ETSI TR 102 927 V1.1.1 (2010-11)

Technical Report

Speech and multimedia Transmission Quality (STQ); Packet Loss Concealment (PLC) performance measurement setup for home networks

Reference

DTR/STQ-00177

Keywords

analysis, quality, voice

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ).

1 Scope

The present document describes the principles of performance measurement of PLC mechanisms for Voice over IP, based on a simulation of mastered degradations (IP packet loss with growing importance), assessment of QoS (through MOS-LQO scores) at the output of PLC. It defines the measurement setup needed and describes the detailed analysis of its behaviour.

The present document presents a detailed and practical description of the corresponding test setup and procedure.

Annex A provides examples of results of measurements done according to the test procedure defined in the present document.

2 References

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

2.1 Normative references

The following referenced documents are necessary for the application of the present document.

Not applicable.

2.2 Informative references

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ITU-T Recommendation G.711: "Pulse code modulation (PCM) of voice frequencies" (with Appendix I).

[i.2] ITU-T Recommendation G.722: "7 kHz audio-coding within 64 kbit/s".

[i.3] ITU-T Recommendation G.729: "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)".

[i.4] ITU-T Recommendation P.862: "Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs". .

[i.5] ITU-T Recommendation P.862.1: "Mapping function for transforming P.862 raw result scores to MOS-LQO".

[i.6] ITU-T Recommendation P.862.2: "Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs".

[i.7] ITU-T Recommendation P.862.3: "Application guide for objective quality measurement based on Recommendations P.862, P.862.1 and P.862.2".

[i.8] ETSI EG 202 425: "Speech Processing, Transmission and Quality Aspects (STQ); Definition and implementation of VoIP reference point".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

home gateway: gateway between the Access Network (AN) and the Customer Premises Network (CPN)

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

µDSLAM	Micro-DSLAM
ADSL	Asymmetrical Digital Subscriber Line
AMR	Adaptive Multi-Rate codec
DECT	Digital Enhanced Cordless Telecommunications
DSL	See ADSL.
DSLAM	Digital Subscriber Line Access Multiplexer
DTMF	Dual Tome Multi Frequency
FTTH	Fibre to the home
FXS (port interface)	POTS interface of home Gateways
GSM	Global System for Mobile communication
HATS	Head and Torso simulator
IP	Internet Protocol
MOS-LQON	Mean Opinion Score - Listening Quality, objective, Narrow-band
PC	Personal computer
PLC	Packet Loss Concealment
POTS	Plain old telephone service
PSTN	Public Switched Telephone Network
SIP	Session Initiation Protocol
VoIP	Voice over Internet Protocol
WiFi	Wireless Fidelity

4 Application area

4.1 Codecs

In the present document, we focus on PLC mechanisms associated with G.711 [i.1] decoders only. In next revisions of the present document, the declination of this method for other codecs with optional PLC, like G.722 [i.2] and AMR codec family, will be presented.

Since PLC has been by default embedded in newer standardised decoders (like G.729 [i.3]), the performance of PLC mechanisms does not need to be assessed for these other codecs.

4.2 Devices

The present document addresses all the following types of Home network devices likely to implement PLC:

- DSL home gateways
- FTTH home gateways
- VoIP applications on PC or IP phones
- Wireless terminals (WiFi, Bluetooth, DECT) with embedded VoIP application

5 Test environments

The tests can be done in laboratory environment or in live situations. Both approaches are considered in the present document.

It may also be sometimes difficult to make the tests: e.g. some test laboratories may not have the relevant tools for initiating the connection, or on real networks the performance may not be guaranteed.

6 Detailed description of the test bed

Since PLC is meant to compensate for lost packets, the principle is to simulate such losses and to monitor the output signals.

For this purpose a dedicated IP degradation simulation tool is used. This simulation is performed on the IP flow and requires:

1) to be done as close as possible to the measurement point (in order to avoid IP packet losses between the simulation and the measurement points),

2) on an Ethernet link (or any similar physical support where IP packets can be directly processed).

It is also necessary to ensure, as far as possible, that the link on which tests are performed is clean, that is exempt of degradations likely to cause packet losses before the simulation is performed. This is generally the case in a laboratory environment, as well as in live network under unstressed conditions.

At the other end of the connection, a reference end-point is also necessary in order to send reference signals without extra impairments.

6.1 For DSL home gateways

On DSL architectures, it is difficult (and most of the time impossible) to perform test in an isolated test laboratory environment, since the gateway must be registered by a service plat-form to be allowed to set up calls.

The simulator has then to be inserted in series on the link between the phone plug on the wall inside home and the residential gateway. Since this link is not using Ethernet and does not allow access to the raw IP traffic, it requires having a derivation to do so, composed by an ADSL modem router and micro-DSLAM and an Ethernet link between them, as shown in Figure 1.

