

**Telecommunications and Internet Protocol
Harmonization Over Networks (TIPHON) Release 4;
End-to-end Quality of Service in TIPHON Systems;
Part 5: Quality of Service (QoS) measurement methodologies**



Reference

DTS/TIPHON-05008R4

Keywords

IP, methodology, quality, service, telephony,
QoS, 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, send your comment to:

editor@etsi.org

Copyright Notification

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2003.
All rights reserved.

DECTTM, **PLUGTESTS**TM and **UMTS**TM are Trade Marks of ETSI registered for the benefit of its Members.
TIPHONTM and the **TIPHON logo** are Trade Marks currently being registered by ETSI for the benefit of its Members.
3GPPTM is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.

Contents

Intellectual Property Rights	5
Foreword.....	5
1 Scope	6
2 References	6
3 Definitions, symbols and abbreviations	8
3.1 Definitions	8
3.2 Symbols.....	8
3.3 Abbreviations	8
4 Test Set-up for Terminals and Systems Including Terminals	9
5 Call Establishment Measurements	9
5.1 Start Dial Signal Delay	9
5.2 Post Dial Delay.....	9
5.3 Call Duration	10
5.4 Release on Request.....	10
6 Speech Quality Measurements	10
6.1 Subjective Speech Quality.....	10
6.2 Objective Speech Quality.....	11
6.3 Advanced Objective Speech Quality Parameters	11
6.4 Mean One Way Delay	12
6.5 Echo Path Loss	12
6.6 Loudness Ratings	12
6.7 Overall Transmission Quality Rating [R].....	13
7 Transport layer measurements.....	13
7.1 One way transmission time	13
7.2 Roundtrip transmission time.....	13
7.3 2 Point packet delay variation	14
7.4 1 Point packet delay variation	14
7.5 Network packet loss	14
7.6 Effective packet loss.....	14
7.7 Packet errors.....	15
7.8 Mis-sequenced packets.....	15
7.9 Voice client induced PDV	15
7.10 Packet loss correlation.....	15
8 QoS mechanism tests.....	16
8.1 Simulated media for QoS calibration	16
8.2 Passive media path monitoring for QoS	16
Annex A (normative): Call establishment measurements.....	17
Annex B (normative): Speech quality measurements	18
B.1 Delay measurement	18
B.2 Loudness rating	19
Annex C (normative): QoS mechanism tests.....	20
C.1 Simulated media for QoS calibration	20
Annex D (normative): Testing procedures for TIPHON terminals and systems.....	23
D.1 Introduction	23
D.2 Parameters, defining speech transmission quality including terminals.....	23

D.3	Test set-up for terminals and systems including terminals	24
D.3.1	Test Signals	24
D.4	Measurement of standard parameters	24
D.5	Advanced Measurements, Taking Into Account the Conversational Situation.....	25
D.5.1	Measurement set-up for objective tests	27
D.5.2	Practical realization of test signals	27
D.5.3	Test procedures	27
Annex E (informative): Method for determining an equipment impairment factor using passive monitoring		28
E.1	Introduction	28
E.2	Passive QoS Monitor Framework	28
E.3	Determining equipment impairment factor for packet loss	29
E.4	Determining Equipment Impairment Factor for Packet Delay Variation.....	31
E.5	Measuring Delay	32
E.6	Determining Equipment Impairment Factor for CODEC	32
E.7	Determining Overall Equipment Impairment Factor	32
E.7.1	Determining Average Equipment Impairment Factor	32
E.7.2	Recency Effect	33
E.8	Use of the E Model.....	33
Annex F (informative): Bibliography		34
History		35

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 Specification (TS) has been produced by ETSI Project Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON).

