Measuring Quality of Service in an Experimental Wireless Data Network

M. Ivanovich, P. Fitzpatrick, J. Li, M. Beresford, A. Saliba

Telstra Research Laboratories
Clayton, Vic 3168
{milosh.ivanovich, paul.g.fitzpatrick, jon.li, michael.beresford, anthony.saliba}@team.telstra.com

Abstract – This paper describes the design, implementation and use of a wireless data network test-bed to support a facilitator-led experiment measuring the quality of service (QoS) of selected mobile applications. The test-bed comprised a wireless LAN, wireless-capable Personal Digital Assistants acting as "3G-like" terminals, with a network performance emulator at the IP-layer providing controlled, reproducible mappings between application-specific QoS and relevant network parameters. The mapping process uses a QoS framework with a top-down view of quality, and a layered OSI-based hierarchy with linkage between performance attributes of adjacent layers. Finally, we illustrate a technique to estimate whether a wireless access technology could satisfy customers' QoS expectations.

I. INTRODUCTION

Wireless communications is one of the fastest growing telecommunications sectors, and developments in wireless data are seen as a means for delivery of new services and continuing growth in this sector. Customer satisfaction with wireless networks is driven not only by cost but also by service quality [1]. The expectation is that wireless access must resemble wired access as much as possible and provide a range of services including voice and data [2]. These services have different network requirements, given the diverse functionalities offered by each. However, functionality is not always enough, because the demands of the user often exceed the worst-case performance level of the network.

Today's fixed access users of the Internet application space have certain levels of expectation or demand when it comes to the service quality they receive. It is well known users are increasingly dissatisfied with the 'best effort' quality of service (QoS) provided by the current Internet [3]. Meeting customer expectations is clearly an important goal, but we find that much of the literature is more concerned with performance attributes of specific network layers (e.g. IP- or PPP-layer bandwidth, TCP segment or radio link layer block or queueing priority) than with measures of quality as experienced by users delay (for example see [4] and [5] and the references therein).

While understanding and developing techniques to control access and system performance will ultimately contribute to delivering QoS, evaluating the impact of such techniques on users and their satisfaction with any given service is an area which is not as well understood.

The main goal of this paper is to describe the process of designing, implementing and using a wireless data network test-bed in support of a facilitator-led experiment measuring the quality of service (QoS) of specific applications delivered in a mobile context. The experiment itself revolves around a "wireless QoS framework", which we also briefly outline. Its purpose is the modelling, assessment and evaluation of the impact of wireless network performance on service quality, taking into account the particular application and framework layer under study.

The paper is organised as follows. In Section II we explore the various definitions of QoS in the domain of voice and data services, as well as the feasibility of developing quantitative user satisfaction models. This is followed by a description of the hierarchical framework for wireless QoS, in Section III. Section IV discusses the topology and components of the experimental wireless test-bed, including the applications chosen for study in mobile conditions. In Section V the focus is on using the described QoS framework and wireless test-bed in a practical experiment to predict the extent to which a given wireless access technology can satisfy customers for a specific application. Finally, conclusions are drawn in Section VI.

II. DEFINING AND MEASURING QOS

Quality as a generic performance measure can have many meanings, as discussed in a companion paper [6]. The ETSI definition of QoS in the context of digital networks [7] is quite appropriate: "...the collective effect of service performances which determine the degree of satisfaction of a user of the service". The user most often does not need details about how a service is provided; their primary concern is the resulting quality.

A complicating factor is the individual nature of how users rate the quality that they receive. Any two users who may be sharing a common experience (i.e. identical applications) are likely to have significantly different views of the QoS. The first step in understanding how such individual views about service quality are formed is being able to create a quantitative linkage between wireless network parameters and the user perception of QoS provided over that network. This linkage will typically take the form of a numerical mapping (mathematical relation) between some measure of the user-perceived quality (e.g. the Mean Opinion Score

(MOS), as discussed later) and a particular set of network parameters under study (e.g. available bandwidth), as illustrated in Figure 1.

