Of Packets and People: A User-centered Approach to Quality of Service

A. Bouch , M. A. Sasse, H. DeMeer

Department of Computer Science

University College London

Gower Street, London

WC1E 6BT

UK

{A.Bouch, A.Sasse, H.DeMeer}@cs.ucl.ac.uk

Abstract

Multimedia communication has gained increasing attention, both from the application side and the network provider side. While resource provisioning for QoS support in packet switched networks has lead to the design and development of sophisticated QoS architectures, notably ATM, IntServ or DiffServ, research has not exactly been user or application-context centered. In the cause of the evolution of QoS architectures, the integrated service network approach has lost momentum, and with it, the notion of QoS guarantees. Differentiation of QoS classes within the DiffServ framework is based on the definition of various per-hop behaviors. What is currently missing is a technique for specification and mapping of application and user QoS preferences onto evolving service profiles. In addition, adaptation of applications (and users) is becoming increasingly important in the face of dominating weak QoS-assurance paradigms, both in wireline and wireless environments. As a prerequisite, this paper investigates cognitive and perceptive conditioning of users and applications in a situated setting. The contribution of this paper is twofold: First, essential empirical results on user QoS preferences and QoS graduations are presented, and second, methodological foundations are laid for investigating user-centered QoS.

Keywords

User-centered Quality of Service, task oriented performance.

1. Introduction

Usage of packet networks, especially of the Internet, has increased rapidly over the past few years [1]. Arguably, this increase is not solely driven by the improvement in the network itself, but by a range of new applications and services attracting users, not least within the confines of the World Wide Web. Among these applications are many which are time critical, loss sensitive or throughput dependent. To accommodate these applications with highly variable degrees and categories of Quality of Service (QoS) requirements, a major effort has been invested to overcome the traditional best-effort service quality of the Internet [2,3].

In the wake of early enthusiasm, though, it has been painfully realized that the integrated service model would imply serious drawbacks if applied on a wide scale for data communication. Scalability issues, flexibility concerns and reliability matters soon emerged vigorously among other causes of hindrance. The integrated service model, which is built on the principles of admission control, flow-based resource reservations, the existence of appropriate QoS interfaces and signaling protocols, and techniques for specification and definition of QoS semantics, would provably be able to provide QoS support with guarantees in packet-switched networks. Nevertheless, a paradigm shift occurred due to the very nature of the Internet and its properties, which almost oppose the integrated service model's principles. Immediate and global connectivity, egalitarian packet switching and scheduling, cooperative control mechanisms, self-organizing routing, statelessness, and oblivion to connection properties on the network level were in fact identified as among the main warranties of the success of the Internet. Therefore, a shift towards weak QoS-assurance paradigms occurred based on differentiation of service classes rather than on the granularity of flows [4].

Luckily, the paradigm shift was accompanied by the insight that many of the new multimedia applications based on video, audio and voice communication constituents would in fact require much less stringent QoS assurance than originally anticipated. Streaming video and audio, for example, became very popular within the confines of the World Wide Web, contributing now almost three quarters of overall Internet traffic. Delay and loss became much less critical under these circumstances so that throughput could be identified as the dominating QoS factor governing non-real-time multimedia communications. Playout buffers and retransmissions serve well those types of applications in the given context.

On the other hand, the almost ubiquitous existence of the Internet has considerably shaped the organization of the work of many users. As a result, traditional applications such as email, not considered as time critical, have become regularly embedded into the workflow process so that de-facto delay requirements have emerged. These may be less severe than real-time requirements as for IP telephony, imposing a delay bound of a few hundred milliseconds, but can nevertheless be identified as significant constraints in the end-to-end path. Internet access providers, for example, apply different policies in the time pattern of delivering emails to their subscribers, sometimes without considering the significance of delivery time to users.

Differentiation of QoS classes within the DiffServ framework is based on the definition of various per-hop behaviors (PHB). Expedited (EF) and Assured Forwarding (AF) classes have been defined in RFC2598 [5] and RFC2597 [6], respectively. EF can be used to build virtual leased-line services for real-time applications. Simulation studies underpin evidence that even IP telephony with relatively restrictive delay and jitter requirements can be supported sufficiently well, as for example shown in [7]. The picture is much less clear with respect to AF PHBs. Four levels of forwarding assurance have been defined, each one allowing to choose from three different degrees of dropping preferences. The current double challenge is to build services from these 12 classes of assured forwarding behaviors that both meet applications' and users' QoS requirements and simultaneously are transparent to potential cost models. Charging for services based on AF PHBs may be more difficult to sell than earlier models that focused on the integrated service network model [8,9]. Conceptually, customers' willingness to pay should be inversely proportional to their willingness to adapt. If adaptation is required, it remains to be decided on how to adapt, that is, what information is considered salient in a given context from the user's perspective. Of course, further PHBs are likely to be defined in the future, challenging pricing and mapping over again.

