Quality of Experience in real-time person-person communication – User based QoS expressed in technical network QoS terms

Bjørn Hestnes
Telenor Research – Norway
bjorn-olav.hestnes@telenor.com

Peter Brooks
Telenor Research – Norway
peter.brooks@teolys.com

Svein Heiestad
Telenor Research
Norway
svein.heiestad@telenor.com

Trond Ulseth
Telenor Research
Norway
trond.ulseth@telenor.com

Carl Aaby
Tandberg ASA
Norway
cha@tandberg.no

Abstract

There are many network and terminal characteristics that may interfere with real-time person-to-person communication services and that may result in a negative user perception of the service without fully understanding the reasons for apparent problems. This paper analyses what reduces the quality in new packet switched networks and identifies 'delay' and 'packet loss' as the important quality parameters for real-time communication. It shows how the technical QoS (Quality of Service) parameters of delay and packet loss can be expressed as a measure of QoE (Quality of Experience) by using a guideline derivation approach for expressing results from user-based laboratory experiments.

Key words: QoE, Quality of Experience, QoS, Quality of Service, real-time, person-person communication, Fitness-for-Purpose, user-based QoS, Internet protocol, UDP, User Datagram Protocol

1. Introduction

In real-time person-to-person communication services there are many network and terminal characteristics that may interfere with human communication. These technical characteristics are typically a topic of Quality of Service (QoS) parameters, with measures of QoS being based on theoretical mathematical and engineering principles and removed from measures of quality as perceived by end-users. In order to better represent the perspective of end-users the term 'Fitness-of-Purpose' has been applied and defined as: *The correct balance between technological performance and human performance, such that the interaction is both sufficient and beneficial for person-person communication and consistent with human expectations from face-to-face communication* [Brooks et al. (1999, 2003)]. Based on empirical tests of users, Fitness-for-Purpose has been measured using a range of objective

Hestnes, B., Brooks, P., Heiestad, S., Ulseth, T., Aaby., C. (2003). Quality of Experience in real-time person-person communication - User based QoS expressed in technical network QoS terms. . Proceedings of the *19th International Symposium on Human Factors in Telecommunication*, Berlin, Germany, December 1-4 2003, pp. 3-10.

measures such as task completion and communication behaviour and subjective measures such as task perception, person perception and service suitability [O'Malley et al. (2003)].

Similar concepts to Fitness-for-Purpose are 'user-based quality of service' and 'user perceived quality', which have not received detailed study or definition, but tend to have informal usage to describe something other than technically-oriented QoS. However, the concept of 'Quality of Experience' (QoE) is attracting more formalised and growing attention. QoE has originated from user-centred work on user-system interaction design [e.g. Alben (1996)] rather than user-user interaction design. QoE has recently been considered in relation specifically to eCommerce, Web design and multimedia services [Aldrich and R. T. Marshak (2000); van Moorsel (2001); Khirman and P. Henriksen (2002), Siller and Woods (2003)] and for mobile data services [Pearce (2002)]. QoE is also becoming recognised as an important construct in real-time human communications. Within standardisation the ITU-T lead Study Group SG 12 on QoS and performance is considering a proposal to start work on QoE [Nortel (2003)]. Early work has also been initiated in relation to Advanced Collaborative Environments [Corrie et al. (2003)] and on IP videoconferencing [O'Neil (2002)].

Although QoE is therefore attracting interest, the term has not been generically defined. Current descriptions tend to reflect a rather specific working focus. By developing definitions proposed by Nortel (2003) and Siller and Woods [2003] the current paper suggests that QoE is *The user's perception of what is being presented by a communication service or application user interface. It presents the overall result of the individual Quality of Services and is a measure of overall acceptability of a service or application that includes factors such as usability, utility, fidelity and the level of support from the application or service provider (e.g., sales, delivery, error corrections).*

This is intended as a rather comprehensive definition that can encompass more specific concerns in this very complex area. As the current paper focuses on the phase of actual user-user communication, it does not concern QoE aspects of sales process, delivery process, error situations or compensation situations. Also it does not address set-up, disconnection and user device control.