The present document is part 5 of a multi-part deliverable covering end-to-end Quality of Service in TIPHON systems, as identified below:

- TR 102 024-1: "General aspects of Quality of Service (QoS)";
- TS 102 024-2: "Definition of Speech Quality of Service (QoS) Classes";
- TS 102 024-3: "Signalling and Control of end-to-end Quality of Service";
- TS 102 024-4: "Quality of Service Management"
- TS 102 024-5: "Quality of Service (QoS) measurement methodologies";**
- TR 102 024-6: "Actual measurements of network and terminal characteristics and performance parameters in TIPHON networks and their influence on voice quality";
- TR 102 024-7: "Design guide for elements of a TIPHON connection from an end-to-end speech transmission performance point of view".
- Part 8: Void.
- TS 102 024-9: "Call performance Classification (Voice)".
- Part 10: To be published as part of release 5 (see note) "QoS Requirements for TIPHON Terminals";
- Part 11: To be published as part of release 5 (see note) "Domain by domain performance planning guidelines for end-to-end QoS objectives associated with TIPHON speech QoS classes".
- TS 102 024-12: "IP Telephony Service Availability".

NOTE: Quality of Service aspects of TIPHON Release 5 systems will be covered in TS 102 025 (see bibliography), and more comprehensive versions of the Release 4 documents listed above will be published as part of Release 5 as work progresses.

1 Scope

The present document applies to IP networks that provide voice telephony in accordance with any of the TIPHON scenarios.

It contains:

- test methodologies for end to end QoS parameters;
- test methodologies for network performance parameters.

It should be noted that the work has tried to reference already developed measurement techniques rather than defining new techniques unnecessarily.

Background information and discussions are contained in the General Aspects of QoS document TR 102 024-1 [1].

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication and/or edition number or version number) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.

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

- [1] ETSI TR 102 024-1: "Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 4; End to End Quality of Service in TIPHON Systems; Part 1: General aspects of Quality of Service (QoS)".
- [2] ITU-T Recommendation P.800: "Methods for subjective determination of transmission quality".
- [3] ITU-T Recommendation G.131: "Control of talker echo".
- [4] Draft ITU-T Recommendation P.861: "Perceptual Evaluation of Speech Quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs".
- [5] ITU-T Recommendation G.169: "Automatic level control devices".
- [6] ITU-T Recommendation P.340: "Transmission characteristics of hands-free telephones".
- [7] ITU-T Recommendation P.76: "Determination of loudness ratings; fundamental principles".
- [8] ITU-T Recommendation P.79: "Calculation of Loudness Ratings for telephone sets".
- [9] ITU-T Recommendation G.107: "The E-Model, a computational model for use in transmission planning".
- [10] ITU-T Recommendation G.108: "Application of the E-model: A planning guide".
- [11] ITU-T Recommendation G.177: "Transmission planning for voice-band services over hybrid Internet/PSTN connections".
- [12] ITU-T Recommendation P.59: "Artificial conversational speech".
- [13] ITU-T Recommendation P.501: "Test signals for use in telephony".

- [14] ITU-T Recommendation P.502: "Objective test methods for speech communication systems, using complex test signals".
- [15] ITU-T Recommendation P.581: "Use of head and torso simulator (HATS) for hands-free terminal testing".
- [16] ITU-T Recommendation P.831: "Subjective performance evaluation of network echo cancellers".
- [17] ITU-T Recommendation P.832: "Subjective performance evaluation of Hands-free Terminals".
- [18] ITU-T Recommendation P.51: "Artificial mouths".
- [19] ITU-T Recommendation P.57: "Artificial ears".
- [20] ITU-T Recommendation P.58: "Head and torso simulator for telephonometry".
- [21] ITU-T Recommendation P.64: "Determination of sensitivity/frequency characteristics of local telephone systems".
- [22] ITU-T Recommendation P.50: "Artificial Voices".
- [23] ETSI TBR 8: "Integrated Services Digital Network (ISDN); Telephony 3,1 kHz teleservice; Attachment requirements for handset terminals".
- [24] ITU-T Contribution COM12-42 Federal Republic of Germany: "Listening only test results for hands-free telephones and their dependence upon room surroundings (1997)".
- [25] ITU-T Recommendation G.711: "Pulse code modulation (PCM) of voice frequencies".
- [26] ITU-T Recommendation G.726: "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)".
- [27] ITU-T Recommendation G.113 Appendix 1: "Transmission Impairments - Appendix I: Provisional planning values for the equipment impairment factor I_e ".
- [28] ITU-T Recommendation G.114: "One-way transmission time".
- [29] ITU-T Contribution COM12-D139 France Télécom R&D (Q14/12): "Study of the relationship between instantaneous and overall subjective speech quality for time-varying quality speech sequences: influence of a recency effect (Delayed Contributions 9-18 May 2000)".
- [30] ANSI T1A1.7/98-031: "Testing the quality of connections having time varying impairments".
- [31] ITU-T Recommendations P.831: "Subjective performance evaluation of network echo cancellers".
- [32] ITU-T Recommendations P.501: "Test signals for use in telephonometry".
- [33] ITU-T Recommendations G.108: "Application of the E-model: A planning guide".

