QOE–DRIVEN MOBILITY MANAGEMENT – INTEGRATING THE USERS' QUALITY PERCEPTION INTO NETWORK–LEVEL DECISION MAKING

Martín Varela, Jukka–Pekka Laulajainen

VTT Technical Research Centre of Finland
PL 1100, Oulu 90571, Finland
{martin.varela | jukka-pekka.laulajainen}@vtt.fi

ABSTRACT

One of the most interesting applications of single–sided Quality of Experience (QoE) metrics is their use in improving the quality of the service, as perceived by the user. This can be done either at the application level – by for example changing the encoding in use, or the level of error correction applied – or at the network level, for instance by choosing a different DiffServ marking strategy, or changing the access network in use. To this end, the QoE metric used needs to be fast and accurate, and the context in which the application will be used needs to provide the opportunity for performing some sort of control operation. In this paper we describe the application of QoE estimations for VoIP to improve existing network–level mobility management solutions.

***Index Terms*—** Quality of Service, Internet telephony, Computer network management, Wireless networks, Multimedia communication

1. INTRODUCTION

Quality of Experience has been, as of late, receiving an ever–increasing amount of attention from both the research community and the industry. The term itself lends to definitions of varying width. In this paper we will use the term interchangeably with *"the quality of a (media) service as perceived by its end users"*, and will not consider other aspects of User Experience, which are sometimes included in the QoE usage. In any case, QoE is currently a hot buzzword in both academia and industry, and with good reason. With the proliferation of networked services, and media services in particular, their quality becomes a significant issue from both the users' and the providers' points of view.

Delivering the best possible QoE for a given service, if possible at a minimum cost, has become a common research issue. In this work, we address an instance of this issue as it pertains to the quality of Voice over IP (VoIP) services in a mobile context. More precisely, we refer to the user's mobility, which can happen over a variety of network technologies, and in particular across them. The most commonly used mobility solution nowadays is Mobile IP [1], which provides a mechanism for mobile hosts to connect to networks in different administrative domains while maintaining a single IP address in its *home network*.

The possibility of moving between different access networks, coupled with the current ubiquity of devices with multiple network interfaces for different technologies (e.g. laptops, smart-phones, tablets), allows for new quality–improvement mechanisms to be developed. The solution we present in this work uses Pseudo–Subjective Quality Assessment (PSQA) [2] to estimate the call quality, and it allows for automatic access network switching when said quality degrades beneath a certain threshold.

The rest of the paper is organized as follows. In Sections 2 and 3 we briefly describe the technologies involved (for mobility management and QoE estimation, respectively). We describe our proposed control mechanism in Section 4. Section 5 describes the experimental setup we used for validation and Section 6 describes the results obtained. We conclude the paper in Section 7 providing pointers for further research in this area.

2. ABOUT MOBILE IP

Mobile IP (MIP) [1] provides a solution for users who need to attach to different networks[1], while being still reachable through their own *home network*. To this end, an IP tunnel is set between the *mobile node* and a *home agent*, which allows relaying traffic to the mobile node at its *home address* in the home network. Several techniques can be used for tunneling the connection, since the different *foreign networks* to which the mobile node can attach might have different features. It is common for NAT (Network Address Translation) traversal techniques to be used, since NAT is used in a very large proportion of today's networks.

This work was partly supported by the CELTIC Easy Wireless 2 and CELTIC IPNQSIS projects, Tekes, and VTT Technical Research Centre of Finland.

[1]Different, as in different administrative domains

As the mobile node moves between networks, *mobility bindings* are created, which establish the link between the node itself and its *care–of address*, namely the address it is assigned in the foreign network. These bindings are what allow traffic to be directed to the mobile node's current address in the foreign network.

In the case of multi–access devices, MIP can be used to roam (seamlessly, in many cases) between different access networks, while still having a single IP address visible to the outside world. Hence, for example, a user could be working with a laptop connected to an Ethernet link, unplug it, thus roaming to a WiFi network, and as he walks away from the coverage area, roam once again to an HSPA network, all without losing connectivity at any point, and still being reachable on the same IP address.

Current MIP solutions implement priority–based policies for choosing the interface to be used, in the case several are present. These policies tend to be based on expected network performance and connection costs, and thus end up favoring wired connections, and putting 3G/HSPA ones a lower priorities, as they are usually more costly[2] and typically have lower performance.