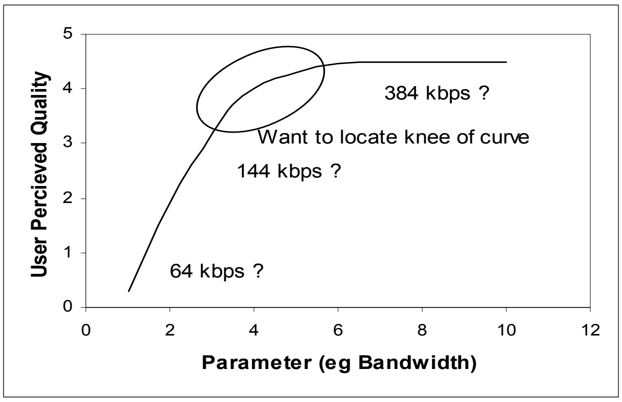

Figure 1 : An Example of a Quantitative Mapping Between Network Parameters and User-Perceived Quality of Service

This simplified figure shows only a single parameter as the independent variable - in this case the available network bandwidth. Usually the mapping becomes significantly complicated by its multi-dimensional nature, as user perception of QoS invariably depends on multiple factors, some examples being the inherent packet transfer delay, delay variance (jitter) and packet loss probability in packet switched networks. The need to control such complexity is the primary driver for restricting the number of network parameters studied in the experiment outlined in Section IV.

The chief motivator for the development of a quantitative mapping such as the sample presented here, is the need to search for the so-called "Knee of Satisfaction" shown in Figure 1. The x-axis position of the knee then identifies the value range of the network parameter(s) being studied, beyond which there are significantly diminishing returns in achieving greater perceived quality of service. A secondary benefit of the mapping is a better understanding of what constitutes the *customer value proposition* in accessing diverse internet applications in wireless environments across a range of service qualities. With reference to Figure 1, this may be interpreted as looking for the y-axis plateau of the QoS curve, rather than just where the "Knee of Satisfaction" occurs. This knowledge is very useful because it can be used to determine the minimum resource allocation to meet user expectations of quality.

In the domain of voice services, MOS is a well known measure for assessing subjective quality. The value of the MOS approach is that it measures the users perception of quality for the conditions of the particular experiment. When dealing with data services, the problem tends to be multi-dimensional primarily because of the many different applications which may be carried simultaneously by the same data network connection. An example of one particular data-based service for which a quantitative model measuring user satisfaction has been developed, is the *E Model* for Voice over IP (VoIP) [8].

In the following section we outline the QoS framework developed in [9] and [10] that provides the methodology for the experiments reported in this paper.

III. THE HIERARCHICAL FRAMEWORK FOR WIRELESS QOS

The facilitator-led experiment conducted on our test-bed network revolves around a "wireless QoS framework" as shown in Figure 2. Full details of the development of this model are in [9] and [10]. The formulation of the framework is based on two steps: (i) at each intermediate OSI layer ***identify*** those parameters which will have an impact on the highest layer (User-perceived) QoS; (ii) at each layer interface, by a combination of experimentation and/or analysis ***construct*** a quantifiable "upward" mapping between the adjacent layers' constituent parameters. By "upward" mapping, we mean that the parameter values of a higher layer depend on, and are caused by those of its adjacent lower neighbour, while the converse is **not** true. As an example, we might be able to say that an Application's slow perceived speed of content delivery is caused by "low" IP layer network bandwidth and "high" packet loss probability, while the converse ("downward") mapping is not true.