3 Definitions, symbols and abbreviations

3.1 Definitions

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

codec: combined speech encoder and decoder

3.2 Symbols

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

ms	milliseconds
s	seconds

3.3 Abbreviations

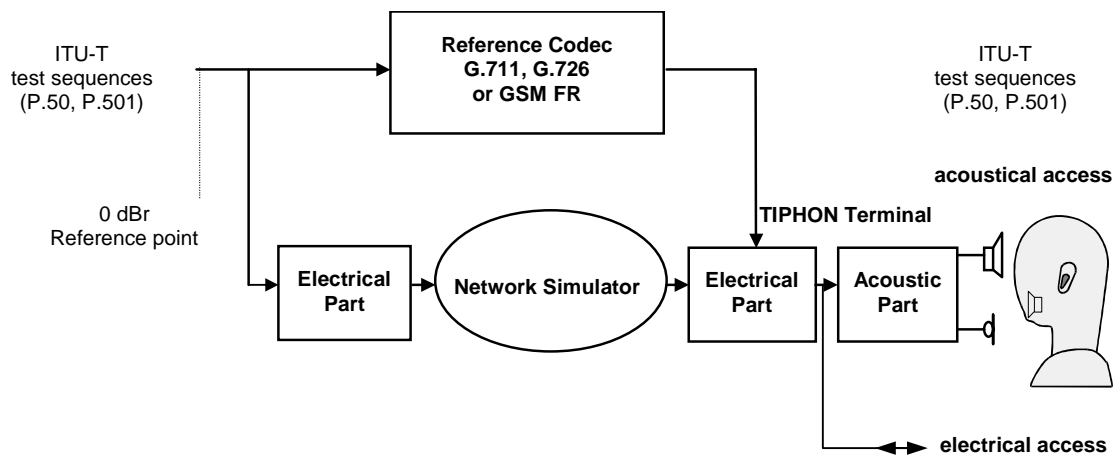
For the purposes of the present document, the following abbreviations apply:

CRC	Cyclic Redundancy Check
DTX	Discontinuous Transmission
FFT	Fast Fourier Transform
GSM FR	Global System for Mobile, Full Rate codec
IP	Internet Protocol
LR	Loudness Ratings
LSTR	Listener Sidetone Rating
MNRU	Modulated Noise Reference Unit
MOS	Mean Opinion Score
Nc	Circuit Noise referred to the 0 dBr-point
OLR	Overall Loudness Rating
PDD	Post Dial Delay
PDV	Packet Delay Variation
qdu	Number of Quantizing Distortion Units
QoS	Quality of Service
RLR	Receive Loudness Rating
SCN	Switched Circuit Network
SDSD	Start Dial Signal Delay
SLR	Send Loudness Rating
STMTR	Sidetone Masking Rating
TCLw	Terminal Coupling Loss (weighted)
TELRL	Talker Echo Loudness Rating
UDP	User Datagram Protocol
VoIP	Voice over Internet Protocol
WEPL	Weighted Echo Path Loss

4 Test Set-up for Terminals and Systems Including Terminals

The general access to terminals is described in figure 2. The preferred way of testing is the connection of the terminal to a network simulator or a complete network. When testing without acoustical access, the test sequences can be fed to the electrical interface as indicated in figure 2. The test sequences are fed in either electrically, using a reference codec or using the direct signal processing approach or acoustically using ITU-T specified devices such as artificial ear and mouth according to the Recommendations P.51 [18], P.57 [19] and P.58 [20]. The positioning and set-up for handset type telephones is described in ITU-T Recommendation P.64 [21], for hands-free type telephones the set-up is described in ITU-T Recommendation P.581 [15]. The test set-up can be used on both sides of a connection if complete configurations are tested.