The output signal for measurement and analysis purposes are captured at the output of the FXS port of the home gateway under test.

In this case, the far-end reference point is a PSTN access, allowing a full G.711 [i.1] communication.

This set up has one drawback: it is based on the assumption that the IP CORE and DSL access networks are not causing any packet loss. Therefore an analysis of the contribution of these networks to the end-to-end packet loss rate is highly recommended in prior to the start of the test measurement procedure.

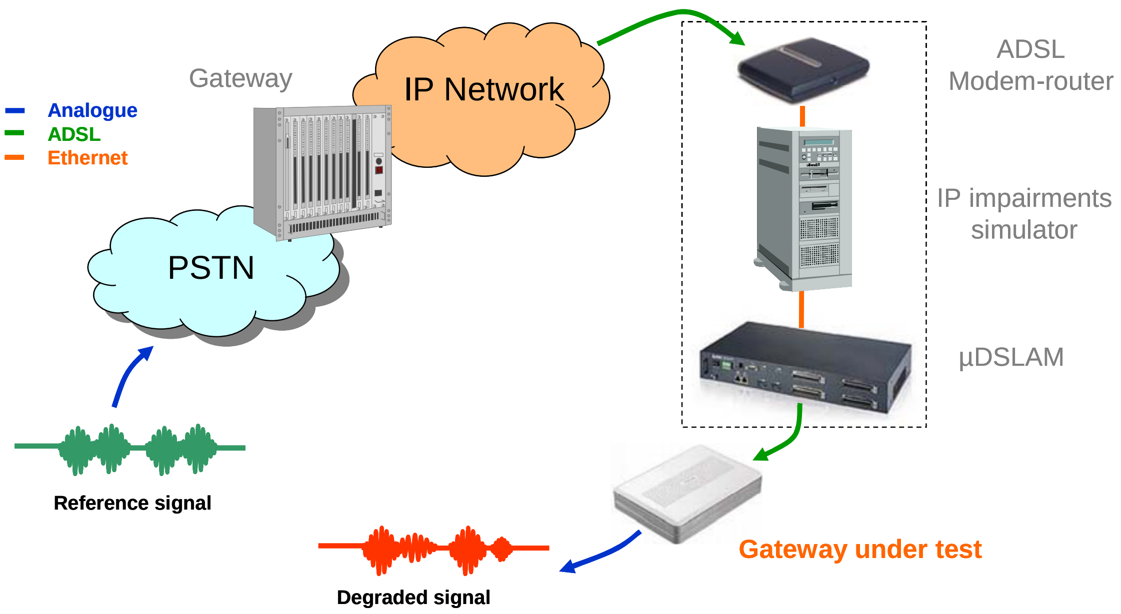

Figure 1: Test set up for ADSL home gateway

6.2 For FTTH home gateways

For further study. No test method is currently ready to be proposed.

6.3 For VoIP applications on PC and IP-phones

In fact, a distinction has to be done here between two situations:

- VoIP clients requiring a registration process on a commercial platform to be allowed to use the service
- Other VoIP clients

The first case is in fact similar to the one of home gateways, in the sense that it requires access to real networks, with possible resulting degradation of the IP flow. But the test set up is simpler, since the IP flow is directly accessible at the input of the PC or of the IP-phone. The extra modem-router and the µDSLAM are not necessary; the IP degradation simulation can be performed directly in series with the PC or the Ip-phone, as illustrated by Figure 2.

It should be checked that PC client on reference access is not creating unexpected impairment (e.g. packet jitter). Its performance should be in line with EG 202 425 (VoIP reference point) [i.8].