Service specific networks for real-time communication have been used for a long time, but now most network traffic migrates towards packet switched networks. Packet switching is used both in fixed and mobile networks and was developed primarily for transmitting non real-time traffic. When introducing real-time communication over packet switched networks, new characteristics are implied to the "old" and well-known services. The consequences for the users and their communication services may therefore be more difficult to predict.

This paper therefore analyses what reduces the quality in new packet switched networks and identifies the important quality parameters for real-time communication. It describes the most used enhancement mechanisms in networks and terminals. The paper then shows how these technical QoS parameters can be expressed as a measure of QoE by using a guideline derivation approach for expressing results from user-based laboratory experiments.

2. What decreases quality?

It is possible to analyse the communication traffic between users engaged in real-time communication by analysing the traffic from one user to the other(s) through the terminals and network and identifying where and in which way quality is reduced. Figure 1 shows the 'worst-case connection' and components that influence the quality for real-time person-to-person communication services such as voice telephony, video conferencing and multimedia conferencing [Heim et al. (2000)]. Each component is examined below with regard to videoconferencing, as this service covers those real-time applications that are most often thought of as similar to audio telephony and in addition put a still higher demand on the network and terminal.

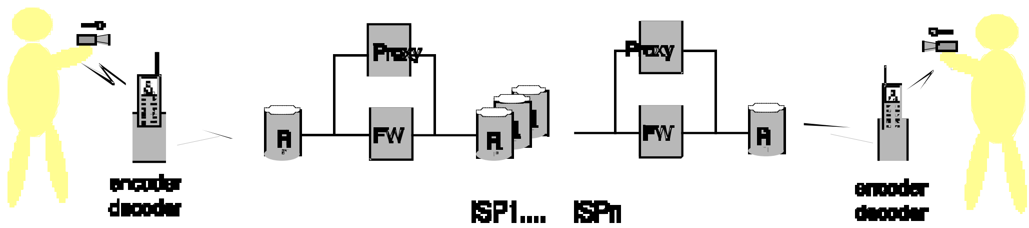

Figure 1. The physical QoS-view of a real-time person-to-person communication

2.1 Input devices

Users have two input devices that communicate with the corresponding output devices at the remote end: a **microphone** sends information to the remote **loudspeaker** and a **camera** sends information to the remote **screen**. The microphone reduces the bandwidth, introduces distortion and mixes sound with the room's (or environment's) acoustic characteristics. The **camera** reduces resolution, zooms in and shows parts of the site. A narrow angle will give more pixels per person than a wider angle for a group. In a CIF-mode (Common Intermediate Format which is 352x288 pixels) a view of one-person will represent about 50% of the pixels in a head-and-torso view. In a 3-group view, only 1/9 of the pixels will be available per user. That gives 7% or effectively 90x75 pixels per person in today's highest space resolution. In addition, the screen reduces visibility due to false light (e.g., sunlight) on the screen's surface that results in low contrast. A loudspeaker introduces acoustic distortion in its mechanical parts, reduces the bandwidth and mixes the sound with the environment's (e.g. room's) acoustic characteristics.

2.2 Components involved in a communication

Between the input and output devices there can be a number of components:

- **PANs** (Personal Area Networks) can be a cable-based Firewire or a wireless Bluetooth II connection/network. Quality can be reduced due to delay of information, loss of information and loss of connection.
- **Encoders** have the function in the codec to encode the information from an input device and compress it so it is made ready for transmission on the network. Quality for audio is influenced when the information is compressed with the introduction of distortion and delay. For video there will be a reduction in the number of frames, number of pixels in the horizontal and vertical direction and the introduction of delay and distortion.
- **Wireless network access** (for mobile networks) can reduce quality because of delay, loss of information and loss of connection.
- **Routers** introduce a longer delay of information and may introduce loss of information (depending on traffic conditions).
- **Tunnels**, **firewalls** and **encryption** introduce delay.
- **Domains** can reduce quality because of different QoS classes and models.

- **WANs** (Wide Area Networks) introduce delay in addition to the LANs (Local Area Networks), Routers and the Domains because electrons require time to be transported between sites.
- **Decoders** have the function in the codec to decode the information from the network, extract it and make it available for the output devices. Quality can be affected as described above for encoders.