Figure 1: Void



NOTE: Packet loss distribution is for further study.

Figure 2: Methodology for testing TIPHON Terminal/Systems Speech Quality

5 Call Establishment Measurements

5.1 Start Dial Signal Delay

Definition

Time in milliseconds for the dial tone to be audible after the phone is placed off-hook from the idle state.

Test Metrics

- start dial signal delay/ms;
- percentage of calls with no dial tone.

Comments

None.

5.2 Post Dial Delay

Definition

Time in milliseconds between dialling the last digit and an audible tone being heard at the originating end. The audible tone is typically ring-back or the engaged tone.

Test Metrics

- post dial delay/ms.

Comments

Some systems have shown to present the user with a ring-back tone before a connection has been established, this gives the impression that the PDD is low. If the connection fails this is later switched to an engaged tone. This is unacceptable operation and should be tested.

5.3 Call Duration

Definition

The time in seconds between bi-directional media path establishment and media path closure at both ends of the connection.

Test Metrics

- Call duration/s (accurate to 1 ms);
- Percentage of premature releases.

Comments

The call duration information can be used to check billing system accuracy.

5.4 Release on Request

Definition

Check to identify that connection is released when placing phone on-hook.

Test Metrics

- Percentage of correctly terminated calls.

Comments

None.

6 Speech Quality Measurements

6.1 Subjective Speech Quality

Definition

A subjective quality measure, or Mean Opinion Score (MOS), is determined from performing a subjective test in accordance with P.800 [2]. A MOS is an average opinion of quality for a system based on asking people their opinion of quality under control conditions. Further evaluation procedures specifically for echo canceller and hands-free terminal testing can be found in ITU-T Recommendations P.831 [16] and P.832 [17].

Test Metrics

- Listening quality absolute category rating (P.800 [2], annex B);
- Listening distortion category rating (P.800 [2], annex D).

Comments

A subjective test, for TIPHON QoS class classification, should include the following reference conditions:

- Clean speech;
- G.711 [25] with no additional distortions;
- G.726 [26] at 32 kbit/s with no additional distortions;
- GSM FR with no additional distortions;
- MNRU conditions (Q = 6, 12, 24 and 30).

6.2 Objective Speech Quality

Definition

A measure of speech quality by a computer based software program. The ETSI TIPHON project endorses the ITU-T Recommendation P.861 [4].

Test Metrics

- Speech Quality MOS prediction (ITU-T Recommendation P.861 [4]).

Comments

It is paramount to use an appropriate speech or speech-like test signal. The signal should be at least 8 to 10 seconds in duration to ensure both system stability and the opportunity for errors to be assessed.

6.3 Advanced Objective Speech Quality Parameters

Definition

Various measures of speech quality parameters based on test signals and procedures are described in ITU-T Recommendations P.501 [13], P.502 [14] and P.340 [6].

Test Metrics

- Convergence parameters of echo cancellers (section 4 of ITU-T Recommendation P.502 [14]);
- Speech quality parameters during double talk (section 5 of ITU-T Recommendation P.502 [14]);
- Companding and AGC characteristics (section 6 of ITU-T Recommendation P.502 [14]);
- Quality of background noise transmission (section 7 of ITU-T Recommendation P.502 [14]);
- Switching parameters (section 8 of ITU-T Recommendation P.502 [14]).

Comments

The tests require various (speech-like) test signals as described in ITU-T Recommendation P.501 [13] and ITU-T Recommendation P.50 [22]. The test duration should be at least 8 seconds to 10 seconds in duration to ensure both system stability and the opportunity for errors to be assessed. In addition to the test method described in clause 6.2 individual parameters influencing the speech quality can be assessed, in single and double talk situations. The test methodology allows the assessment of terminals as well as of network components and configurations.