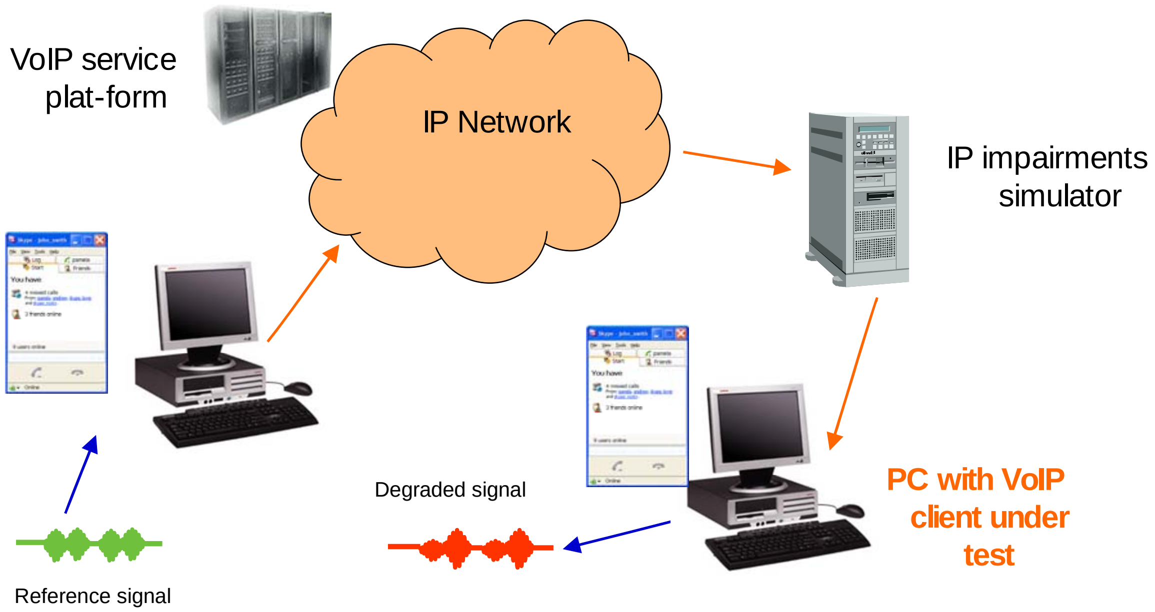

Figure 2: Test set up for VoIP on PC with registration on commercial platform

The second case is much simpler, and can be implemented in a fully controlled environment in laboratory. The test set up, as shown in Figure 3, is then simply composed by the PC or IP-phone under test, the IP simulator and an optional PC where the registration and authorisation functions are hosted (SIP proxy, for instance).

It should be checked that PC client on reference access is not creating unexpected impairment (e.g. packet jitter). Its performance should be in line with EG 202 425 (VoIP reference point) [i.8].

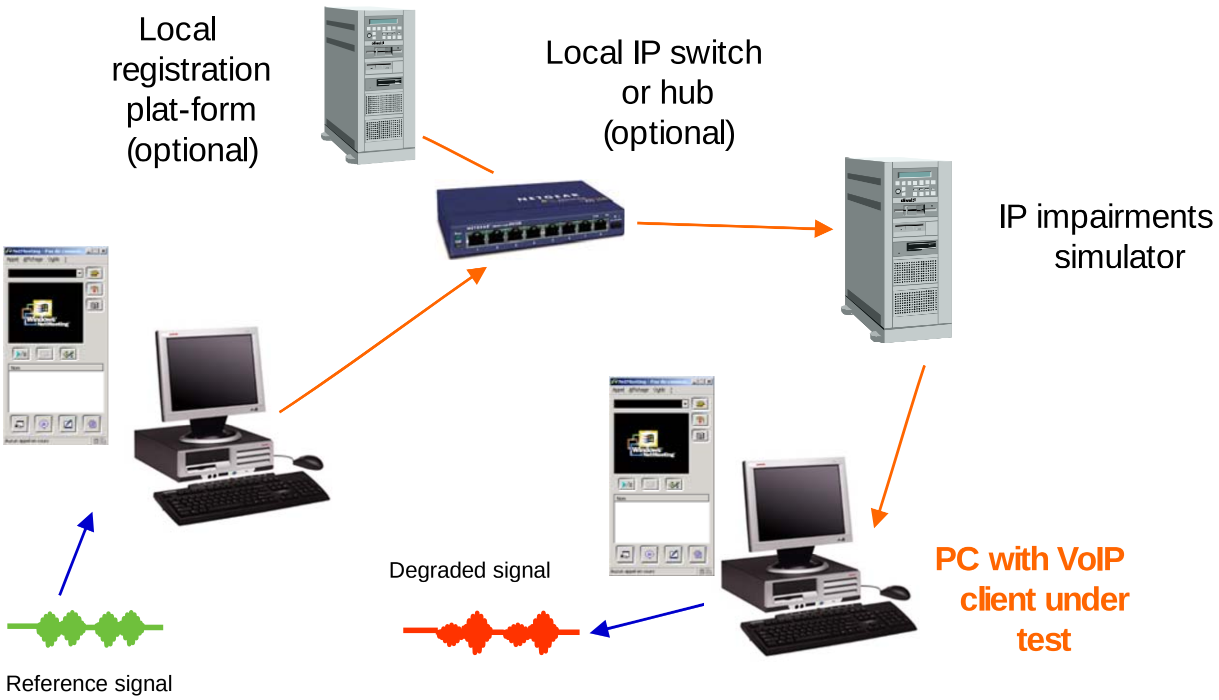

Figure 3: Test set up for VoIP on PC without registration on commercial platform

In both cases:

- the output signal for measurement and analysis purposes are captured at the output of the sound card of the PC under test or at an electrical output of the IP-phone (jack or handset plug),
- the far-end reference point is another PC or IP-phone embedding the same VoIP client with the same G.711 [i.1] codec.