A common problem, however, is that current MIP solutions tend to neglect the actual performance of the network and applications when considering which access network they will attach to. That is, in the scenario described above, as long as the home network could be reached through the Ethernet access, the mobility manager would not switch to a different network even in the case where the QoS of that particular Ethernet access, and hence for example the QoE of media applications, were to become unacceptable for the user. This kind of situation arises often in the face of congestion, especially in wireless networks, and it is this problem that we address with our proposal.

3. ABOUT PSQA

Pseudo–Subjective Quality Assessment (PSQA) [2] is a generic methodology for estimating QoE for media applications. It is a parametric model, which takes both network QoS parameters and application–specific parameters, such as encoding, error correction, and so on. Put very simply, PSQA makes a mapping between the parameters considered, and the subjective quality as measured for example by standard assessment methodologies, such as ITU-T's P.800 [3] in the case of listening quality for voice applications. Other possible quality estimators can be considered, although most of the PSQA implementations to date produce MOS–type values. The mapping is usually done by having a neural network (a Random Neural Network, or RNN [4], in most cases) learn the relation between a carefully selected set of parameters and their values, and their impact on subjective quality.

[2]This has become less of an issue in the last few years, though.

PSQA has been successfully used to estimate VoIP listening quality [5], VoIP conversational quality [6], and video quality[7, 8, 9], yielding very high correlation with subjective scores (usually in the high 90's for all applications considered so far[3]) while having a trivial computational cost, and being able to produce the estimations in real–time even in very resource–constrained devices.

PSQA has also been used in more theoretical contexts as a way to supplement traditional performance evaluation methods [10], and as a way of defining utility functions for pricing purposes [11].

4. SMART MOBILITY MANAGEMENT

We propose to address the problem stated at the end of Section 2, namely, the current MIP solutions' lack of consideration for application– and network–level quality and performance, by using PSQA–based estimations as an additional driver for mobility.

To this end, we have implemented a proof–of–concept smart mobility manager, which extends the functionality of Birdstep's[4] MIP solution. At a very high level, our extension performs passive network QoS monitoring (using the QoSMeT [12] tool) for the VoIP traffic, and feeds that network QoS information to PSQA. Both listening and conversational quality estimations can be used to drive the mobility manager. The application and network parameters considered vary according to the type of estimation used. We used PSQA implementations based on data from [5] and [6] for listening and conversational quality estimations, respectively. The listening quality estimations consider the following parameters:

- Codec used (either GSM or PCM Linear–16)
- Media dependent FEC (as described in [13])
- Packetization interval (from 20ms to 80ms of speech per packet)
- Packet loss rate
- Temporal distribution of the losses (as mean loss burst size)

For the conversational estimation, the following parameters are considered:

- Bit–rate (only Speex was used for encoding, with a wide variety of bit–rates)
- FEC (media independent, as described in [14])
- Packet loss rate

[3]This is comparable to PESQ's for voice applications, but without the need for a reference signal

[4]Formerly Secgo.

- Temporal distribution of the losses (as mean loss burst size)
- One–way delay
- Jitter (average variation of one–way delay inter-packet arrival times)

The proposed solution uses the QoE estimations provided by the QoSMeT + PSQA combination to decide whether a handover to a different access network is needed. In order to avoid a very "jumpy" behaviour, several strategies can be chosen for the decision making. The strategy we find most usable considers a sliding weighted average of the MOS estimations over a configurable time window, which offers a good compromise between stability and reaction time. A simple, non-weighted average over said time window can also be used for a more stable behaviour, and it is also possible to consider only the last estimation before deciding whether to perform a handover. This last approach, however, is not very usable in practice, as the estimations are done every second, and minor impairments in the network might result in unwarranted handovers.

In order to avoid unstable behaviour, a configurable time period between successive handovers can be enforced, and it is also possible to configure the manager to roam back to the original access after a certain period of time, which is useful in the case of temporary congestion or interference.

The actual handovers are implemented by the Birdstep MIP software, and are usually seamless, as connections to the underlying access networks are kept open, even when not in use. The QoE–based handover can be either forced (which results in an immediate handover), or suggested to the Birdstep software by changing the priorities of the available access networks, in which case the actual handover might take a few seconds longer. In this latter case, the Birdstep MIP software retains control over further handover decisions, whereas in the case of forcing the handover, it does not.