Further details can be found in annex D.

6.4 Mean One Way Delay

Definition

Mean one way delay is the time taken in milliseconds for a test signal to go from the near-end voice test point, traverse the network, get looped back at the far voice test point and arrive back at the near voice test point divided by two.

Test Metric

- Mean one-way delay/ms;
- Average of 10 delay measures or 90 % of largest delay (whichever is greatest).

Comments

VoIP systems exhibit bulk delay variations and therefore a number of delay measures should be made to have a statistical average. Delays should be measured over 30 seconds.

One methodology to measure delay is described in annex E.

6.5 Echo Path Loss

Definition

The ratio of r.m.s values of the incident to reflected speech signals with the echo path delay removed.

Test Metrics

- Steady state residual and returned echo level (ITU-T Recommendation G.169 [5], test 1).

Comments

The electrical performance of echo cancellers in a TIPHON system should conform to ITU-T Recommendation G.169 [5], guidelines for measuring acoustical echo are given in section 10 of ITU-T Recommendation P.340 [6].

6.6 Loudness Ratings

Definition

A Loudness Rating (LR) is a single-figure weighted average of the frequency-dependent loss between two reference points.

Test Metrics

- SLR (ITU-T Recommendation P.76 [7] and P.79 [8]);
- RLR (ITU-T Recommendation P.76 [7] and P.79 [8]);
- OLR (ITU-T Recommendation P.76 [7] and P.79 [8]).

Comments

LR calculations are traditionally performed using sine waves placed at 1/3 octave centre frequencies. However, when assessing complex non-linear systems there is a need to use a speech-like test signals to pass through low-bit rate codecs. The LR calculations are performed on the speech-like signal by calculating Fast Fourier Transform based 1/3-octave parameters.

6.7 Overall Transmission Quality Rating [R]

Definition

The "Overall Transmission Quality Rating [R]" is the output of the e-model, a planning tool, which relates aspects of telephony transmission performance to a single figure R. R is representative of a users perceived conversational performance of a system. The E-model is described in ITU-T Recommendation G.107 [9] and guidance can be found in ITU-T Recommendations G.108 [10] and ITU-T Recommendation G.177 [11].

Test Metrics

- R-value.

Comments

A default telephone-handset profile is used for TIPHON classification. This profile is based on a "traditional" telephone handset by using the default values for e-model calculations. Acoustic characteristics of TIPHON terminals are not considered in order to focus on the parameters specific to TIPHON network related issues (i.e. where TIPHON networks differ from existing SCN networks).

One methodology to passively monitor the overall transmission quality is described in annex E.

7 Transport layer measurements

7.1 One way transmission time

Definition

Time in milliseconds between the emission of a signal and the time it is received, includes delays due to equipment processing as well as propagation delay.

Test Metrics

- Mean packet transmission time/ms;
- Minimum and maximum packet transmission times/ms.

Comments

Measurement requires two synchronized test boxes.

7.2 Roundtrip transmission time

Definition

Time in milliseconds for a packet to be transmitted from host A and received at host B and to be re-transmitted from host B and received back at host A.

Test Metrics

- Mean roundtrip packet transmission time/ms;
- Minimum and maximum packet transmission times/ms.

Comments

The reflection of a packet for roundtrip measurement should be at the protocol layer that the measurement is addressing.

7.3 2 Point packet delay variation

Definition

PDV is the difference between upper and lower percentiles on the packet delay distribution. 2pt PDV uses 2 monitoring points. The measurement uses the difference between the inter-packet sending and inter-packet arrival times.

Test Metrics

- 2pt packet delay variation/ms.

Comments

Measurement requires two synchronized test boxes.

7.4 1 Point packet delay variation

Definition

PDV is the difference between upper and lower percentiles on the packet delay distribution. 1pt PDV uses only 1 monitoring point. The measurement is based on the inter-packet arrival times.

Test Metrics

- 1pt packet delay variation/ms.

Comments

Measurement requires a single test box and therefore no synchronization. This measure gives a clear illustration of the end-systems view of PDV but cannot be used so easily to quantify where any PDV has occurred.