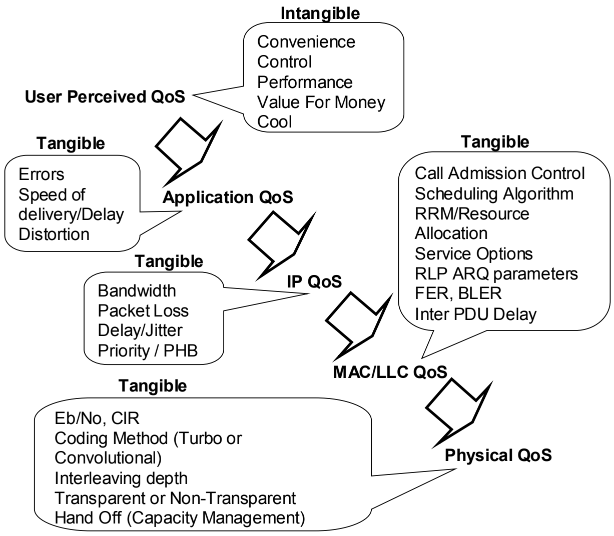

Figure 2 : A hierarchical framework for wireless QoS

Figure 2 also shows that all of the parameters at layers below User-perceived QoS are considered "Tangible" due to the relative ease with which they can be measured, as compared to some of the very subjective and "Intangible" parameters of the (highest) User-perceived QoS layer. *Coolness* and *value for money* are two such "Intangible" parameter examples.

Another important factor in step (ii) is that the mapping is likely to be a non-linear many-to-many mathematical relation. Therefore when constructing the mappings in the course of experimentation or analysis, it may be necessary to limit the underlying complexity by studying only the most prominent parameters at a given layer – a *core subset*, which is deemed to have the most impact on the ultimate (User-perceived) QoS measure (as illustrated in Section V).

IV. TEST-BED COMPONENTS, TOPOLOGY AND SUPPORTED APPLICATIONS

Figure 3 depicts the wireless data network topology which formed the backbone of our test-bed. Detailed information on the experimental method and techniques from a human factors perspective, as well as the participant demographics, can be found in [10].

There are three key architectural components - the server farm subnet, the client devices subnet and the IP layer network performance emulator (a Linux PC running NISTNet, and equipped with at least two network interface cards (NICs)). NISTNet is a general-purpose software tool, running on a routing-capable computer, for manipulating critical IP-layer parameters and therefore network performance levels [11]. Examples are packet loss probability, available bandwidth, packet delay and packet delay variation (jitter).

This topology allowed targeted IP traffic from the server farm to the user terminals to be shaped (bi-directionally) according to the settings on the NISTNet-equipped routing PC. This, in turn, translated into the different QoS levels that the user participants experienced during the facilitator-led trials.

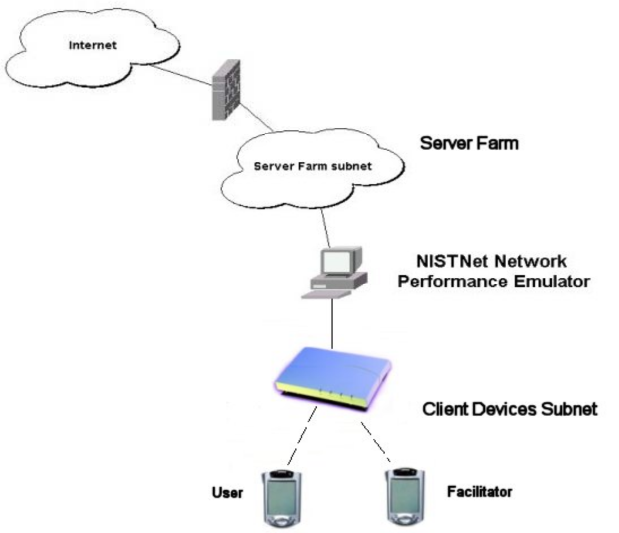

Figure 3: Wireless Data Network Test-bed used in Measuring Application-specific QoS in a Mobile Context

Significantly, NISTNet allows individual IP-flows to be treated with the use of address filters, allowing flexibility to impose certain network performance conditions on some devices (e.g. those of test participants) and not others (e.g. the facilitator's "control" device used for running and timing the experiments). The latest version of NISTNet also introduces advanced multi-protocol features such as support for (i) TCP explicit congestion notification (ECN) and (ii) the class of service (CoS) field in IP packet headers. Such new features present a good opportunity for further research in the field of user perceived wireless QoS emulation and measurement.

The client devices subnet was realised through use of the IEEE 802.11b wireless local area networking (WLAN) standard. The idea was to provide a wireless "bit pipe" large enough to permit the NISTNet settings to remain the only performance bottleneck in the end-to-end path (for experimental control, reliability and reproducibility), while retaining the untethered WLAN property in order to emulate the mobility associated with typical 2.5 / 3G cellular access technologies of interest (e.g. GPRS, 1xRTT, W-CDMA).