What is currently missing is a technique for specification and mapping of application and user QoS preferences on the evolving, yet to be defined, service profiles. In addition, adaptation of applications (and users) is becoming increasingly important in the face of dominating weak QoS-assurance paradigms, both in wireline and wireless environments. As a prerequisite, this paper tries to fill that niche and investigates cognitive and perceptive conditioning of users and applications in situated and context-sensitive settings. For dynamic application adaptation playing an essential, complementing role in the face of the emerging DiffServ framework, not only application-type specific knowledge is required, but also context and usage sensitive information will be valuable for controlling the adapatation process. Semantic-driven codecs such as MPEG4 may play an important enabling role in that respect [10].

The contribution of this paper is twofold: First, essential empirical results on user QoS preferences and QoS graduations are presented. Second, methodological foundations are laid for investigating user-centered and context-specific QoS (Section 2). The reason for emphasizing differences in methodologies is due to the fact that current approaches for investigating user-centered QoS are largely confined within the concept of mean opinion scores (MOS) [11]. Note, we do not reject MOS-based techniques as such, but regard them as one building block to be used among others, like task orientedness, in an overall framework. In the current paper we elaborate further on earlier results on task-oriented approaches [12,13,14], capturing the requirements of situated users interacting with an application. Due to the limit of space, more details on the whole framework must be given elsewhere. Empirical work, reported in Sections 3 and 4, establishes that users' assessments of the *elasticity* of certain QoS dimensions varies within the same network application for different tasks. This work combines the advantages of both qualitative and quantitative methods to elicit a more complete set of data. In Section 5, we outline the implications of our findings for defining the utility of a delivered QoS. We argue that it is necessary for applications to adapt to changing network conditions in order to maintain valuable levels of service to users.

2 Investigating the value of quality

Studies that focus on users' perceptions of QoS aim to discover the factors that bridge the relationship between objective and subjective QoS. Table 1 presents established minimum QoS required by users, for example applications.

Table 1: Minimum thresholds for acceptable QoS

	Application	*Minimum user requirement*
Inelastic applications	Video	5 frames per sec [15] [13]
	Audio	< 30% packet loss• [16] [13] Latency < 400ms
	Interactive real-time multimedia	Delay < 200ms Jitter < 200ms
Elastic applications	Web page access	Latency < 11 seconds [17][18]

The statistics presented in Table 1 can be used as guidelines for application-specific QoS requirements. However, the context in which applications are used, and users' intentions, make a difference on top of inherent media qualities. Quality can be defined from a lower level, using metrics such as packet loss, to a higher level of granularity. At the highest level, the definition of quality is valid insofar as it carries semantic value to the user. For example, users may conceptualize the quality they receive in terms of attributes of the media stream (e.g. *smoothness*), or attributes of the entire session (e.g. *reliability*). From a user's perspective, quality associated with an application has been shown to be multi-dimensional, variable concept [19]. It has been recognized that what determines quality is dependent upon the purpose of interaction. For example, Virtanen et. al [20] identified continua within which users could describe a particular level of QoS, according to the particular context in which they were operating. Evaluating users' QoS requirements is further complicated by the affect on subjective perceptions of quality created by the interaction between different media. For example, in multi-media conferencing users typically require higher audio quality relative to video quality [16]. However, if the user's task is such that lip-synchronization is required between the audio and video, the quality required from the video stream is 16 frames per second as a minimum [21]. Furthermore, it has been found that lip motion can alter users' voice perception [22].

To effectively relate network state to users' demands for network resources, methods are needed to assess which contextual factors influence users' subjective perceptions of QoS. Table 2 is a summary of user-study methods appropriate for establishing QoS requirements. These methods are explained in more detail below.

Table 2: User-study methods

Method	*Main advantages*	*Evaluation*	*Data gathered*	*Where applicable*
MOS	Standardized scales	Subjective	Quantitative	Experimental conditions. Where controlled levels of QoS used.
Continuous assessment	Immediate reaction to changing levels of QoS	Subjective	Quantitative	Experimental conditions. Systems with variable QoS. Tasks of low cognitive load.
Task performance measures	Direct measure of QoS adequacy.	Objective	Quantitative	Experimental conditions. Specific tasks, with high cognitive load.
Qualitative methods	Grounds data in users' perspective	Subjective	Qualitative	Definition of salient variables User- evaluation of quantitative data

2.1 MOS scores

A method to find out users' opinions on the quality presented to them is the Mean Opinion Score (MOS). This involves users rating the quality in a post-hoc fashion using a 5-point rating scale [11]. As such MOS are a

• Where the packet size is 40ms

subjective method that produces data that can be assessed statistically. Such methods are agreed to be more reliable than objective methods involving instrumental assessment [23].

MOS scales are typically used with small samples of users, in conjunction with controlled experimental conditions and have traditionally been used in speech quality assessment where the assessed speech approaches, telephone or 'toll' quality. Recent work has shown that MOS-based techniques are inherently deficient to reveal all essential properties and qualities of multimedia communications [13]. There are several effects that may bias MOS-based results substantially. For example, research has shown that the intervals on the ITU scale are not conceptually equal in size [24]. Nor is it clear if subjects are being asked to evaluate QoS parameters that are important to them.