As a result of these options, the mobility management can be finely tuned to suit a wide variety of application and network contexts and users' needs.

5. EXPERIMENTAL SETUP

Unfortunately for this kind of application, no amount of curves and figures can replace the experience of seeing (and hearing) the system perform live. For this paper, however, we crafted two scenarios which help convey how the improved, QoE–driven mobility management works in what could be considered typical usage contexts.

5.1. Scenario 1: Network Congestion

One common problem in any best–effort network is the eventual emergence of congestion. This is usually a temporary problem, but it happens often enough that it is probably the most common cause for bad quality in networked media. To demonstrate the system's performance in a congestion scenario, we set-up a mobile node with both Ethernet and WiFi interfaces. The Ethernet interface was connected to a Linux–based router running NetEm [15]. We used NetEm to emulate the congestion, by setting a relatively high delay value for a wired network (around 100ms), and then adding losses (up to 8%) to emulate a period of more intense congestion in the network. While the setup was done as described for practicality's sake, very similar behaviours can arise in basically any network access, making this scenario relevant in most usage scenarios. Counterpath's X–Lite was used to generate the voice traffic.

Our QoE–driven mobility manager was configured to use a weighted average of MOS estimations over a $10s$ sliding window, and to roam back to the original access network after about $35s$ had elapsed since the handover. The decision to make the handover was based on a listening quality estimation (conversational quality was also estimated and recorded).

5.2. Scenario 2: Interference in WLAN

The second scenario considered is that of radio interference when using a wireless access network. To this end, we equipped our mobile node with a WiFi card and an HSPA card. The WiFi network was configured to use fixed channel and pulsed White Gaussian Noise (WGN) was generated using a Rhode & Schwarz SMBV100A signal generator, taking care to cover the whole bandwidth of the chosen WiFi channel. The QoE–driven mobility manager was configured in the same way as in the first scenario.

6. RESULTS

The results presented herein are representative of the overall behaviour of our proposed QoE–driven mobility management solution.

Figures 1 and 2 show results for a test run under the conditions of the first scenario, as described in Section 5.1 above. It can be seen that losses start occurring at about the $16s$ mark. As seen in Figure 1 the listening quality degrades noticeably, at this point, but the weighted average remains above toll quality (the QoE threshold for the handover was set at 2.9 in a 5–point MOS scale), and as quality is still acceptable, no handover takes place. Over the next $20s$, the quality keeps degrading, until a handover takes place at about the $36s$ mark. Finally, at around $76s$ a handover is performed back to the original network, which is no longer so congested. In this particular instance, a slightly more reactive configuration might have been useful, as the weighted average hovered just around toll quality for several seconds. Figure 2 shows the evolution of conversational quality over the course of the same test. Delay is also shown in this plot. Several interesting things can be gleaned from these two figures. The first is that the handover manager does perform as expected. Had it not been present, the normal MIP software would have continued on the first access network, despite the lower quality, as seen on Figure 3. Secondly, some differences in behaviour are observed in the listening and conversational QoE estimations. Firstly, the application contexts used for the estimations are very different, and the corresponding subjective tests campaigns were also different, hence a direct comparison of the corresponding MOS values does not make sense. However, the interactive campaign showed that at similar bit-rates, the Speex codec provides significant better quality than GSM. Also, despite the difference in absolute values, both listening and conversational quality estimations present a very similar behaviour in the face of network congestion.

As can be seen, the lack of consideration for jitter in the listening QoE estimation results in a "flat line" in the absence of losses, whereas the effects of jitter are visible in the conversational QoE plot. Moreover, it can be seen that WiFi network has significantly more jitter than the NetEm link, as displayed by the more variable MOS values during the time in the WiFi network. The reduced one–way delay in the WiFi link also results in a very slight increase in overall conversational quality, although, as reported in [6, 16], the impact of delay (in the absence of echo) on conversational quality is seemingly less significant than expected according to common wisdom (usually based on the ITU-T G.114 recommendation [17]).