7.5 Network packet loss

Definition

Percentage of packets lost at an IP test point; this metric does not include any losses due to the end-terminal equipment.

Test Metric

- Percentage network packet loss;
- Total number of lost packets.

Comments

None.

7.6 Effective packet loss

Definition

Percentage of packets lost as measured at the input of the speech codec, affecting the speech coder performance.

Test Metric

- Percentage network packet loss;
- Total number of lost packets;
- Packet loss distribution.

Comments

None.

One methodology to measure effective packet loss is described in annex E.

7.7 Packet errors

Definition

Packets that fail the CRC when received at an IP test point.

Test Metric

- Percentage of errored packets;
- Total number of errored packets.

Comments

Errors in a data packet will normally result in a packet being dropped by the layer 2 protocol which have checksums for the whole packet. However CRC can sometimes fail and this can be monitored using the test tools available.

7.8 Mis-sequenced packets

Definition

Out of sequence packets at the receiving IP test point.

Test Metrics

- Number of mis-sequenced packets.

Comments

A large number of mis-sequenced packets may indicate a congested network or that load balancing is in use.

7.9 Voice client induced PDV

Definition

A measure of the inter-packet delay variations as packets are transmitted onto the network by a voice client.

Test Metrics

- Client transmit PDV/ms.

Comments

Client Induced PDV is a significant contributory factor in the total delay variation experienced on a connection.

7.10 Packet loss correlation

Definition

A description of the "burstiness" of packet losses at a test point.

Test Metrics

- Average number of successive lost packets;
- Distribution of burst loss lengths;

- Markov loss model (as described in annex E).

Comments

None.

8 QoS mechanism tests

8.1 Simulated media for QoS calibration

Definition

A simulated media stream is used to determine a network's ability to deliver a required QoS level.

Test Metrics

- Delay variation;
- Packet loss;
- Packet loss correlation;
- Packet delay.

Comments

The use of a simulated media stream ensures that the QoS mechanism is fully tested for it is in-service use.

8.2 Passive media path monitoring for QoS

Definition

A non-intrusive monitoring of media paths to determine customers QoS.

Test Metrics

- Packet loss;
- Packet loss correlation;
- Delay variation.

Comments

None.

One methodology to estimate effective packet loss and the overall transmission rating is described in annex E.

Annex A (normative): Call establishment measurements

This annex provides a description of how to calculate the call establishment measurements in described in clause 5.

From a TIPHON QoS perspective call set-up measures are generally time related, although there are other equally important measures such as the number of correctly connected calls. With a TIPHON system, there are two different scenarios to consider:

- The traditional telephone user;
- The TIPHON terminal user.

Figure A.1 shows the call set-up sequence used by general telephony services. However, with TIPHON terminal equipment, such as PC clients, there is likely to be no off-hook, dial tone sequence. For this situation the act of a user pressing the "connect button" is regarded as step C - last digit dialled.

NOTE: Connect button describes the process by which a user instigates a call.

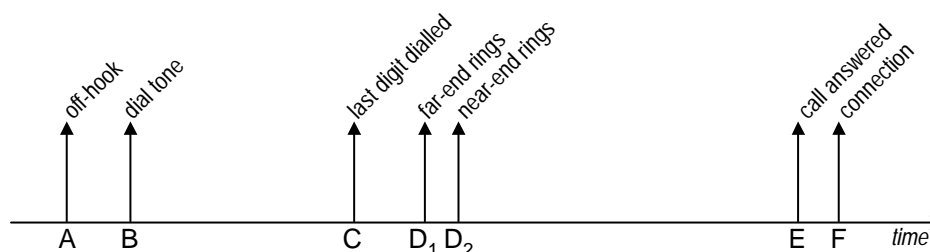


Figure A.1: Call set-up sequence

From the users perspective the significant time sequences are:

Start dial signal delay (SDSD): B - A

Post dial delay (PDD): D₂ - C

NOTE 1: Traditionally, call set-up progression is signified by audible tones, this is now being supplemented or replaced by text based messages.