At the time the experiment was conducted there was an absence of true 3G devices, and as a result we chose to simulate the use of these by using the most likely evolution of technology over the next few years – essentially a "marriage" between a Personal Digital Assistant (PDA) and a mobile phone: specifically the Compaq iPAQ 3800 series. Although this device did not have an integrated mobile data network capability, this was simulated through the use of a PCMCIA sleeve allowing the iPAQ to accept an 802.11b WLAN card.

The same client device was used in all of the experimental tests, removing it from consideration as a variable in the work which was conducted [10]. An interesting future area of research would be to consider the impact of the client device itself on the user perceived wireless QoS for a range of applications. This will be particularly relevant as a wide range of true-3G device form-factors become commercially available.

The application suite used throughout the QoS-measurement experiment was chosen on the basis of contemporary popularity in today's wireline Internet. A further narrowing down to four applications was performed on the basis of estimated suitability to a mobile 2.5 / 3G environment – the final choice is shown in Table 1. Niche mobile phone-specific applications such as general Wireless Access Protocol (WAP) browsing, Multimedia Message Service (MMS) and m-commerce (via SMS or WAP delivery mechanisms) are not in this list because it was felt that the functionality they encompass is adequately captured by web surfing and E-mail retrieval [10].

Application	***Suggested QoS Class*** *(ETSI spec.,[12], [13])*	***Transport Protocol***
Web Surfing (HTTP)	Interactive	TCP
E-mail Retrieval (IMAP)	Background	TCP
Large File Download (FTP)	Background	TCP
Video Streaming (RTP)	Streaming	UDP

Table 1: Applications Used and Suggested QoS Classes

The programs that were used for each of the above applications were Pocket Internet Explorer (web surfing), Pocket Outlook (E-mail), Rucksun Scotty FTP, and Pocket Windows Media Player 8.0 (video streaming). As can be seen from Table 1, three of the four applications use the connection-oriented TCP protocol, while the connectionless User Datagram Protocol (UDP) is used for the most delay sensitive video streaming application. Standard protocols such as Hyper-Text Transfer Protocol (HTTP), File Transfer Protocol (FTP), Internet Message Access Protocol (IMAP), and Real Time Transport Protocol (RTP), are used at the application layer. The second column in Table 1 suggests for each application a classification using the four standard ETSI traffic classes, also known as QoS classes [12]. 1. ***Conversational*** - real time traffic flows, greatest delay sensitivity, e.g. voice or video telephony. 2. ***Streaming*** - real time traffic flows, medium delay sensitivity, e.g. one-way streaming media. 3. ***Interactive*** - used for interactive but delay tolerant traffic flows which require smaller data error rates, e.g. web browsing or chat. 4. ***Background*** - used for non-urgent, delay tolerant traffic flows that require smaller data error rates, e.g. large file download or email retrieval.

The mix of chosen application types is well spread across the traffic classes. This ensures that our testing scenarios model a diverse range of future mobile data customer uses, in terms of the likely user experience and quality expectations. The fact that the *Conversational* traffic class is not covered is mitigated by the substantial body of existing research assessing customer perceptions of quality for this type of traffic (see [8] for example, with regards VoIP quality in traditional wireline environments).

V. A Practical Experiment : Answering Useful Questions

In this section we describe the use of the presented QoS framework and wireless test-bed in support of a practical experiment. We focus on one of the core experimental aims - predicting the extent to which a given wireless access technology can satisfy customers for a specific application. The other objectives of this experiment, as well as a detailed treatment of results and their interpretations are not the topic of this paper and for this the reader is referred to [6], [10].