It has been shown that overall subjectively perceived speech quality is not necessarily improved alongside an increase in intelligibility [25]. This illustrates the complexity of measuring subjective quality. Indeed, it is not possible for MOS techniques to establish the perspective from which users define and judge the quality [14]. An advantage of using MOS is that a standardized scale on which quantitative data relating to one level of quality can be directly compared to another. By allowing the same subjects to evaluate different levels of quality we can ensure that the same conceptual value is associated with internals of the scale. However, the use of a purely quantitative method, such as MOS, cannot assess these a priori differences between users, nor how QoS received during an earlier part of an interaction can influence judgements later on in that interaction. A level of *task*-oriented analysis, is therefore an essential component of the subjective assessment of QoS to be used in combination with MOS.

2.2 Continuous assessment

Traditional methods for extracting users' opinions about the levels of quality they receive were not designed to apply to packet networks using IP. This is because characteristics of packet-based traffic is bursty in nature, particularly for compressed multimedia. From a users' point of view this means that the quality received can be extremely variable over time.

Methods that ask users to provide a rating of QoS after an interaction cannot assess to what part of that interaction they are referring when making the judgement. Furthermore, it has been shown that users memory for more recent stimuli increased when the duration of the stimulus was increased from 10 to 30 seconds [24]. In order to provide a valid representation of users' experiences with the media they evaluate a method is needed to capture users' opinions in a dynamic fashion, in response to dynamically changeable conditions.

Continuous rating scales have been developed to allow users to assess both audio and video in video-conferencing applications [26][27]. SSCQE (Single Stimulus Continuous Quality Evaluation), was designed to incorporate a 5-point labeled scale, implemented as hardware [26]. Although users' performance of interactive tasks was compromised by the cognitive load introduced by having to manipulate the slider, it was shown that there was a difference between ratings given for quality that subjects gave using the SSCQE method and those given at the end of the session.

Other research has used an unlabelled scale where subjects are required to move a virtual slider-bar corresponding to a 0-100 scale, to indicate their opinion about the QoS they receive. Using this slider it has been shown that users will accept lower quality as long as it remains stable and that their expectancies of QoS influence their requirements [28]. Without the use of continuous assessment methods it would not be possible to show how users' QoS requests change dynamically throughout a single interaction.

2.3 Task-oriented performance measures (tpms)

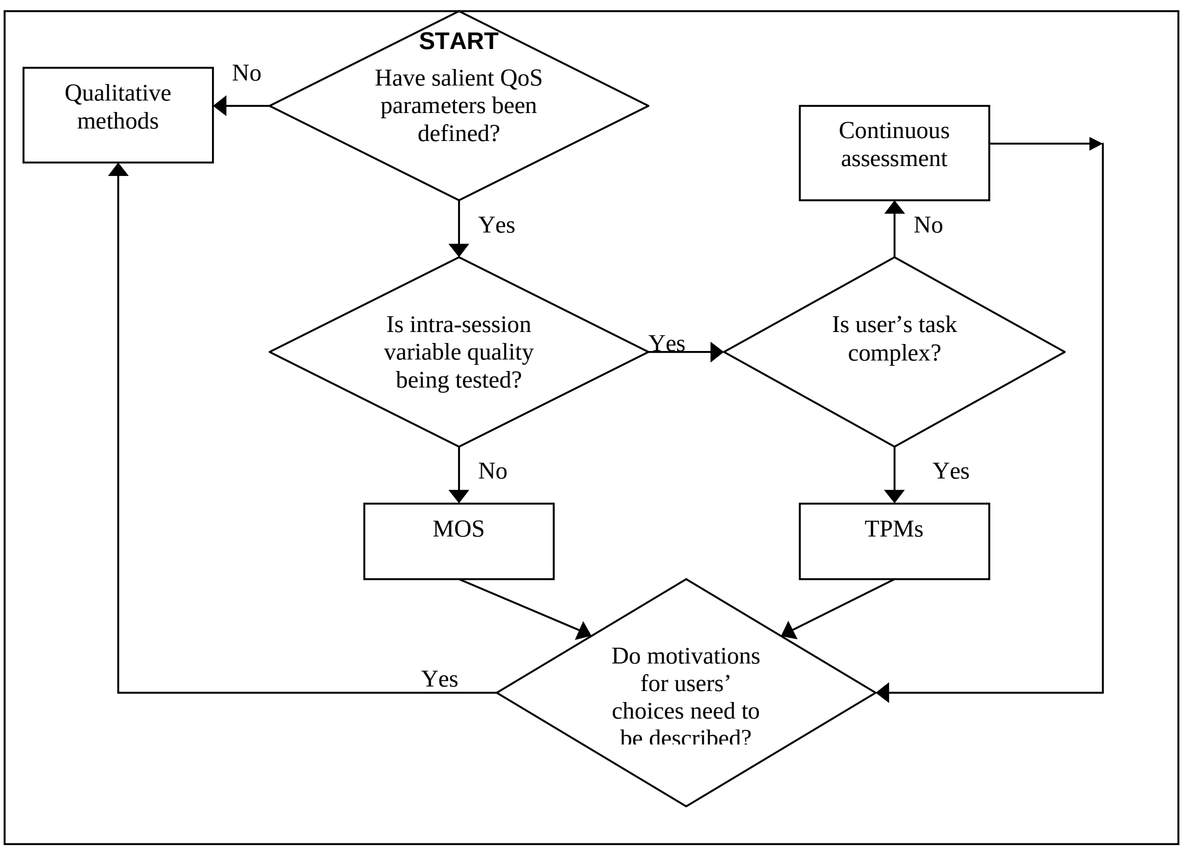
Task-oriented performance measures involve exposing subjects to varying levels of quality under experimental conditions. Their ability to complete the task set is then objectively measured. The advantage of using this method is that the measured performance is relevant to the application that supports the particular task [14]. Tpms contrast with other forms of user-study method as they involve objective evaluation (see Table 2). This means that the element of self-reflection that is necessary in asking subjects to make subjective evaluations is eliminated - the subject doesn't have to be conscious of all the affects of QoS for the adequacy of that QoS to be assessed•. This means that tpms are applicable where subjects are asked to perform a complex task.

