particularly, we have extended

Fig. 2. Overall user-aware adaptive applications architecture

ISABEL-Lite (see Fig. 3), a reduced version of the ISABEL [14] application, to dynamically (and in real-time) adapt its behavior to the available resources. The implemented extensions allow the adaptation logic to choose in real time different audio (GSM, G.722, G.711 and PCM) and video (MPEG, H.261, H.263 and MJPEG) codecs, and dynamically change the video frame rate (0..25 fps), the video size and the quantization factor of the video codec. A screenshot of the modified application is shown in Fig. 3.
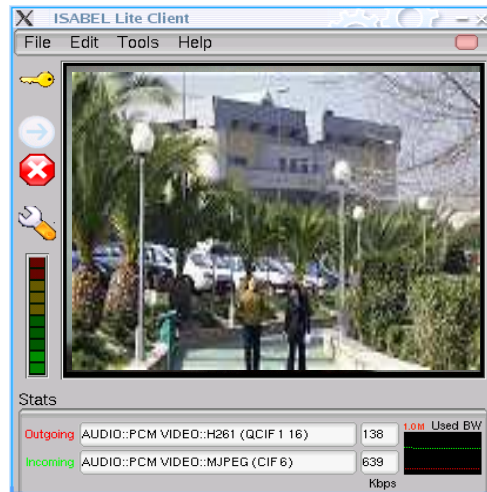
Fig. 3. Screenshot from our ISABEL-lite adaptive application

3.2 QoS Signalling Mechanism

The QoS signaling is a key point of the adaptation architecture as it is the only feedback that the source has from the other end. The received QoS information is used by the adaptation logic to make the decision on which re-configuration of the application settings to perform. As we target heterogeneous networking

scenarios, we use an end-to-end feedback mechanism. The QoS signaling module calculates the loss-rate and mean delay experienced by the data packets in the network. This information is carried in a special signaling packet, called "QoS Report", which is sent back to the source. A sequence number is used to deal with delayed QoS Reports. The rate at which QoS Reports are sent has to represent a good tradeoff between highly dynamic adaptation and a reduced traffic load. At the moment in our implementation it is fixed at one QoS report per second. However, the use of adaptive rates is being investigated to further reduce the signalling overhead.

Although the proposed end-to-end QoS signalling mechanism could have been implemented just using RTCP control packets, the advantage of having a different mechanism is that it allows for a quicker reaction to varying conditions. In addition, the sending of QoS report messages at fixed rates (unlike RTCP) facilitates a quick detection of lost packets as well as problems on the underlying network.

To overcome the loosing of signalling messages one can consider among other alternatives: (i)prioritizing QoS reports or (ii)allowing the applications to detect such losses. As the former approach requires a tight coupling with the network layer, we have found easier for heterogeneous networks to make the receivers send periodic reports towards the sources. In this way, whenever network problems come up, the adaptation logic at the sender can detect missing reports. This information can be used by the adaptation logic to implement some heuristics based on that information (e.g. downgrading the bandwidth consumption when a certain number of QoS Reports is lost).

## 3.3   User-Aware Adaptation Logic

The adaptation logic can be seen as a function which uses the QoS Reports and additional local information to decide which settings need to be configured in each of the different multimedia components. The adaptation logic solves the problem of adapting multimedia flows to the characteristics of the different networks or terminals.

Traditional adaptation logics not considering the user-perceived QoS, are usually driven by fixed thresholds like the loss rate to reduce the data-rates, the number of consecutive 0% QoS reports before trying to increase the data rates, etc. In our case, instead of having fixed thresholds based on network-layer parameters, the user-aware adaptation logic will use QoS perceptual thresholds. Therefore, application reconfigurations will only occur whenever the user-aware adaptation logic detects that the user-perceived QoS is not good enough.

If that happens, the application logic selects from all the feasible configurations of the multimedia components the ones which produces the best user-perceived QoS with those network conditions. Finding such an optimal set of settings requires the adaptation logic to be able to assess the perceptual QoS using the rules shown in section 2.

4 Empirical Results

In order to evaluate the effectiveness of our proposal, we have set up a real wireless ad hoc testbed, on which we have compared the performance of real-time videoconferencing both with traditional applications and with user-aware adaptive applications. The testbed has been deployed in the basement of the CS Faculty at the Univ. of Murcia (see Fig. 4).
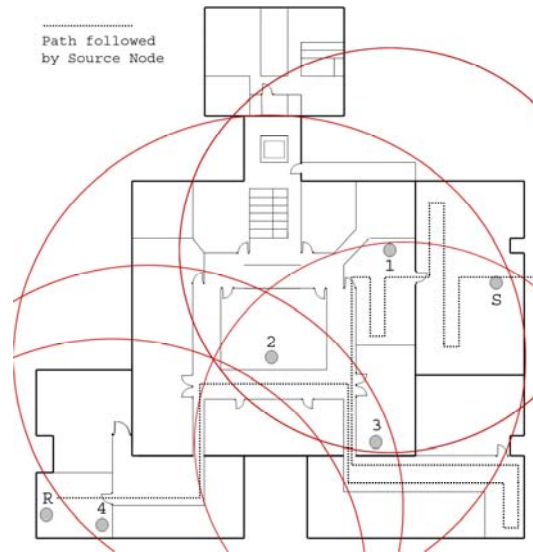

Fig. 4. Map of the testbed scenario

We use the Multicast MANET Ad hoc Routing Protocol (MMARP[16]). The source (S) follows (at walking speed) the path which is shown in Fig. 4 while it runs a videoconference session with node (R). The trials have been performed using the extended ISABEL-lite[14] videoconferencing application described in section 3.

