The challenges paving the way toward a revised ITU-T P.561 Recommendation

Vincent BARRIAC, France Télécom R&D, D.I.H./E.Q.S. Laboratory, fax : +33 (0)2 96 05 13 16,
electronic address : vincent.barriac@francetelecom.com

Summary

This document, issued by the Rapporteur of Q.16 on non-intrusive measurement techniques at ITU-T Study Group 12, is intended to be a basis for the discussion on the features that need to be updated in the ITU-T P.561 Recommendation. After a review of the actual state of this Recommendation, which is well adapted for transmission in time division multiplexing mode but not for packet-based networks, several proposals of improvement are made, to adapt the future INMDs to the context of VoIP, by specifying the way the signal is captured and reconstructed from the Ethernet streams, adding new measurement parameters and a new class of devices, and measuring some parameters on the basis of multi-point signal acquisition. A revised version of P.561 is waited to be approved by the ITU-T before the end of this year.

Keywords

Standard, non-intrusive, VoIP

Introduction

As Rapporteur for Q.16 on non-intrusive measurement techniques at ITU-T Study Group 12 for the study period 2001-2004, the author of this document is, amongst other tasks, responsible for the update of the existing ITU-T Recommendations under the responsibility of this question (i.e. G.115, P.56, P.561 [1] and P.562 [2]). The two first recommendations are related to speech levels in networks and do not require any revision. The situation is different for the two others, which describe how to measure different voice signal parameters in a non-intrusive way and how to interpret the measurement results, both tasks highly depending on the transmission technologies used in networks, and thus needing to be adapted to the new problematic brought by the use of packet-based techniques (like IP) for the transport of voice communications.

This contribution has no specific novelty to present. It is only intended, as a starting point for a discussion between experts and interested people, to pinpoint the actual study items that need to be addressed during the necessary revision of P.561 (which is meant to be completed during the year 2002) and how they can be treated, based on the knowledge of experts from the fields of voice over the Internet protocol (named VoIP in the following of this document), signal processing and measurement techniques. The subject of P.562 will not be addressed directly in this document, although some of the considerations below will be useful for a future update of this recommendation.

State of the art

The standardisation of non-intrusive measurement techniques started in northern America in the late 80s (first standard issued in 1991 [3]). In 1994 (but officially approved and published in February 1996), the ITU-T issued its own standard in this topic, under the reference P.561.

This international standard gives a list of mandatory parameters to be measured by In service Non-intrusive Measurement Devices (INMDs), in which range and with which accuracy, and describes a rough testing procedure to verify that INMDs are compliant with the standard.

These parameters and their respective measurement ranges and accuracies are given in Table 1 below, taken from section 8.3 of P.561.

These parameters are not related to a specific transmission technology, but the standard has been issued at a time when only Time Division Multiplexing (TDM) technologies were used for the transmission of voice traffic. Now that packet-based technologies are more and more widespread, a need raises to revise the content of P.561 Recommendation, in order to keep it up to date by making it applicable for any type of transmission technology. Otherwise, measurement device manufacturers can be tempted by developing and selling INMD-like tools that will measure inappropriate indicators with an unverified accuracy.

Measurement		Range		Mean accuracy (Note 2)	Resolution (Note 4)
		Lower	Upper		
Active speech level (dBm) (Note 1)		–35 –14.9	–15 0	± 0.3 ± 0.3	0.1 0.1
Speech activity factor (%) (Note 5)		0	100		0.1
Noise level (dBmp)		–70	–40	± 0.3	0.1
Echo Loss (dB) Echo path loss (dB) Speech Echo path loss (dB) (Note 3)	A[a)]	6	25	± 0.3	0.1
	B[a)]	6	35	± 0.3	0.1
	C[a)]	6	45	± 0.3	0.1
Speech Echo path delay (ms) (Note 3)	A[a)]	0	50	± 0.3	0.1
	B[a)]	0	150	± 0.3	0.1
	C[a)]	0	1000	± 0.3	0.1

[a)] Refer to different classes of device (A : Local –national for many countries- networks, B : medium delays networks, C : long delay networks).

NOTES

1 Active speech levels greater than –15 dBm may cause saturation on digital connections, i.e. A-law or μ-law. The requirement for ± 0.3 accuracy may need revision if the reference measurement is not limited by a digital codec.

For measurements on an analogue connection, the analogue to digital converter of the INMD may cause saturation.

2 For any single measurement, the difference between the reference measurement shall not exceed ± 2 dB (or ± 2 ms).

3 The upper limits of the echo loss and echo path delay measurement ranges for classes A, B and C devices are consistent with each other, for in each case, from the point of view of the near end user, the talker echo loudness rating given by:

TELR = (EPL upper range limit + near side SLR + near side RLR).

TELR = (EPL upper range limit + 10)
corresponds to the smallest acceptable loudness ratings according to Figure 2/G.131 if the one-way transmission delay does not exceed the overall mean one-way transmission time given by:

OMOTT = (EPD upper range limit + near side mean one-way transmission time)
where the near side mean one way propagation time is great enough for the INMDs of the corresponding class to take into account any end user connection to the national network covered by the specification of this class.

4 A 0.1 resolution is required to meet the ±0.3 mean accuracy. For reporting measurements a greater resolution of 1 or 0.5 is acceptable.

5 Accuracy and resolution is for further study.

Table 1 : P.561 required parameters, ranges and accuracies

Proposals

P.561 needs to be updated and improved in many several of its aspects.