An analysis of these reductions in quality [Heim et al. (2001)] has concluded that they can be categorised into three groups: **Network characteristics**, **Codec characteristics** and **Environment characteristics**. The network characteristics can be explained for the packet switched network as bandwidth, packet size, delay, delay jitter, packet loss, burst packet loss and sequencing. The codec characteristics are the media protocols (i. e. G.7xx, H.26x, MPEGx), video space resolution, video time resolution, delay, distortion and monitor size. The environmental characteristics are lighting conditions, background patterns, colour and reflex, acoustics, audio echo degradation, viewing distance, camera position and camera parameters. All of these parameters are therefore contributors to the reduction of quality in the communication phase. Further description of these technical parameters is provided in Heim et al [2001].

3.3 TCP-IP and UDP-IP protocols

TCP-IP (Transmission Control Protocol – Internet Protocol) is mostly used for non real-time applications such as Internet surfing and e-mail. However, it is also used for real-time person-person communication services that involve very low bit-rate (e.g. real-time text) and low bit-rate and push-to-talk applications (e.g. avatar-telephony). Media information is sent from the sender to the receiver as packets. The packets may not arrive at the receiving end before it is timed-out because of too long delay or too big a difference in delay jitter, packet loss, burst packet loss or sequencing. Then the packet will be retransmitted until it is received correctly. Therefore, of all the network characteristics, only the delay (including delay jitter) will be a quality parameter.

For real-time person-to-person communication where a continuous stream of media information is transmitted, TCP-IP will not be used because of its unpredictable delay when the network is overloaded. UDP-IP (User Datagram Protocol) is used for real-time person-to-person communication services such as VoIP (Voice over IP), video conferencing and multimedia conferencing. UDP-IP does not retransmit. If a packet is lost for any reason (arriving too late, lost or becoming corrupted) then the application looses it. Therefore UDP-IP communication results in two quality parameters: delay (including delay jitter) and packet loss (including burst packet loss).

4. Quality enhancement mechanisms

Within this paper's technical quality focus it is relevant to consider two areas of enhancement mechanisms, concerning either the network or the terminal.

4.1 Network enhancement mechanisms

In an IP network the most frequent way to solve the reduction in quality for real-time communication is to implement a priority mechanism for certain packets. The three main QoS classes are Guaranteed QoS, Real-time QoS and Best effort QoS. For a Guarantied QoS class, a mechanism in the network (e.g. IntServ, RSVP, IP over ATM) must be implemented from end-to-end so the connection has a reserved bit-rate. For a Real-time QoS class a priority is made based on statistical multiplexing. If the traffic is much more than planned, more packets are lost and the QoS is reduced. For a Best effort class, no priority is made and QoS is dependant on the concentration factor among subscribers and their usage pattern.

4.2 Terminal enhancement mechanisms

If there is a large delay in the network, the terminal can increase the jitter buffer to adapt to this situation by increasing the length of the buffer. This will lead to an even longer delay. It may be better not to wait and instead lose the packet. The terminal developer has to decide when the delay should be considered too long to deal with, so that network delay is regarded as packet loss.

If packet loss occurs, the audio and video quality will suffer: The audio will have a lot of extra noise and the video image will contain distortions. The terminal can reduce the effect of the packet loss by different mechanisms for the video and the audio.

For video, packet-handling mechanisms such as Intelligent Packet Loss Recovery [Patent pending] may be implemented in terminals to decrease the user awareness of packet loss. In the event of packet loss, the encoder will enter a 'robustness mode', making the encoder begin transmitting blocks of I-frames. An I-frame is a compression technique for still images that exploits the redundancy within the image and that can be applied to individual frames of a video sequence. It is distinct from P-frame compression that is applied to a sequence of video frames rather than a single frame and exploits similarities between frames. The I-frame blocks are distributed over several frames in such a way that the image is completely rebuilt over a number of frames. Hence the video data received by the decoder is a combination of I-frame blocks and P-frame blocks, instead of all blocks being received as P-blocks. This will increase the need for bandwidth for the video and the consequence is that the frame rate drops. It is expected that reduced frame rate will be better for users than a distorted and 'blocky' image.

For audio, as packet loss will result in noise it is common to decrease the volume of the output when data is lost; i.e. instead of playing the sound as normal with noise in the audio stream, the volume is decreased when one or more packets are lost. Therefore the user will experience a 'clipping' in the sound instead of arbitrary noise. This is considered less annoying for the user and tests have shown that even a relatively large amount of packet loss does not impact the communication [Hestnes et al., 2003].