The results which we present are extracted from the RTP traces which are generated by the videoconferencing application. We have used the same route, at the same speed in the same network conditions for the adaptive and non-adaptive trials.The path followed by the node R has been specifically selected so that link breaks and MMARP route changes take place during the videoconferencing session. Furthermore, there were changes in signal strength due to the variation of the distance to MMARP nodes and the number of intermediate walls to traverse. This made the available bandwidth vary during the session.

The results presented in Fig. 5(a) show that the use of adaptive applications is able to reduce the overall packet losses both for audio and video to approximately one third. As expected, the differences are higher in the periods in which there

is less bandwidth available. This is also noticed in the variation of the delays depicted in Fig. 5(b). In the same critical periods, the non-adaptive approach is not able to control the growing of the end-to-end delay, whereas the adaptive one is able to quickly restore the original steady-state.


(a)

(b)
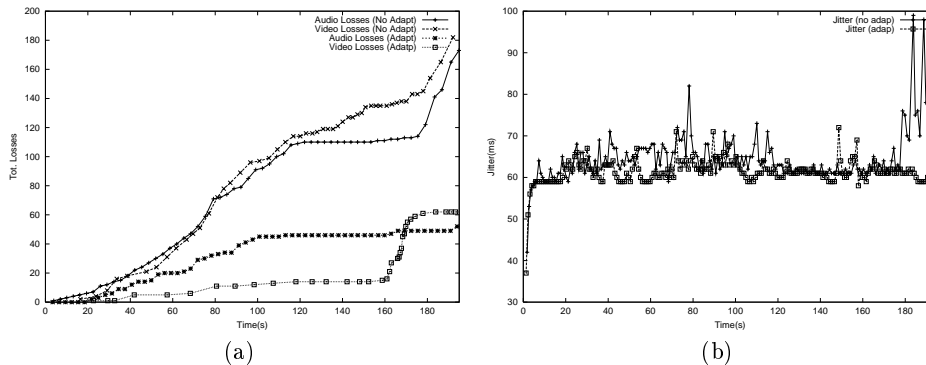
**Fig. 5.** Total losses (a) and audio jitter (b) over time.

As long as the instantaneous loss-rate is the responsible of most of the user's annoyance, in Fig. 6(a), we compare the statistical histogram for the distribution of the audio loss-rate for both approaches. The same statistical analysis is performed for the video flow in Fig. 6(b). For the audio flow, the adaptive application approach is able to keep the loss-rate below 10% all the time. In fact, it keeps the loss-rate below 5% during the 91% of the time. For the video flow, the loss-rate is kept under the 5% the 64% of the time, and its has been under the 10% the 78% of the time.

Although both Fig. 6(a) and Fig. 6(b) demonstrate that the user-perceived QoS is improved when using our novel machine learning driven adaptation, we have conducted an additional evaluation test with real users giving scores of their QoS perception during this trials as well as an overall session score. As we did in the generation of the learning examples (see section 2), we have tried to follow the MOS ITU-recommendation [15] as much a possible. Eight different users have given their QoS evaluations with scores in between 1 and 5. The higher the score the better the quality (i.e., 1=poor, 2=bad, 3=fair, 4=good, 5=excellent). Fig. 7 shows the mean evaluation from these 8 users in different points of the path. The figure also depicts the mean of all the overall QoS perception values. These points are identified by the elapsed time since the starting. These specific moments in time in which we asked for the evaluation were specifically selected in points in which the network conditions were most critical.

As it is depicted in the figure, the user-aware adaptation logic clearly outperforms the traditional multimedia adaptation because the proposed approach is able to maintain a good QoS level even when the network resources are extremely
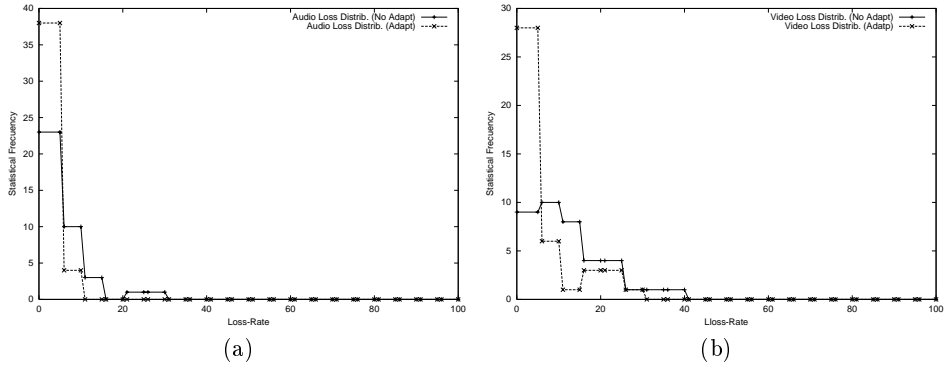
Fig. 6. Histogram for audio loss-rate (a) and video loss-rate (b)
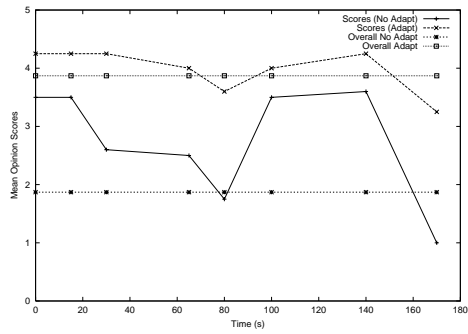
Fig. 7. User scores during the path in the ad hoc testbed