1) First, the physical measurement interface specifications (restricted to DS1 in the 1994 standard, in section 1.3 and 4.1) should be broadened to Ethernet-like connections (at least 100 MB), to make it possible to apply INMD techniques on IP network elements like IP switches, gateways or routers. The philosophy of the acquisition of the signal remains the same, since instead of high impedances probes connected to 2Mbits streams, the measurement devices can get the information from an Ethernet stream by a port mirroring (highly preferable to a serial connection between two network elements).
2) A definition must also be found for the so-called "class D", left for further study in the 1996 version of P.561, and initially intended for use on arbitrary, possibly non-linear and time-variant, networks which include processing devices such as LPC coders. Should it be finally rather devoted to the supervision of voice quality inside packet-based networks ? This solution seems very difficult to implement, since the interconnections between IP islands and PSTN islands are very common, and sometimes unknown by the users of INMDs. So how to make a clear difference between class C and class D devices ? Another solution is a separation based on the type of interface, the class D being dedicated to devices with Ethernet interfaces. An obvious advantage of this solution is that specific parameters, derived from the analysis of the IP/RTP/RTCP protocols, could then be allocated to this class of devices, allowing no fundamental change in the three other classes. If other transmission techniques appear in the future, new classes of devices will then be added.
3) Then, a new section of P.561 (maybe inside section 4) should specify how the audio signal has to be reconstructed after being extracted from an Ethernet stream. This is a very critical point since, in opposition with the classical classes A, B and C INMDs, the probes connected to an Ethernet stream do not access directly to the PCM signal as it will be delivered to the access network of the final user. Rules have to be given on how to implement inside the measurement devices (and not inside a reference circuit, since P.561 deals with the conformance of measurement devices, not of signal processing algorithms) the following features, which highly impact the reconstructed signal and the way it can be perceived by the users, and how to test that the devices on the market fulfil the conditions to be declared P.561-compliant (as it is already done in section 9 of this Recommendation for classes A, B and C INMDs) :
   - The voice decoder : a list of standard ITU codecs (at least G.711; G.729 and G.723.1) to be supported by INMDs (i.e. the implementation of the decoders inside the measurement device is fully compliant with these standards) must be given. For proprietary codecs or codecs standardised outside the ITU, P.561 cannot ask for a formal compliance but only quote an up to date list of references.
   - The de-jitter buffers : as long as no ITU-T Recommendation exists on this topic, a description of the features of a de-jitter buffer and how they should be implemented inside INMDs has to be added to P.561. The three major features of a buffer are its default depth, its maximal depth and its degree of dynamics (how it adapts to the quality of the network). We believe that the user should be left free to choose (and change if needed) the values or characteristics of all of these features. If a specific new recommendation is issued in the future to specify those features of the de-jitter buffers, it will be used as a reference in a next revision of P.561.

- The comfort noise generator : like for standard codecs, INMDs must be compliant with their annexes (for some codecs, this is even not an annex but an integral part of the standard) using voice activity detection. That means that the comfort noise generation schemes implemented inside the INMDs are the ones described in those annexes.
- The error concealment process : again, some ITU standard codecs contain packet loss concealment (PLC) features in their main bodies, and others only in annexes. INMDs should be able to implement them in a correct way, and independently according to the need of the user (for instance, a user should be able to choose whether the PLC should be used or not with the G.711 decoder).

4) New parameters can be added to the mandatory ones in P.561, with measurement ranges and required accuracy (specified in a revised table 1 or in a similar new table), as well as testing procedures for the compliance of measurement devices (as it is already the case in section 9 of P.561 for active speech level, psophometric noise level, echo path loss and echo path delay). They can also be added only to the list of optional parameters (and thus not require ranges, accuracies and testing procedures). All these new parameters can be either derived from the analysis of protocol (jitter, packet loss, etc.) or resulting from the most recent developments in signal processing (voice distortion, voice quality score, etc.).
5) Finally, another interesting field to investigate inside Q.16/12, although not related only to VoIP, is the use of INMDs to perform measurement at two points and to use a comparison of the measurement results at those points to derive parameters (i.e. : one-way transmission delay, exact jitter). It is not clear whether such a subject falls within the scope of P.561 (new parameters have to be defined) or of P.562 (the results depend on an analysis of single ended INMD measurements).

Hopefully, studies have been carried out in manufacturers' or network operators' laboratories on non-intrusive measurements for VoIP for about two years, and first results are now available (for instance in [4] and [5]). The evolution of VoIP itself is becoming slower, the technologies used are more stable, and the network operators are waiting for the network equipment manufacturers to bring them not only a good (or differentiated) quality of service, but accurate supervision tools. All of these factors and material make it possible now to envisage a realistic standardisation of the non-intrusive measurement techniques for voice quality in packet-based networks, as it is already the case for end-to-end measurement techniques since the standardisation of P.862. We believe that the goal of a revised P.561 Recommendation before the end of the year 2002 is reasonable and can be achieved, by simply implementing the proposals exposed above.

Rather than issuing new standards dedicated to VoIP from all the available data (a possibly unavoidable solution in some cases, like for a future single-ended PESQ-like model), it seems reasonable to envisage to enlarge the scope of existing standards. The proposals above apply for P.561 only, but a similar approach will be followed for P.562, leading to possible new sections or annexes describing models similar to the E-model or the Call Clarity Index (CCI) already included in P.562, but dedicated to VoIP scenarios.

References

[1] ITU-T P.561 Recommendation (1996) : In service Non-intrusive Measurement devices ; Voiceband services measurements

[2] ITU-T P.562 Recommendation (2000) : Analysis and interpretation of INMD voice-service measurements

[3] ANSI Standard T1M1-221 (1991)

[4] ITU-T Contribution COM 12-23 : Non-intrusive assessment of perceived voice quality impairments (France Télécom, 2000)

[5] ITU-TU Contribution COM12-D49 : Non intrusive monitoring of speech quality in voice over IP networks (Psytechnics, U.K., 2001)