6.4 For wireless terminals with embedded VoIP application

We are dealing here specifically with devices attached to the home network, that is their radio base station (whatever radio technology is used) is located inside home (i.e. it cannot be a WiFi hostpot around the next corner). The radio base station is therefore a peripheral unit of a central IP termination point. This central point can be any of the terminals presented in the cases above.

The output signal is captured at the corresponding electrical interface of the device, generally a headset jack plug. If the test facilities are also equipped with acoustical recording tools (HATS, etc.), it is also a possibility.

The far end reference points remains also the same.

Therefore the test set up for this case is similar to the one exposed above, with the only extension of the radio link. In Figure 4 we show the example with the DSL gateway.

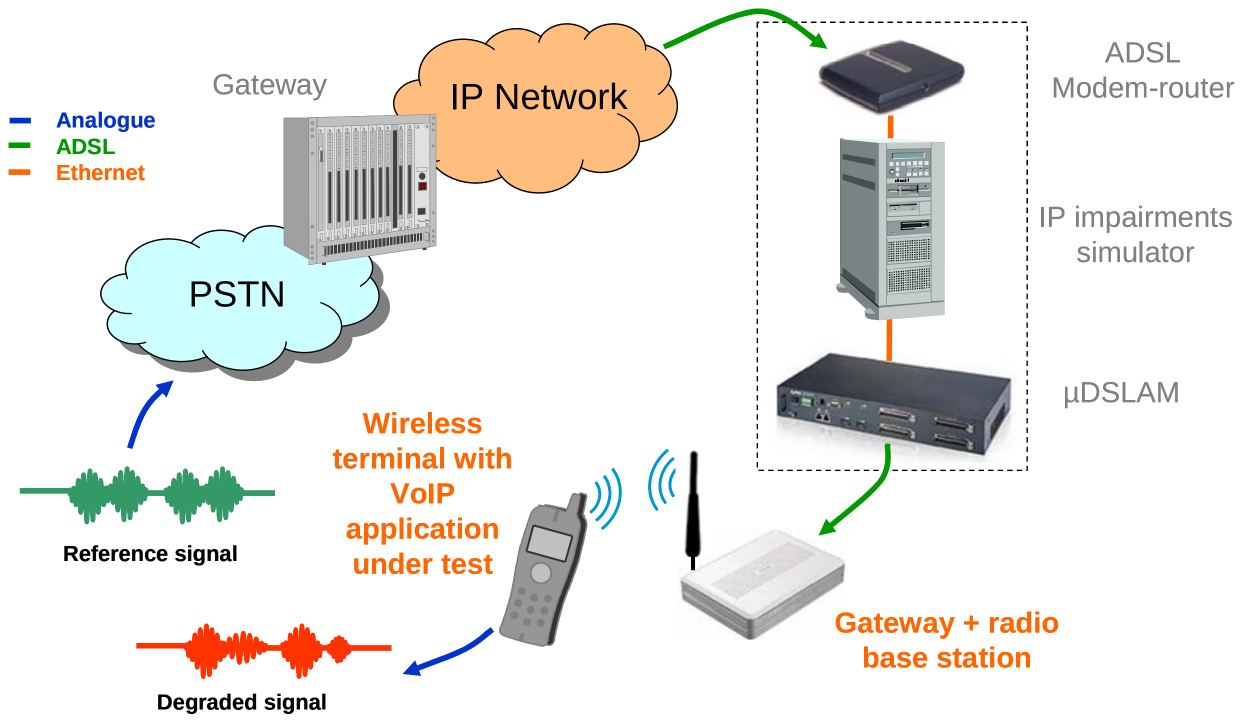

Figure 4: test set up for wireless terminal with embedded VoIP application and base station attached to an ADSL home gateway

7 Test procedure

The test procedure is based on the comparison of the device performance with two reference performance (one with reference PLC implementation and one without). These reference performance are made available in Table 1 and Figure 5, and can be used by the test lab as references.

7.1 PLC performance assessment (in terms of resulting voice quality)

Test scenarios are exactly the same for all cases exposed above, as follows:

1) a call is set up between the home network VoIP device under test and the reference point, using codec G.711 [i.1];

2) a test signal (voice sample compliant with ITU-T Recommendation P.862.3 [i.7]) is sent from reference to the home network VoIP device under test;

3) this signal is impaired by the IP impairment simulation at the receiving side;

4) the signal is recorded at the output of the home network installation under test, it can be:

   a) an FXS port for a home gateway (clauses 6.1 and 6.2);

   b) the electrical audio output (clause 6.3, jack plug for a DECT phone, sound card for a PC, etc.);

   c) the acoustical audio output for a wireless terminal (clause 6.4).