This method was used to show that, in a consonant identification task, task performance decreases drastically at an audio-visual skews of 160ms [14]. As the measurements taken by this method are defined according to the particular task it is most applicable where that task is not generalizable outside that task context. The detection of skew is a good example of this. To conduct experiments that use tpms effective researchers need to be clear that experiments are testing QoS parameters that are considered salient to users. Qualitative methods can be used for this purpose.

2.4 Qualitative methods

It is difficult to speculate on individual stress factors that may arise in particular situations if media distortions should arise. Different tools are necessary to find out more on these issues. Qualitative methods, as they involve asking users to *describe* their experiences of QoS, can provide a richer set of data than quantitative methods such as the use of MOS. The former methods attempt to extract the motivations behind users' decisions made during interaction. They are most applicable in the early stages of a user-study to establish a framework from which experiments can be designed but should also be used in conjunction with experimental data to enable users to evaluate their objective experiences. Research reported in [12] used qualitative methods to develop a set of conceptual models that describe what factors are important to users when they assess network QoS. It was found that users make a Risk Assessment of whether future QoS will conform to their task-specific QoS expectancies. These models were used to predict which variables will have significant statistical effects when placed in experimental context [28].

We are working towards establishing a hierarchical framework of user study methods. Such methods should include the capture of qualitative data and task-oriented performance measurements, besides more well know MOS-driven approaches. Figure 1 is designed to aid researchers in deciding the most appropriate user-study method to use when investigating users' QoS requirements.



• This is similar to the idea of using physiological measures to assess the cost of providing users with variable levels of QoS [29]

Figure 1: User-study methods: decision chart

We conducted empirical work to investigate the QoS dimensions that are salient to users. For completeness, both quantitative and qualitative methods were used in the study. The validity of the study was increased by embedding the questions asked in a number of real-world scenarios. The results therefore provide a basis upon which future user studies can build valid experiments.

3. Empirical Results

3.1 Tools

The survey information was gathered by presenting a questionnaire in HTML form, which was made publicly available over the World Wide Web (WWW). A Common Gateway Interface (CGI) program was written to collect the results in a compressed file as raw data. This data was then piped through another program for presentation in a tab-delimited format that could be easily read by spreadsheet programs.

3.2 Participants

The survey was conducted through a WWW form, accessible to all Internet users. Anyone who had used all network applications covered by the survey was invited to complete it. Users were asked to describe their own level of experience and expertise with these applications. We tried to attract a heterogeneous sample of users of network applications, ranging from casual users to experts. We emailed a request to complete the survey form to several mailing lists at a Telco research lab, academic research groups at several universities, and discussion groups on network service management; we also emailed staff and student lists in 'non-technical' academic departments.

3.3 Survey Structure

To establish a baseline for how the notion of *value* might influence respondents' attitude to quality, users were asked for initial descriptions of quality for each of the 6 applications. If a respondent stated that *throughput* was the most important dimension for email, for example, we can use this to see if the answer changed when the respondent was presented with the task scenarios.

In order to evoke real-world situations in which network applications are used, the survey asked users to consider task scenarios in which they used 6 network applications. Table 3 summarizes the survey. For each application, 2 scenarios with different tasks were presented, and respondents asked to state the most valuable determinant of quality. For each application we tried to make the 2 scenarios semantically different. Respondents were then asked if their answer would be different if they had to use the application for a second – different –task. If they chose a different determinant, they were asked to explain why it had changed. The influence of QoS dimensions on the listed applications is hypothesized in Table 3.