The experiment shapes an *empirical quantitative user satisfaction model* by defining and measuring mappings of user satisfaction versus wireless network parameters at the top three levels of the framework shown in Figure 2: the User-perceived, Application and IP QoS layers. When conducting an experiment of this nature, limiting complexity of the inter-layer mappings becomes paramount. It was decided that trying to understand the QoS implications of all inter-layer mappings in the hierarchy simultaneously would have made the task unrealistic, and overshadowed the key relationships that we hoped to illustrate. As a result, we simplified the process by: 1. performing the mapping directly between the User-perceived and IP QoS layers, 2. taking into account the impact of the Application QoS layer by separately studying 1. on a per-application basis and 3. reducing the number of IP layer parameters to a *core subset*, deemed to have the most impact on the ultimate (User-perceived) QoS measure.

With regards to 3., the core set was made up of (i) the available IP-layer bandwidth and (ii) IP packet loss percentage. With one exception, the application suite used TCP as the transport protocol (see Section IV), and these two parameters were selected on the basis that they had the most significant impact on TCP and hence application performance - more detail can be found in [10]. To a certain extent TCP can cope well with a range of packet delays, although not with extreme short-term variations in the same. The applications which utilise UDP on the other hand, can generally tolerate moderate to large packet losses as well as changes in bandwidth, but are very sensitive to packet delay and jitter, much more so than applications which are carried by TCP. In this we see a pattern arising – IP is the network layer common to both TCP and UDP, and it appears that IP-layer delay as well as delay variation in the form of jitter becomes an important future extension to our work, particularly due to the wireless nature of the environment.

The skeletal structure of the user satisfaction model for this experiment was formed when, based on the consumer behaviour literature discussed in [6], we found that four main factors influenced the user perception of QoS given a particular network condition. These were *reliability*, *efficiency*, *predictability* and *satisfaction*, forming the core set of measurable parameters at the User-perceived QoS layer (see Figure 2).

Reliability examines how important and useful it is to know in advance the level of network performance. *Efficiency* represents a measure of how quickly the system responded to requests and may be interpreted as the perceived speed of network delivery. *Predictability* is concerned with the degree to which the user experience followed the expectations of the user. Finally, *satisfaction* measured the overall "rating" assigned by the user to each experience, taking everything into account.

The purpose of the experiment was to flesh out the user satisfaction model by assessing each of these parameters with a question in an application-specific and network condition-specific context. Some parameters were not relevant to each application, and *reliability* was not assessable experimentally (examined in a pre-study questionnaire). Although questions were application-specific, all were assessed on the same 5-point scale [6]. The application-specific nature of the questions and their impact on core parameters shows the effect of the Application layer being accurately recorded in our experiment, without the more complex alternative of individually defining and mapping all intermediate parameters.

The practical result was a series of points, mapping the **two** controlled network parameters (IP-layer bandwidth, loss) to a **single** user-perceived parameter (one of either *efficiency*, *satisfaction* or *predictability*), which we can refer to here as merely "Opinion Score" for simplicity and illustrative purposes.

As illustrated by Figure 4, we plot the mean user rating (i.e. Opinion Score) as a surface for each of the network parameter conditions taken across all participants who experienced the given application. The black dots on the grid line vertices were actual conditions for which we have a (mean) measurement, while the surface itself is an interpolation. For each of the three quantifiable parameters (*efficiency*, *predictability* and *satisfaction*), an Opinion Score of '3' was required for participants to be satisfied with the service quality on an every-day basis.

Having obtained such a surface, we developed an important wireless access technology "positioning technique". As detailed in [10], [6] the technique involves superimposing the typical network parameter value ranges for a given technology such as 1xRTT (cdma2000) or the General Packet Radio Service (GPRS), onto an appropriate multi-dimensional "Knee of Satisfaction" curve such as that given in Figure 4.

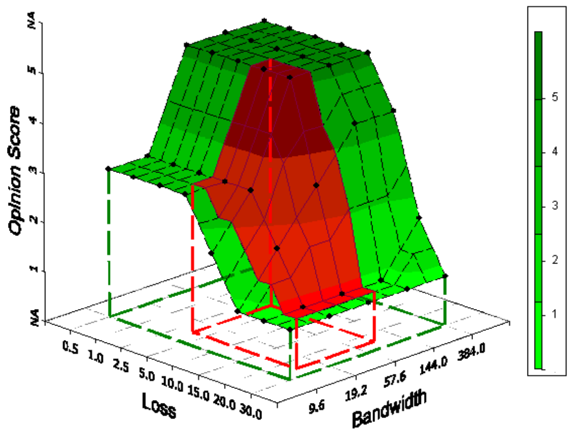

Figure 4: Wireless Access Network Technology 'positioning' on user-perceived "Opinion Score" (QoS) surface