5) the received signal is recorded and compared to reference signal with the method defined in ITU-T Recommendation P.862.1 [i.5] to compute of a MOS-LQO score.

Different levels of degradation are simulated and tested (always with the same reference signals as defined in ITU-T Recommendation P.86x [i.4] to [i.7]) in order to see the performance of PLC with signal losses of growing size. We will only consider here the classical case of 20 ms packets. For other sizes, it is relatively easy to extrapolate proper values.

- no degradation
- 1 lost packet every second (2 % loss)
- 2 consecutive lost packets (i.e. 40 ms) every second (4 % loss)
- 3 consecutive lost packets (i.e. 60 ms) every second (6 % loss)
- and so on until at least 10 %

In each case, at least 10 measurements are performed in order to have a good and reliable statistic on MOS-LQO scores, since we are not sure if the lost packets are in active speech periods or in silence. This number can be reduced in very specific cases where it is possible in laboratory environment to make all losses of packets occur exactly during active speech periods.

It is possible to make a fully automatic test in which all impairment levels will be applied in series during the same call or during successive calls.

The obtained mean LQO scores are then compared to references computed off line with codec executables:

- low reference: G.711 [i.1] without PLC (lost audio frames replaced by silence)
- high reference : G.711 [i.1] with Appendix I PLC (active on up to 60 consecutive lost milliseconds)

Table 1 and Figure 5 give an example of typical mean values obtained in these two reference situations with test signals (according to ITU-T Recommendation P.862 [i.4]) in French languages (2 male and 2 female talkers). They are provided in the case when the tester has no access to an executable of G.711 [i.1] and its Appendix I.

Table 1: Example of reference results obtained according to ITU-T Recommendation P.862 [i.4]

Packet loss rate (%)	G.711 [i.1] with Appendix I PLC	G.711 [i.1] without PLC
0	4,5	4,5
2	3,9	3,5
4	3,5	2,7
6	3,3	2,3
8	2,8	2,1
10	2,3	1,9

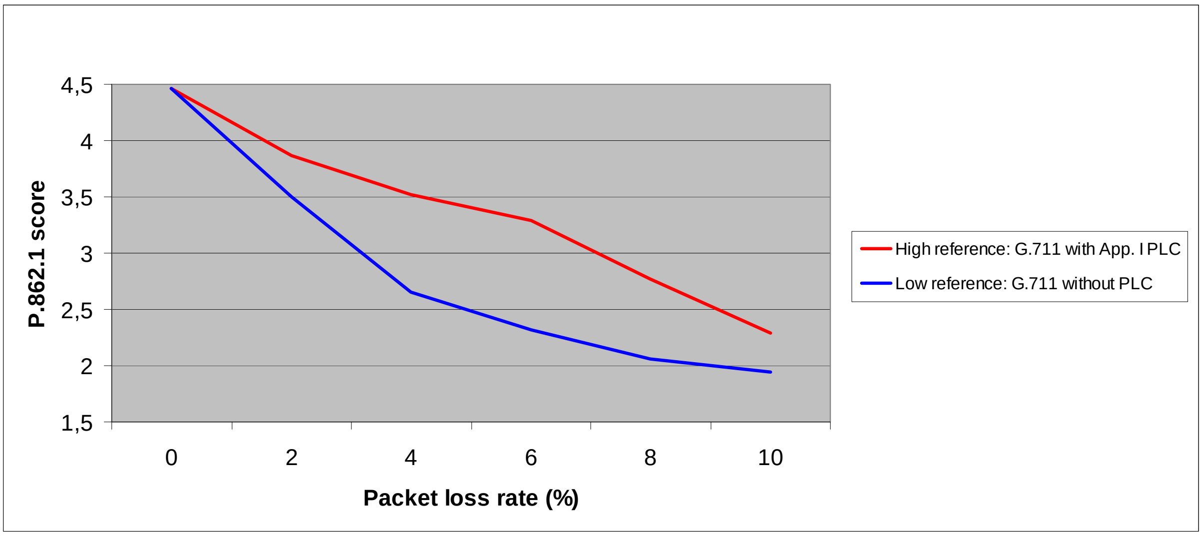

Figure 5: Example of reference results obtained according to ITU-T Rec. P.862 [i.4]

An example of simple acceptability (or pass/fail) criterion for a PLC mechanism under test can be a distance lower to the higher reference that to the lower reference for packet loss rates of 4 %, 6 % and 8 %.

To be considered as efficient, a PLC mechanism must have performance levels closer to the higher reference that to the lower reference for packet loss rates of 4 %, 6 % and 8 %. This means:

- at 4 %: MOS-LQON > 3,1
- at 6 %: MOS-LQON > 2,8
- ate 8 %: MOS-LQON > 2,4

7.2 PLC behaviour analysis

This second type of test is performed on the same test set up as for the voice quality measurement.