Table 3: Structure of survey questions

Application	*Scenario 1*	*Scenario 2*
Email	Delay-tolerant	Delay-intolerant
File Transfer Protocol	Delay-tolerant	Delay-intolerant
Video	Background	Foreground
Audio	Foreground	Background
World Wide Web	Finding information for leisure	Finding information for work
Audio and Video	Non-interactive	Interactive

Email

In the first of the two scenarios asked respondents to imagine they wanted meet with a colleague in 5 minutes and were using email to tell them. Respondents were asked to contrast this with a situation where the same meeting was scheduled for the following week. This question investigates the influence of urgency on the use of this application, by implicating the function of email as being similar to a telephone. It is hypothesized that ‘speed’ will be an important dimension in the initial scenario, but not in the second.

FTP

The first scenario for this application asked respondents to imagine they had to print out a file, which they were retrieving, using FTP, for a presentation starting in 10 minutes. The second scenario contrasted this situation with one in which the presentation was in 3 weeks’ time. Like email, FTP is considered to be delay-tolerant, application. This question is designed to add value to the speed at which the document can be downloaded by the application.

Video

The first of the two scenarios for video asked respondents to imagine the visual background (‘wallpaper’) set up on their PC was from a video connection to a camera showing a river scene. The contrasting scenario involved watching a real-time security video. This question is designed to examine the role of the content of a video stream in determining the value placed on specific QoS dimensions. It is hypothesized that answers might still specify a delay-dependent QoS dimension as being the most important for these applications. However, it is also probable that the value of the video update will be indicated as different between the two applications suggested.

Audio

In the first scenario respondents were asked to imagine they were conducting an important business conversation using network audio. This scenario was contrasted with using network audio for listening to background music. This question is designed to examine the role of the content of an audio stream in determining the value placed on specific QoS dimensions.

WWW

For WWW applications respondents were initially asked to imagine that they were browsing the Internet for leisure interest. They were asked to contrast this with searching for an important document. Arguably, *browsing* the Internet carries different connotations to searching the Internet. The element of speed required in searching may be because, in the latter circumstance, the task involves the access to valuable information.

Audio and Video

In the first scenario participants were asked to imagine that they were receiving a lecture taking place remotely, and were following it whilst completing other tasks on their computer. Respondents were asked to contrast this scenario with a situation where they were taking part in an interactive tutorial. As the latter scenario demands more concentration than the former it is hypothesized that greater value will be ascribed to time dependent QoS variables in the latter case.

4. Results

211 completed survey forms were collected over a period of 3 months. 146 respondents described themselves as members of the network management or research community, and 59 respondents were from a (non-technical) academic background. 6 respondents did not classify themselves in these terms. Data gathered from the initial question regarding the most important QoS dimensions for each application (Table 4) was analyzed separately from responses to the two scenarios (Figures 2-7). In the initial question not all participants gave opinions for all applications. Participants indicated more than one dimension in some instances.

Table 4: Number of responses by QoS dimension

QoS Dimension	*Email*	*FTP*	*Real-time Video*	*Real-time Audio*	*WWW*
Speed	105	105	60	55	87
Reliability	35	35	10	20	23
Ease of Use	70	45	5	4	14
Accessibility	35	40			32
Smoothness		5	35	70	
Security	10	5	25	5	34
Failure Notification	15				19
Compatibility	10		5		
Size			11		
Resolution			27		
Feedback		5			7

For each application, as displayed in Table 4, the most popular QoS dimension was statistically significant. This means that, for example, the probability that no real trend of opinion is reflected by the high number of respondents assessing that *Speed* is the central requirement for email applications is very low ($p<0.05$).

Figures 2-7 contrast the QoS dimensions respondents selected for first and second scenarios for each application, effectively presenting 2 QoS profiles for the same application used in different contexts. In each figure, only dimensions that were chosen by respondents are presented.

Figures 2 and 3 clearly show that the particular QoS profile required for email and audio depends on the task for which it is used. For both applications respondents selected *speed* as the most valuable characteristic in the first task (*Scenario 1)*. Figures 2 and 3 also indicate that the variation within QoS profiles is more pronounced than the variation between two applications seen as having opposite QoS requirements.