Note that any wireless access technology will necessarily display a range of network conditions, so as our example in Figure 4 shows, the best we can do is to come up with an approximate "band" of likely network parameter values for that particular technology. The results from our experiment do the rest - namely, they allow us (using the interpolation mentioned earlier) to predict with a probabilistic certainty level, what the "Opinion Score" would be, if the given wireless access technology was used to deliver the specific application under study. By mapping the expected position in this way, we are able to gain insight as to the likely user-perceived QoS for each "positioned technology" across a number of different applications.

As simple and powerful as this positioning technique is, it does suffer from the fact that determination of the typical operating ranges for IP-level network parameters for each wireless access technology (x and y axes in Figure 4) is not a trivial task, due to the impacts of a wide variety of factors which would need to be investigated: radio conditions, network architecture, technology specific transmission characteristics, and network load to name but a few. The "technology positioning" results presented in [10] were based on *estimated* network conditions for given wireless access technologies - they depicted the range of performance possibilities of the given technology, rather than being based on an exhaustive set of measurements in a live network. This serves to further highlight the importance of and need for further research in this area.

VI. Conclusions

We have described our experiences in designing, implementing and using a wireless data network test-bed in support of a facilitator-led experiment measuring the quality of service (QoS) of specific applications delivered in a mobile setting. The test-bed was comprised of a wireless LAN (WLAN) and wireless-capable Personal Digital Assistants (PDAs) acting as "3G-like" user terminals with web surfing, E-mail retrieval, file download and video streaming capabilities.

A freeware network performance emulator at the IP-layer was used for controlled and reproducible mappings between application-specific QoS and relevant network parameters such as available bandwidth and packet loss probability. We also illustrated a technique used to predict whether a given wireless access technology could satisfy customers on a per-application basis.

Acknowledgment

The authors acknowledge Telstra for permission to publish this paper.

References

[1] Davidor Y. "Winning the wireless war", Telephony., June 5, 2000. http://telephonyonline.com/ar/telecom_winning_wireless_war/index.htm

[2] Wu J., "Performance analysis of QoS-based voice/data CDMA systems", Wireless Personal Communications, 13, 2000, pp223-236

[3] Lakelin P., "Internet service providers in Western Europe", http://www.isp-planet.com/business/lakelin-exec.html, (2000)

[4] Xylomenos G. and Polyzos C., "TCP Performance Issues Over Wireless Links", IEEE Communications Magazine, April 2001, pp52-58

[5] Pahlavan K. and Levesque A., "Wireless Data Communications", Proceedings of the IEEE, Vo. 82, No. 9, Sept 1994

[6] Beresford M., Saliba A., Ivanovich M. & Fitzpatrick P. "User-perceived Quality of Service in Wireless Data Networks", *submitted*.

[7] ETSI, "Network Aspects (NA); General aspects of quality of service and network performance in digital networks, including ISDN (ETSI ETR 003 ed.1)", 1990.

[8] ITU, "The E-Model, A Computational Model for use in Transmission Planning", ITU-T Recommendation G.107 (05/00)

[9] Fitzpatrick P., Ivanovich M., Beresford M. & Saliba A., "A Hierarchical Model for Quality of Service in Wireless Data Networks", *submitted*

[10] Ivanovich M., Fitzpatrick P., Beresford M. & Saliba A., "Defining and Quantifying the Quality of Service in Wireless Data Networks", *submitted*

[11] National Institute of Standards and Technology (NIST), http://snad.ncsl.nist.gov/itg/nistnet/, May 2002

[12] ETSI, "Universal Mobile Telecommunications System (UMTS); Quality of Service (QoS) concept and architecture *(3GPP TS 23.107 version 4.2.0 Release 4)*", October, 2001

[13] ETSI, "Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Service description; Stage 1 *(GSM 02.60 version 7.5.0 Release 1998)*", July 2000