Test scenarios are exactly the same for all cases exposed above, as follows:

1) a call is set up between the home network VoIP device under test and the reference point;
2) a test signal (see below) is sent from reference to the home network VoIP device under test;
3) this signal is impaired by the IP impairment simulation at the receiving side;
4) the signal is recorded at the output of the home network installation under test, it can be:
   a) an FXS port for a home gateway (clauses 6.1 and 6.2);
   b) the electrical audio output (clause 6.3, jack plug for a DECT phone, sound card for a PC, etc.);
   c) the acoustical audio output for a wireless terminal (clause 6.4).
5) the received signal is recorded and analysed.

The impairment applied here is equivalent to the 4 % packet loss scenario of the voice quality measurements, that is 80 ms of signal lost every second. If the PLC mechanism under study is active beyond 60 ms, it may be useful to use a scenario with longer losses.

The test signal will last 1,5 second, so that we are sure that a loss of 80 ms will occur after at least 500 ms of signal.

This test signal must allow:

- a comparison of behaviours of different PLC mechanisms (what requires a stationary test signal, i.e. not voice)
- a good distinction between the different repetitions of the original frame taken into account by PLC (what requires a rather complex signal, i.e. not a pure sinusoid)
- an application for other codecs with optional PLC (GSM, AMR, etc.), and causing possible distortion (therefore, DTMF could not be the solution)

Studies are going on to find the most suitable test signal. For the time being it is recommended to use a 440 Hz sinewave.

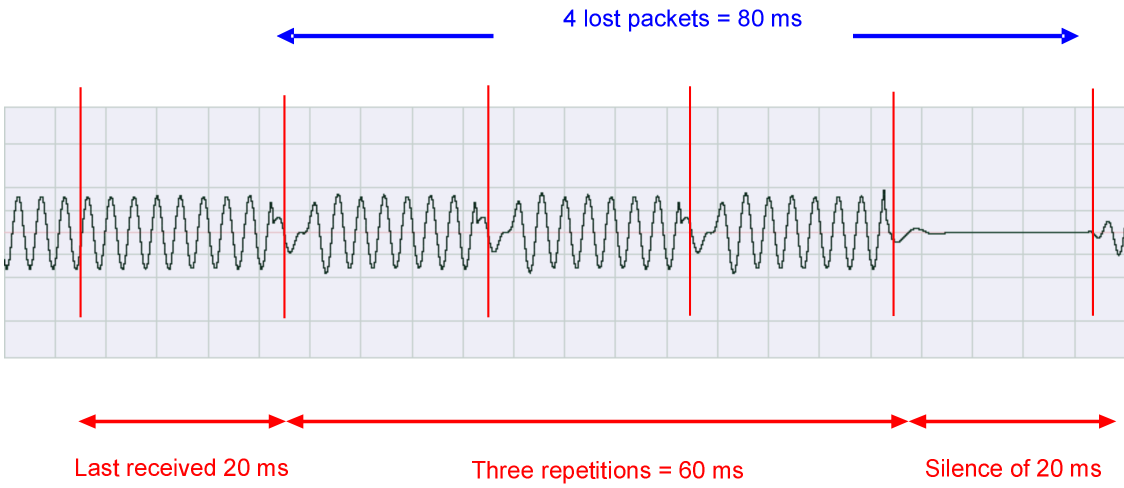

Figure 6: Example of PLC behaviour testing with a 440 Hz sinewave

Here, the result of PLC, as well as the periodicity of its behaviour, is very clear and easy to read. But it will not be necessarily always the case.

In the mean time, the analysis can be performed on the voice test signal used in the previous type of test (according to ITU-T Recommendation P.862 [i.4]), as in the examples shown in Annex A.

On the received signal, the following actions will be undertaken:

- detection of a hole or a cut, corresponding to a packet loss event
- selection of the 80 ms before the end of this cut
- comparison of the content of these 80 ms with the signal in the 20 preceding milliseconds:
  - repetition (size of the copied signal piece, count of the number of repetitions)
  - attenuation (measured in dB)

Annex A:
Examples of test results

In France Telecom residential offers, PLC can be implemented in two places:

- inside the home gateway (called ("Livebox"), for calls on the FXS port using G.711 [i.1]
- inside a wireless terminal (called "Livephone") connected through DECT to the Livebox, for calls in G.711 [i.1] or G.722 [i.2]

This contribution presents only the example of measurement on gateways (G.711 [i.1] on FXS port).

Two types of results/monitoring may be obtained.

A. MOS-LQO value versus packet loss

Figure A.1 shows what we obtained during a benchmark between several models of home gateways for the quantitative part of our tests.