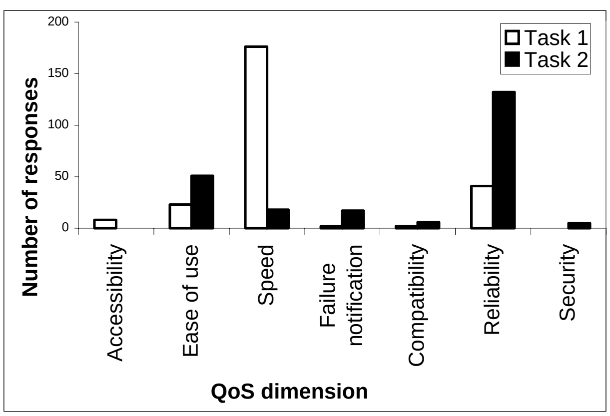


Figure 2: QoS Profile: Email

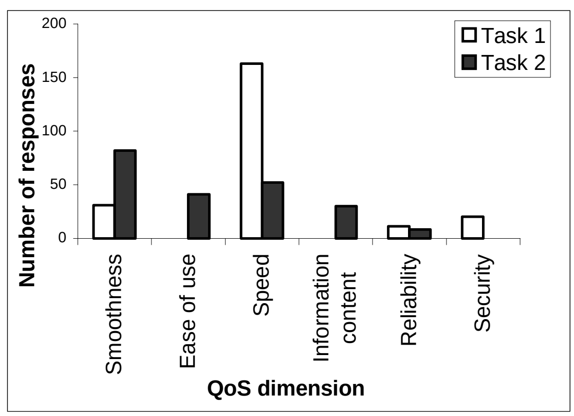


Figure 3: QoS Profile: Audio

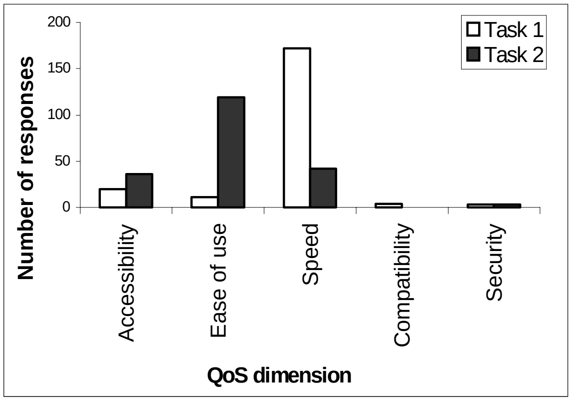


Figure 4: QoS profile: FTP

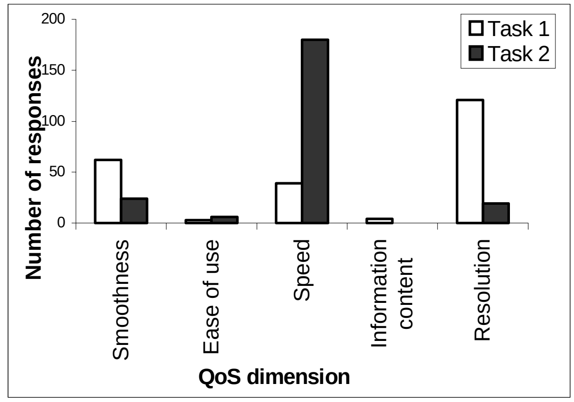


Figure 5: QoS Profile: Video

Figures 4 and 5 tell a similar story for two other contrasting applications. The profile of *Scenario 1* video reflects a similar - relatively delay-tolerant – assessment as *Scenario 2* for FTP. In the contrasting scenarios, both applications were given profiles that indicate that they are not delay-tolerant. Figure 6 shows that profiles for multimedia applications vary dramatically between the 2 different scenarios.

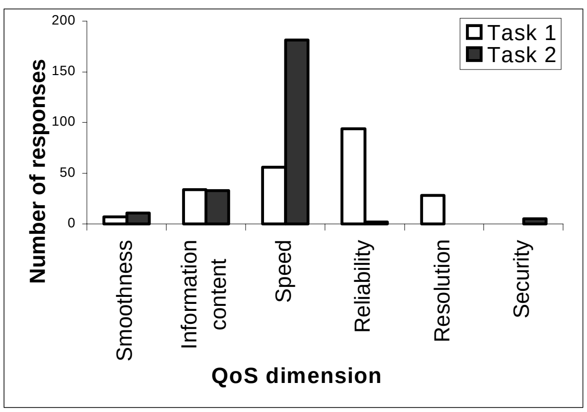


Figure 6: QoS profile: Audio and Video

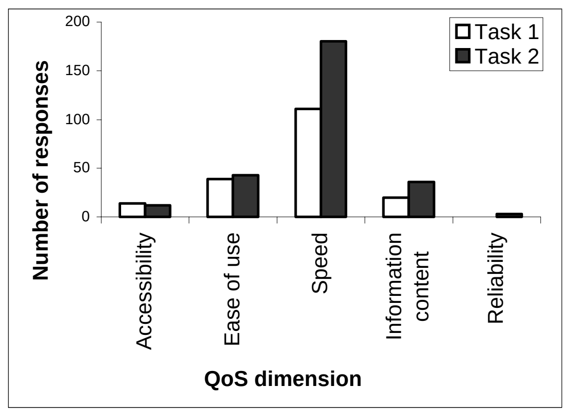


Figure 7: QoS profile: WWW

An additional aim of the survey was to determine whether different user-groups (expert and casual users of networked applications) have different QoS requirements. The survey found no significant differences between responses of these two user groups.