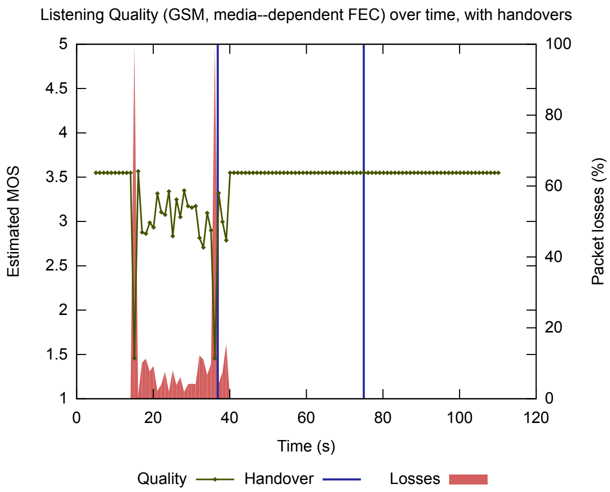

Fig. 1. Listening quality over time for Scenario 1. When problems in the network significantly reduce the perceived quality, a handover to a different network is performed. After a period of about $35s$, the device roams back into the original network.

Figures 4 and 5 show results for a test run of the second scenario. Compared to the first scenario, the network impairment caused by the radio interference is significantly more

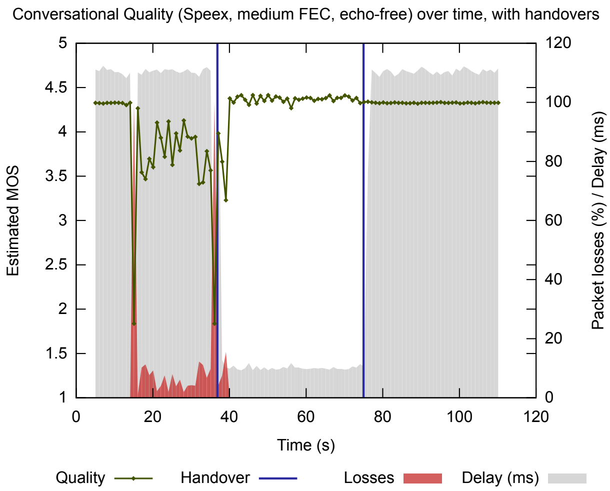

Fig. 2. Conversational quality over time in Scenario 1. This plot shows the evolution of QoE over time, but the handover decision was made based on listening quality

severe than the one caused by slight congestion. The IEEE 802.11 MAC in the WiFi link can't conceal the physical level problems in the presence of severe radio interference and even though over 60% of the traffic at the MAC level are retransmissions of lost frames, the losses start to appear also at the IP level. Those packets that get through suffer from high delay and jitter caused by the MAC level retransmission procedure. It should be noted that the current MIP solutions would not perform a handover even in this extreme case as long as the link to the home agent still existed. The violent impairment of the network results in a very swift handover decision. It can be seen that both delay and absolute jitter are higher in the HSPA network than on the unimpaired WiFi link. However given that the conversational quality estimator resulting from [6] considers jitter values relative to the overall one–way delay values, the estimation is less variable in the HSPA network than on the WiFi one. Whether this is actually accurate or not will depend on how the dejittering buffer is implemented on the VoIP application. In any case, the variation is perceptually insignificant.

Once again, the QoE–driven handover performed as expected.

6.1. Caveats and possible improvements

The current prototype is still at an early stage of development and it lacks some functionality that could yield larger improvements in QoE. The most important limitation is that no probing of the available interfaces is done before performing a handover, so the handovers are directed towards the next available interface, using the same priorities as the MIP software. In order to implement this probing, the passive monitoring approach currently implemented would need to be sup-

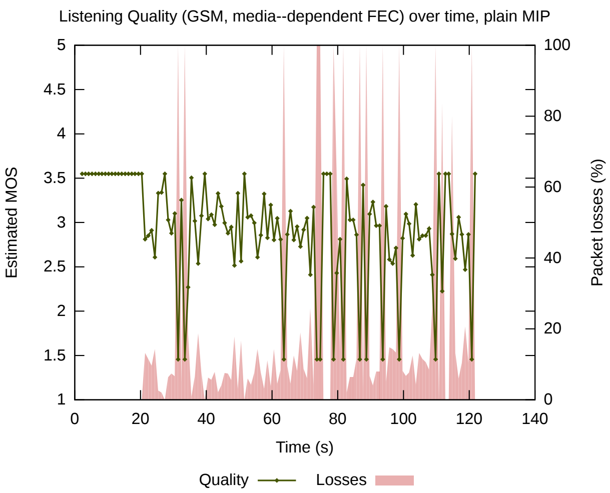

Fig. 3. Baseline MIP performance for Scenario 1. Despite a very noticeable decrease in listening quality, the MIP software does not perform a handover, as it can still reach the home agent through the current connection.