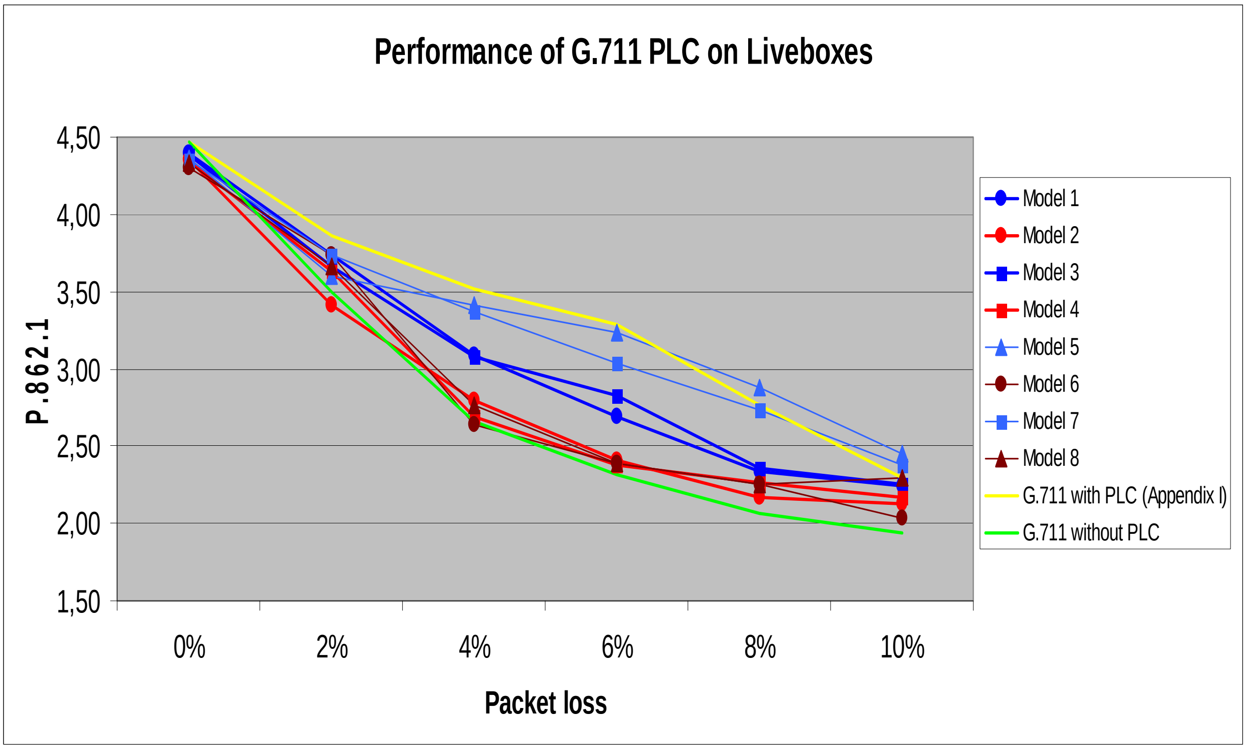

Figure A.1: Example of test results (according to ITU-T Recommendation P.862 [i.4])

It is important to compare measurement results with references. This is why we have the curves in yellow (high reference or target: performance level of G.711 [i.1] appendix 1) and in green (no PLC).

When focusing on the most realistic and important zone (around 4 % and 6 %), one can see three categories:

- in blue : OK, close to the target
- in dark blue: PLC present but not as good as reference
- in red : no PLC at all

Based on these results it could be possible to define a requirement based on the reference curves obtained for G.711 [i.1] with and without PLC.

Such a process could apply in same way for G.722 [i.2].

B. Monitoring of the signals

Figures A.2 to A.4 show the result of the qualitative analysis on three representative gateways for each of these categories, on the case with 4 % packet loss:

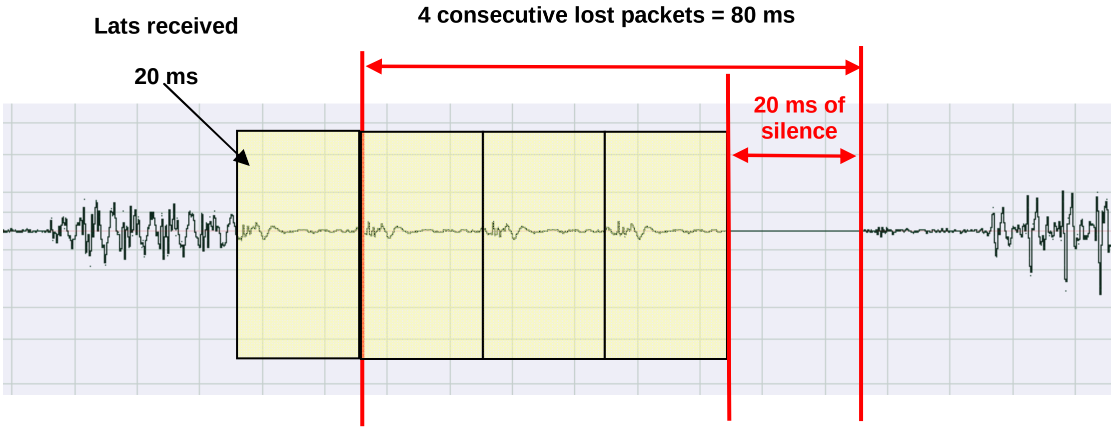

Figure A.2: Example of PLC behaviour equivalent to Appendix I performance

Here the, last 20 ms are repeated until three times. Then the missing signal is replaced by silence.

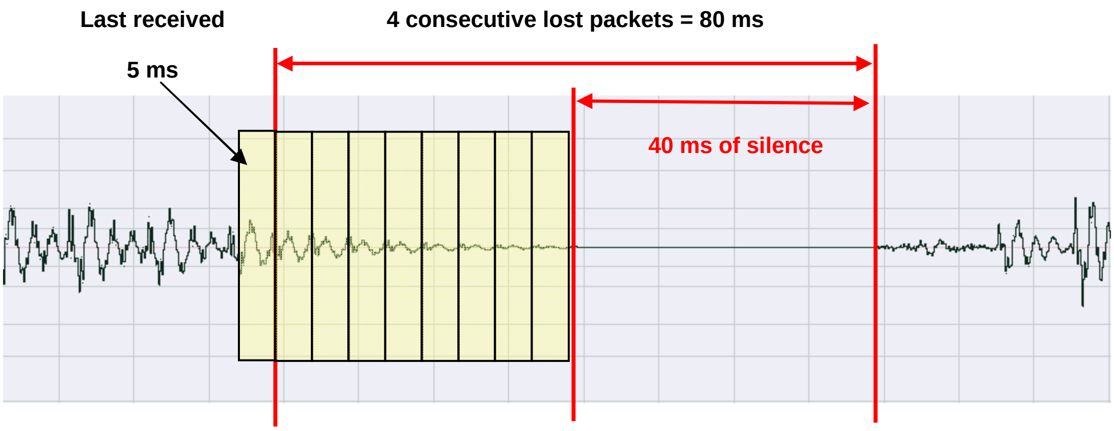

Figure A.3: Example of PLC behaviour worse than Appendix I performance

Here the, last 5 ms are repeated until eight times, with a progressive attenuation. Then the missing signal is replaced by silence.

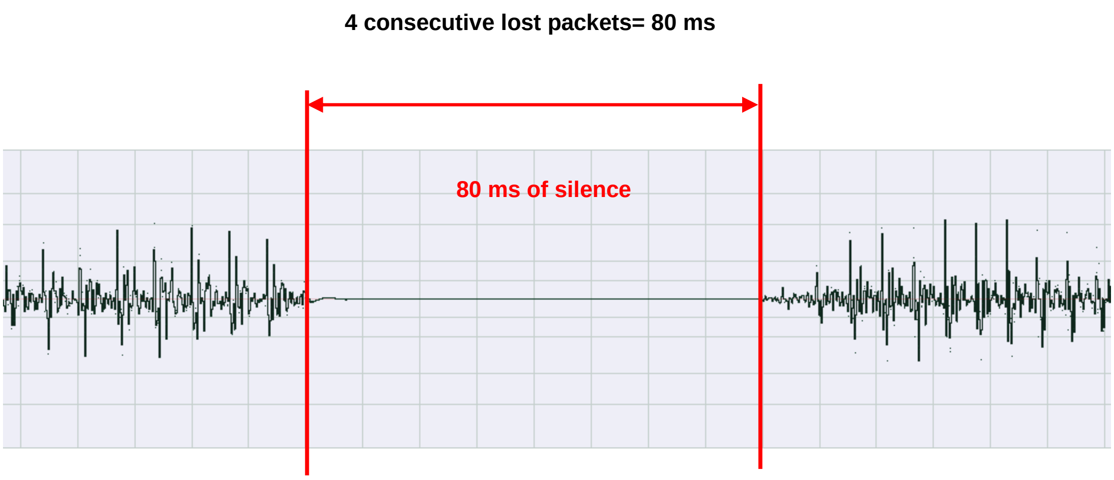

Figure A.4: Example with no PLC at all

It is important to have such a qualitative analysis besides MOS-LQO scores. Indeed, the objective of such tests is not only to assess the performance by itself, but also (and mostly!) to understand why signal processing may not perform well and be able to propose enhancements.

History

Document history		
V1.1.1	November 2010	Publication