4. Discussion

Our results show that users' requirements for QoS vary within the context of their interaction. We have defined the affective context as including several factors. We identified the users' goal of interaction as a strong contextual variable influencing their QoS requirements. Our results show that the salience of particular QoS dimensions vary with the task performed. This means that users' definitions of what QoS parameters are valued depends on the goal of their interaction, and that QoS is not determined on the type of application alone. The use of quantitative methods has enabled us to distill responses into statistical format. In this description the subjective value of QoS parameters are in direct proportion to each other. Coupling this method with qualitative data we can prove that the motivations behind these choices are truly task-driven, in the manner hypothesized (see Table 3). For example, from the qualitative data we captured a respondent stated, when referring to real-time video:

'If I'm a security guard, I'm going to need an immediate update on my screen, otherwise I can't do my job. With the background scene…it's not the same urgency is it?'

For FTP:

If I need something in 10 minutes then it's clearly the speed of the download that I'm concerned with, that I want. If it's three weeks away, I suppose I can wait a week!, maybe it's more important then that it's complete,…'

Both of these quotes from respondents illustrate that fact that the type and level of quality delivered to the user must be configured, not according to a *description* of actions, but in accordance with the goal of that interaction. Our results also show that users may not associate applications with their traditional QoS dimensions, even where the context in which the task was performed was not manipulated via contrasting scenarios. In the initial questions, a significant number of respondents chose *speed* as the most salient QoS dimension associated with email. Thus, it is not true that email can always be delayed when the network is congested. Users may need the ability to distinguish between, and pay differently, for urgent emails and truly delay-tolerant ones.

The current Internet works on the premise that network capacity is shared proportionately between users. This assumes that the utility function that drives QoS allocation is uniform among users. Our research has attempted to show that the definition of the utility of a service is flexible according to users' subjective perceptions. Given this degree of variability it makes sense to allow end-systems to specify QoS requirements in order the users' network usage can be optimized. For example, a congestion event detected on the audio stream of a multi-media conferencing application could be passed to the video stream. Previous work has shown that allowing users themselves to dynamically configure their QoS requirements not only increases user satisfaction, but optimizes system usage [8]. Our results have specified the relative utility of such QoS parameters in situated settings.

The variable nature of network congestion means that, potentially, the number of decisions needed concerning how to react to congestion events would mean an unacceptable cognitive load on users. To provide a service that is both valuable and implementable there is therefore a need for compromise between the abstracting the control of QoS optimization decisions away from users while maintaining flexibility in their representation. The appropriate compromise necessarily involves tradeoffs between the ability of the network to allocate resources efficiently, and the provision of network-level parameters to which users can relate. In this way the interface between the network and the notion of *value* ascribed to an application is not direct, but lies in the description of a set of QoS parameters that form the chosen profile.

Solutions developed in the area of application adaptation may meet the criteria of abstracting resource allocation requirement details away from the network, while allowing the appropriate level of flexibility when allocating resources to users. Reference [30], for example, presents a scheme using layered transmission. At times when network congestion occurs the most *valuable* QoS parameter can be maintained. Our results provide specifications that can aid network designers in understanding which part of a QoS profile can be restricted in which circumstances. The profile representing how an application should adapt to changing conditions is specified by the user and potentially maintained by the user-agent software [31].

The most efficient implementation of resource allocation schemes is the focus of ongoing research. The results reported in this paper show a central consideration for future network QoS management must be an understanding of the utility of QoS from a users' perspective. Further research is needed to assess how the needs (and pockets) of users dictate traffic generation. We have attempted to provide a detailed account of the methodology that can be used to investigate users' subjective QoS requirements. By using these methods we have shown that the utility of a received service is determined by subjective QoS criteria. It is, arguably, only this utility, ascribed to the QoS received, that has true potential to create revenue.

5. Conclusions

The aim of this paper was to show how different methods can be used to elicit users' cognitive and perceptual processes with regard to network QoS. We were motivated in this aim by the need to provide a specification of QoS requirements that could be mapped to evolving service profiles. We have presented results from a number of studies that demonstrate the effectiveness of such methods. The results of our investigations show that the applicability of QoS dimensions varies according to the subjective value ascribed by the user and depends on the purpose of the task. Capturing qualitative data enabled us to understand the motivations behind users' preferences for certain QoS dimensions, and to what extent their choices can be generalized. Further work is now needed to examine users' responses to different levels of QoS under experimental conditions, for the dimensions that this work has specified. This work has addressed issues that are becoming increasingly critical to the optimization of network QoS. Although many implementation decisions will be based on predictions about the nature of technological progress, establishing a framework from which users' needs can be integrated into resource allocation schemes can help to manage current and future demand for QoS. Specifying of the purpose of users' tasks has potentially crucial implications for providing a degree of service that will encourage users to utilize, and pay for, the quality that they receive.

References

[1] Coffman, K.G., & Odlyzko, A.M. (1998). The size and growth rate of the Internet. Available at: http://www.research.att.com/~amo

[2] Gupta, A., Stahl, D.O., & Whinston, A. B. (1995). Priority pricing of integrated services networks. In McKnight, L.W., & Bailey, J.P. (eds). Internet Economics. MIT Press.

[3] Clark, D., Shenker, S., & Zhang, L. (1992). Supporting real-time applications in an integrated services packet network: Architecture and mechanism. In Proceedings of SIGCOMM'92 pp.14-26.