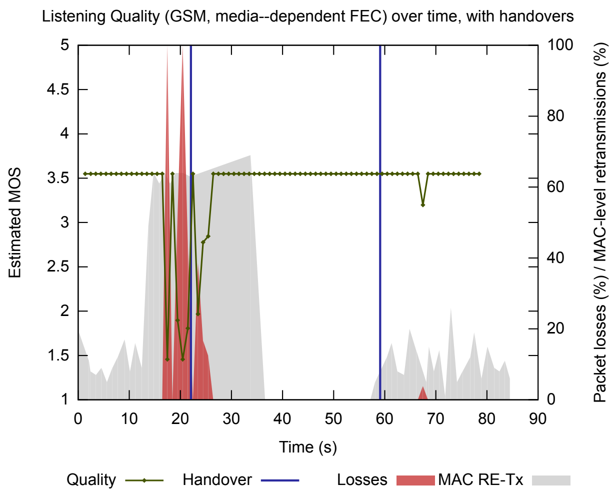

Fig. 4. Listening quality over time for Scenario 2. The radio interference creates a sharp spike in MAC–Layer retransmissions and IP–level losses, and a handover is quickly required.

plemented with periodic active probing of the available interfaces, possibly with synthetic media traffic between the mobile node and the home agent. This would allow for optimal interface selection, which could certainly improve the overall QoE. However, this would also create new issues that require careful consideration. Firstly, tighter integration with the MIP software would be needed, as an extra service (at the home agent) and means to route traffic "below" the MIP software (at the mobile node) would be needed, which is not usually possible in common usage. Secondly, considerations of cost and power consumption would also be relevant, as periodic probing would result in a non–trivial amount of extra network traffic over otherwise inactive interfaces. Finally, there might be scalability issues at the home agent, as presumably many clients would make use of this probing facility. These issues, as well as the likely benefits of active probing need to be carefully considered and weighted.

Another possible improvement would be to allow for shorter inter–handover times if a recently chosen interface does not provide the expected QoE levels. This is trivial to implement, but it needs to be tested with a proper subjective campaign in order to determine whether the expected increase in QoE would outweigh the increased QoE variability that would accompany the more dynamic handover policy.

7. CONCLUSIONS AND FUTURE WORK

In this paper we have presented a working improvement to current IP mobility management solutions, based on the QoE of the applications considered. In particular, we focused on VoIP as an application, but the same solution could in principle be applied to any application with strict QoS constraints. For the case of video, existing PSQA implementations could be easily included into the current prototype.

We have tested our prototype implementation in scenarios representative of real–world usage, and provided a detailed account of our solution's performance. In summary, considering QoE as a factor when making vertical handover decisions in a MIP context is a marked improvement over current available mobility management techniques.

The solution proposed is very lightweight and usable in any kind of device capable of running networked media applications. It relies on PSQA for accurate, real–time QoE estimations.

Future work on this domain spans two main branches. On the QoE assessment side, using other PSQA implementations could expand the usefulness of this approach to other applications, such as video, or eventually on-line gaming. On the network management side, many different new techniques can be thought of based around the same "QoE–driven" principle. Voice–call-continuity (VCC) and its successor in LTE networks is a clear candidate for integrating QoE estimations into the decision making. Moving from the terminal to the network components, several mechanisms, ranging from access control to DiffServ marking can also benefit from having QoE–awareness. This line of research is the focus of the just–started CELTIC IPNQSIS project, so further results in this domain are expected in the short term.