5. Guidelines as a way to express QoE in technical QoS terms

One real-time person-person communication application is remote inspection. Remote inspection is asymmetric video conferencing and is identified as an important way that new mobile terminals will be used. With a 2-way audio-channel (telephony) together with a 1-way video-channel the application is initiated from a mobile person (e.g. field worker) to a remote person who acts on the video and audio information received. Typically both persons see and talk about the area or object at which the camera is pointing [Hestnes et al., 2001].

O'Malley et al. [2002] have performed a laboratory experiment using a remote inspection task. The video and audio channel were divided into separate networks, with the audio channel using a non-packet switched line (an ordinary line switched channel like ISDN or GSM) and the video using a packet switched IP channel. This configuration is the expected implementation of the 3G mobile network UMTS (Universal Mobile Telecommunication

System). The video channel was tested with 0% packet loss together with 1%, 2%, 7% and 15% packet loss.

Based on this work, guidelines have been developed with essential information extracted from the experiments [Hestnes et al., 2003]. The guidelines have a simple clause:

IF <communication situation>
USING <service prescription>
WITH <technical parameters>
THEN <user behaviour>.

The attributes <communication situation>, <service prescription>, <technical parameters> and<user behaviour> have sub-attributes and sometimes sub-sub-attributes in order to cover the problem space and to correspond to existing knowledge of media effects on communication behaviour. For example, the attribute 'Communication Situation' has the sub-attributes 'Task', 'Motive', 'Setting' and 'User'; and 'Task' is defined by 11 sub-sub-attributes including 'Duration', 'Situation formality', 'Difficulty' and 'Urgency'. The ‘Service prescription’ containes the service used (i. e. telephony or video conferencing), the ‘technical parameters’ concern network delay or packet loss and ‘user behaviour’ includes variables such as communication efficiency and user satisfaction.

With this essential information collected and structured, meta-level guidelines have been abstracted that state the main message for each of the detailed guidelines [Brooks and Hestnes (2003)]. For the two critical quality parameters of delay and packet loss, the following guidelines exist:

- Remote inspection with 500ms asynchrony for giving advice on a procedure does not affect task performance (compared with remote inspection with no delay)
- Remote inspection with 7% packet loss for giving advice on a procedure does not affect task performance (compared with remote inspection with no packet loss).

6. Conclusions

A definition of QoE has been introduced that includes user-user interaction in addition to user-system interaction and that emphasises user-based performance as a result of individual technical QoS parameters. This paper has identified that the two essential technical QoS parameters in a packet switched network for real-time person-to-person communication are delay and packet loss. Knowledge of the implications that delay and packet loss have on QoE can be expressed as guidelines intended, in particular, for network operators and equipment manufacturers. This information is identified as useful for network operators who need their quality parameters expressed in a user-centred way in order to know the implications that the network has on end-user usage. Similarly, equipment manufacturers can make use of knowledge of user behaviour in relation to technical QoS parameters because the terminal can be designed to enhance quality in terms of delay and packet loss. Market rejection can happen if users perceive the quality of communication is not high enough in comparison with other services and face-to-face communication. By measuring technical network quality in terms of QoE, network and terminal designs have increased potential to meet users ' communication requirements.

Acknowledgements

The basis of this work was performed as part of the EC project Eye-2-Eye – Fitness-for-purpose of Person-Person Communication Technologies (IST-1999-11577). The authors have benefited substantially from the whole project group: Asbjørn Følstad, Jan Heim, Trond Schliemann, Jan Håvard Skjetne (SINTEF); Han Frowein (Instituut voor Doven); Patrick Brundell, Claire O'Malley, (University of Nottingham); external peer reviewers Pedro Concejero (Telefónica Investigación y Desarrollo) and Manfred Laube (Alcatel SEL).

References

Alben, L. (1996). Quality of Experience. *ACM Interactions*, vol. 3 (1996), pp. 11-15, also available at http://www.albenfaris.com/publications/pub_qofe.shtml.