[4] Nichols., K., & Blake, S. (1998). Differentiated services operational model and definitions. Available at http://www.ietf.org/internet-drafts/draft-nichols-dsopdef-00.txt

[5] Jacobson, V., Nichols, K., & Poduri, K. (1999). An expedited forwarding PHB. RFC 2598.

[6] Heinanen, J., Baker, F., Weiss, W., & Wrowclawski, J. (1999). Assured forwarding PHB group. RFC 2597.

[7] Anurag, A., Muppala, J.K., & De Meer, H. G. (2000). VoIP support on Differentiated Services using Expedited Forwarding. Proc. 19th IEEE International Performance, Computing, and Communications Conference (IPCCC-2000), Phoneix, Arizona, February 2000.

[8] Mackie-Mason, J.K & Varian, H. (1997). Economic FAQs about the Internet. In McKnight, L.W., & Bailey, J.P. (eds). Internet Economics. MIT Press.

[9] Breslau, L., Shenker, S. (1998). A simple comparative analysis. Computer Communication Review, 28, (4), October, 1998.

[10] Requirements Group. (1997). MPEG-4 Overview. ISO/IEC JTC1/SC29/WG11 N1730, Stockholm, July 1997.

[11] ITU-T P.800 Methods for subjective determination of transmission quality.

[12] Bouch, A., & Sasse, M. A. (1999). It ain't what you charge, it's the way that you do it: A user perspective of network QoS and pricing. Proc. IFIP/IEEE International Symposium on Integrated Network Management (IM'99), Boston 24-28 May, 1999,pp. 639-654.

[13] Watson, A., & Sasse, M.A. (1996). Evaluating audio and video quality in low-cost multimedia conferencing systems. Interacting with Computers, 8, (3), pp. 255-275.

[14] Knoche, H., DeMeer, H., & Kirsh, D. Utility Curves: Mean opinion scores considered biased. 7th IEEE/IFIP International Workshop on Quality of Service (IWQoS'99), London, UK, June, 1999.

[15] Frowein, H.W, Smoorenburg, G.F., Pyters, L. & Schinkel, D. (1991): Improved speech recognition through videotelephony: experiments with the hard of hearing. IEEE Journal on Selected Areas in Communication, 9, pp. 611- 616.

[16] Sasse, M.A., Biltung, U., Schulz, C.D., & Turletti, T. (1994). Remote seminars through multimedia conferencing: experiences from the MICE project. Proc. of INET '94/JENC5.

[17] Nielson, J. (1994). Usability engineering. AP Professional Press, Boston, MA.

[18] Bouch, A., Bhatti, N., & Kuchinsky, A.J. (2000). Quality is in the eye of the beholder: Meeting users' requirements for Internet Quality of Service. To be presented at CHI'2000. The Hague, The Netherlands, April 1-6, 2000.

[19] Fluckiger, F. (1995). Understanding Networked Multimedia. Prentice Hall.

[20] Virtanen, M.T., Gleiss, N., & Goldstein. M. (1995). On the use of evaluative category scales in telecommunications. Proc. of Human Factors in Telecommunications 16, 2, pp.16-31.

[21] Kouvelas, I., Hardman, V., & Watson, A. (1996). Lip Synchronisation for use over the Internet: Analysis and Implementation. Proc. of IEEE Globecom '96, November 1996, London UK.

[22] McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. Nature, 264 (5588).

[23] Flanagan, J.L. (1965). Speech Analysis, Synthesis and Perception. Springer-Verlag, Berlin.

[24] Aldrige, R., Davidoff, J., Ghanbari, M., Hands, D., & Pearson, D. (1995). Measurement of scene-dependent quality variations in digitally coded television picture. Proceedings – Vision, Image and Signal Processing 142, 3.

[25] Podolsky, M., Romer, C., & McCanne, S. (1998). Simulation of FEC-based error control for packet audio on the Internet. Proc. INFOCOM '98.

[26] ACTS TAPESTRIES. (1997). Acceptability studies in selected areas of audio-visual communications. ACTS Project AC055, Deliverable R/003/b2.

[27] Bouch, A., Watson, A., & Sasse, M.A. (1998). QUASS - A Tool for Measuring the Subjective Quality of Real-Time Multimedia Audio and Video. Poster presented at HCI '98, 1-4 September 1998, Sheffield, England.

[28] Bouch, A., & Sasse, M.A. (2000). The case for predictable network service. Proc. MMCN'2000. San Jose, CA, January 24-27, 2000.

[29] Wilson. G. & Sasse, M. A. (1999) Proc. of the International Workshop on Interactions, October 21st-22nd 1999, Siena, Italy.

[30] Hoffman, D., & Speer, M. (1996). Hierarchical video distribution over internet-style networks. In ICIP'96. Lausanne, Switzerland. September, 1996.

[31] Nahrstedt, K., & Smith,J.M. (1995). The QoS Broker. IEEE Multimedia, 1995, 2 (1), pp.53-67.