8. ACKNOWLEDGMENTS

The authors would like to thank Dr. Jarmo Prokkola for his help with the testbed setup required for the tests presented in

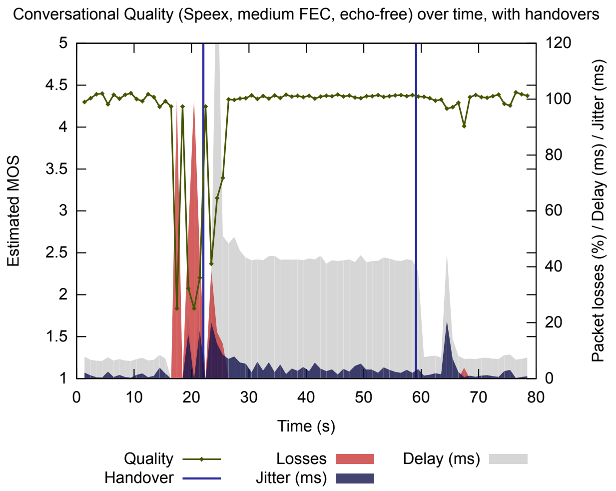

Fig. 5. Conversational quality over time in Scenario 2. This plot shows the evolution of QoE over time. In this case, the handover decision was made based on listening quality. Note the increased one–way delay in the HSPA network, but lower relative jitter values.

this paper.

9. REFERENCES

[1] IETF Network Working Group, "IP Mobility Support for IPv4 (RFC 3344)," Aug. 2002.

[2] Martín Varela, *Pseudo-Subjective Quality Assessment of Multimedia Streams and its Applications in Control*, Ph.D. thesis, INRIA/IRISA, univ. Rennes I, Rennes, France, Nov. 2005.

[3] ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality," 1996.

[4] Erol Gelenbe and Jean–Michel Fourneau, "Random neural networks with multiple classes of signals," *Neural Computation*, vol. 11, no. 3, pp. 953–963, 1999.

[5] Gerardo Rubino, Martín Varela, and Samir Mohamed, "Performance evaluation of real-time speech through a packet network: a random neural networks-based approach," *Performance Evaluation*, vol. 57, no. 2, pp. 141–162, May 2004.

[6] Ana Couto da Silva, Martín Varela, Edmundo de Souza e Silva, Rosa Leão, and Gerardo Rubino, "Quality assessment of interactive real time voice applications," *Computer Networks*, vol. 52, pp. 11791192, Apr. 2008.

[7] Samir Mohamed and Gerardo Rubino, "A study of realtime packet video quality using random neural networks," *IEEE Transactions On Circuits and Systems for Video Technology*, vol. 12, no. 12, pp. 1071 –1083, Dec. 2002.

[8] Julio Orozco, *Quality of Service Management of Multimedia Flows Over DiffServ IP Networks*, Ph.D. thesis, INRIA/IRISA, univ. Rennes I, Rennes, France, Mar. 2005.

[9] Ana Couto da Silva, Pablo Rodriguez-Bocca, and Gerardo Rubino, "Optimal Quality-of-Experience design for a P2P Multi-Source video streaming," in *Communications, 2008. ICC '08. IEEE International Conference on*, May 2008, pp. 22 –26.

[10] Gerardo Rubino and Martín Varela, "A new approach for the prediction of End-to-End performance of multimedia streams," in *First International Conference on Quantitative Evaluation of Systems (QEST'04)*, Sept. 2004.

[11] Yezekael Hayel, Gerardo Rubino, Bruno Tuffin, and Martín Varela, "A new way of thinking utility in pricing mechanisms: A neural network approach," in *Proceedings of the 13th CLAIO (Congreso Latino-Iberoamericano de Investigacin Operativa*, 2006.

[12] Jarmo Prokkola, Pekka H. J. Perälä, Mikko Hanski, and Esa Piri, "3G/HSPA Performance in Live Networks from the End User Perspective," in *ICC*, 2009, pp. 1–6.

[13] Jean Bolot and Andrés Vega Garcia, "The case for FEC-based error control for packet audio in the internet," in *ACM Multimedia Systems*, 1996.

[14] Daniel Ratton Figueiredo and Edmundo de Souza e Silva, "Efficient mechanisms for recovering voice PAckets in the internet," in *Globecomm'99*, 1999.

[15] Stephen Hemminger, "NetEm website," `http://www.linuxfoundation.org/collaborate/workgroups/networking/netem`.

[16] Florian Hammer, *Quality Aspets of Packet–Based Interactive Speech Communication*, Ph.D. thesis, T.U. Graz, Vienna, Austria, June 2006.

[17] ITU-T Recommendation G.114, "One-way transmission time," 2003.