Aldrich, S. E. and R. T. Marshak (2000). *Customers.com® Quality of Experience Benchmark*. Patricia Seybold Group August 31, 2000, available at http://www.qualityofexperience.org/QoE_original.pdf.

Brooks, P., Brundell, P., Hamnes, K., Heiestad, S., Heim, J, Hestnes, B., Heydari, B., O'Malley, C., Schliemann, T., Skjetne, JH., Ulseth, T. (1999). Final Report. ACTS Project AC314 *Vis-à-Vis: Fitness-for-Purpose of Videotelephony in Face-to-Face Situations*. . (Available from www.eye-to-eye.org)

Brooks, P. and Hestnes, B. (2003). User-centred technical guidelines for real-time human communication services: requirements and derivation. In *Proceeding of Human Factors in Telecommunications 2003*, Berlin, December 2003

Brooks, P., Schliemann, T., Hestnes, B., Frowein, H., Aaby, C., O'Malley, C. (2003). *Fitness-for-Purpose of Person-Person Communication Technologies*. Final Report Project IST-1999-11577 Eye-2-Eye. (Available from www.eye-to-eye.org)

Corrie, B., H. Wong, T. Zimmerman, S. Marsh, A. S. Patrick, J. Singer, B. Emond and S. Noël (2003). Towards Quality of Experience in Advanced Collaborative Environments. *Third Annual Workshop on Advanced Collaborative Environments*, Seattle, June 2003.

Heim, J., Asting T., Brandtzæg, P., Brooks, P., Skjetne, J. H., Hestnes, B., Heiestad, S., Kjellmo, I., Nilsen, S. J., Ulseth, T. Frowein, H., Aaby, C., (2000). *Initial Verification of Real-time Communication Requirements*. . (Available from www.eye-to-eye.org)

Heim, J., Brooks, P., Følstad, A., Schliemann, T., Hestnes, B., Ulseth, T., Heiestad, S., Frowein, H., Devoldere, P., Aaby, C., O'Malley, C., Brundell, P., Londsdale,.P., (2001). *Final Verification of Real-time Communication Requirements.* (Available from www.eye-to-eye.org)

Hestnes , B., Heiestad, S., Brooks, P., Drageset, L., (2001). Real situations of wearable computers used for video conferencing – and implications for terminal and network design. *The fifth international symposium on wearable computers*, pp. 85-93, IEEE computer Society. ISBN 0-7695-1318-2, Zürich 8-9 October 2001.

Hestnes, B., Heiestad, S., Ulseth, T., Schliemann, T., Brooks, P., Følstad, A., Frowein, H., Aaby, C., O'Malley, C., Brundell, P., (2003). *Fitness-for-Purpose Guidelines for Person-Person Communication*. (Available from www.eye-to-eye.org)

Khirman, S. and P. Henriksen (2002). Relationship between Quality of Service and Quality of Experience for Public Internet Service. *Passive and Active Measuring Conference* (PAM 2002), March 2002 available at http://www.labs.agilent.com/pam2002/proceedings/.

Nortel Networks (Canada) (2003). *Quality of Experience*. ITU-T COM 12-D95, Geneva 27-31 January 2003.

O'Malley, C., Brundell, P., McFadzean, J., Lonsdale, P., Schliemann, T., Brooks, P., Følstad, A., Heim, J., Hestnes, B., Heiestad, S., Ulseth, T. , Frowein, H., Devoldere, P., Aaby, C., (2002). *Results of Laboratory Experiments of Communication Media*. (Available from www.eye-to-eye.org)

O'Neil, T. M. (2002). *Quality of Experience and Quality of Service for IP videoconferencing*. Polycom White Paper available at http://www.polycom.com/common/pw_item_show_doc/0,1449,978,00.pdf.

Patent pending US provisional 60/423,393

Pearce, J (2002). *Quality of eXPerience (QXP ™)* White paper Argo Group available at http://www.argogroup.com/pressroom/QXP.pdf.

Siller, M. and J. Woods (2003). Improving Quality of Experience for MultiMedia Services by QoS arbitration on a QoE Framework. *Packet Video 2003*, Nantes, April 28-29.

van Moorsel, A.(2001). *Metrics for the Internet Age: Quality of Experience and Quality of Business*. HP Laboratories, Palo Alto, Report HPL-2001-179, July 2001.