scarce and variable. As can be observed in Fig. 7, the mean overall perception for our adaptive approach is around 3.87 which means that most of the users scored the overall QoS as good. However, the non-adaptive approach has scored in average 1.87 which means that most of the users scored the overall QoS as bad.

5   Conclusions and Future Work

We have proposed a novel machine learning driven adaptation approach for real-time adaptive multimedia applications. This approach, that we call user-aware adaptive applications, differs from other analogous systems as it takes into account the instantaneous user-perceived QoS when taking the decisions about when and how to adapt the multimedia applications to the underlying networking environment. Our proposed user-aware adaptation logic based on intelligent machine learning algorithms has demonstrated to be effective compared to traditional multimedia applications when used in real networks. These results are clearly supported by the evaluations which the users have done on the overall performance in the wireless ad hoc testbed.

As short-term future work, we are extending the proposed adaptation logic with reinforcement learning modules. The goal is not only to generate a rule-based model of the user, but to be able to automatically enhance it over time according to the user's profile and scores.

Acknowledgements

References

1. J.-C. Bolot and A. Vega-Garcia, *"The case for FEC-Based Error Control for Packet Audio in the Internet"*, ACM Multimedia Systems, 1998.
2. D. Sisalem, *"End-to-end Quality of Service Control using Adaptive Applications"*. IFIP International Workshop on Quality of Service, 1997.
3. M. Kazantzidis, S.-J. Lee , M. Gerla, *"Permisible Throughput Network Feedback in AODV MANETSs"*. Proceedings of ICC 2001, Helsinki, Finland, June 2001.
4. T.-W. Chen, M. Gerla, M. Kazantzidis, Y. Romanenko, I. Slain, *"Experiments on QoS Adaptation for Improving Enduser Speech Perception over Multihop Wireless Networks"*. Proceedings of QoS Mini Conference in conjuntion with IEEE ICC'99, Vancouver, Canada, June 1999.
5. A. Alwan, R. Bagrodia, N. Bambos, M. Gerla, L. Kleinrock, J. Short, and J. Villasenor, *"Adaptive Mobile Multimedia Networks"* IEEE Personal Communications, April 1996, pp. 34-51.

6. M. Mirhakkak, N. Schult, and D. Thomsom, *"Dynamic Bandwidth Management and Adaptive Applications for a Variable Bandwidth Wireless Environment"* IEEE JSAC, October 2001, pp. 1985-1997.

7. R. Ramanathan, and R. Hain, *"An Ad Hoc Wireless Testbed for Scalable, Adaptive QoS Support"* IEEE WCNC, November 2000, pp. 998-1002.

8. R.-T. Apteker, J.-A. Fischer, V.-S. Kasimov, and H. Neishlos, *"Video Acceptability and Frame Rate,"* IEEE Multimedia, Vol.2, no.3, pp. 32–40.

9. D. Wijesekera, J. Srivastava, A. Nerode and M. Foresti, *"Experimental evaluation of loss perception in continuous media,"* Multimedia Systems, Vol.7, no.6, pp. 486–499.

10. R. Steinmetz, *"Human Perception of Jitter and Media Synchronisation,"* IEEE JSAC, Vol.14, no.1, pp. 61–72.

11. G. Ghinea, S.-Y. Chen, *"The impact of cognitive styles on perceptual distributed multimedia quality,"*. British Journal of Educational Technology, Vol.34, no.4, pp. 393–406.

12. W.-W. Cohen and Y. Singer, *"A Simple, Fast, and Effective Rule Learner"*. Proceedings of the Conference of American Asociation for Artificial Ingelligence, 1999.

13. ITU-T Recomendation E.800 (0894): *Terms and definitions related to quality of service and network performance including dependability*

14. *The ISABEL CSCW application.*[On line], http://www.agora-2000.com/productos/isabel.html.

15. ITU-T Recomendation P.800 (1996): *Subjective Quality test based on Mean Oppinion Scores (MOS)*

16. P.-M. Ruiz, A. Gomez-Skarmeta, I. Groves, *"The MMARP Protocol for Efficient Support of Standard IP Multicast Communications in Mobile Ad hoc Access Networks"*. Proceedings of the IST Mobile & Wireless Summit. Aveiro, June, 2003. Vol II, pp. 478–482.