NOTE 2: It is worth noting that there should not be a significant delay between D₁ and D₂, delays can result in the receiving party answering the phone before the calling party is aware a connection has been made. Similarly it is inappropriate to allow D₂ to occur before the far-end connection has been identified as being accessible; either in terms that it exists or in that it is not engaged.

Measurement of these set-up times shall be based on the progression mechanisms presented to the user, and not lower level signalling information.

Annex B (normative): Speech quality measurements

B.1 Delay measurement

A single assessment of a TIPHON system's delay is inadequate. For VoIP systems, it is important to determine a statistical average of the delay.

It is proposed that TIPHON use the mean delay from at least 10 measurements or 90 % of the largest delay measure, whichever is greatest.

The delay measurement test signal is illustrated in figure B.1.

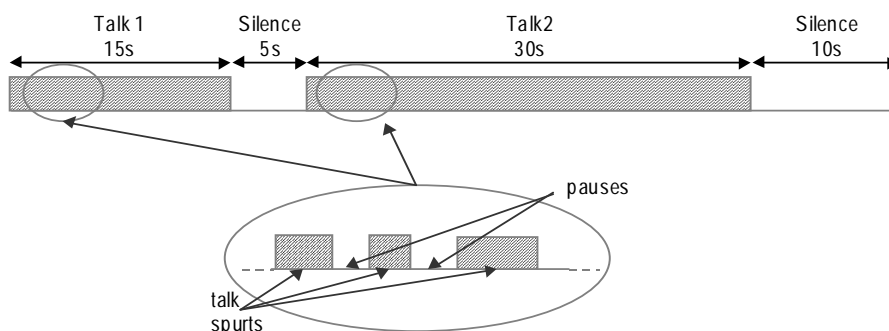


Figure B.1: Delay measurement test signal composition

The test signal contains periods of speech activity (Talk1 and Talk2) and periods of silence. Talk 1 is an initialization sequence, allowing the dynamic jitter buffer to converge.

Both Talk1 and Talk2 contain periods of talk spurts and silence intervals. This is important because jitter buffers are generally designed to adjust their length, so altering the delay of a system, during silence intervals. In ITU-T Recommendation P.59 [12], the average measured talk spurt is 1,0 seconds and the average pause is 1,6 seconds. It is further recommended that the silence intervals are at least 300 ms long.

Delay assessments, using cross-correlation or another appropriate technique, will be made for each talk spurt. At least 10 measurements are required to determine the TIPHON delay measure during "pseudo-stable" operation, which implies 10 opportunities for the jitter buffers to adjust. Therefore, the Talk2 period will contain at least 11 talk spurts.

The measured stable delay for the TIPHON system is the mean delay from all measurements (at least 10) during Talk2 or 90 % of the largest delay measure, whichever is greatest.

Delay measures should be performed for each speech burst during the Talk1 period. Although these measurements are not used for TIPHON classification, they are important, as a slow convergence time maybe unacceptable to a user.

(Study is required into jitter buffer convergence time effects on perceived performance).

NOTE 1: For the delay measurement the talk spurts consist of either a speech-like test signal or natural speech.

NOTE 2: Delay measurements should be accurate to within ± 5 ms.

B.2 Loudness rating

LR calculations are traditionally performed using sine waves placed at 1/3 octave centre frequencies. However, when assessing complex non-linear systems there is a need to use a speech-like test stimulus to pass through devices such as low-bit rate codecs. Since Loudness Ratings are a measure of frequency-dependent loss, it is possible to use wide-band signals to obtain equivalent results. When using a speech-like test stimulus it is important to ensure a reasonable degree of spectrum coverage in the reference signal.

Offsets between reference and recorded signals should be removed and a Fast Fourier Transform (FFT) performed on each. The FFT signals should then be divided into 1/3 octave bands and the loss in each band calculated. These losses can then be used in the LR formulas.

NOTE: The use of the artificial test stimulus to determine LR will be more susceptible to error from circuit noise contributions at higher frequencies.

Annex C (normative): QoS mechanism tests

C.1 Simulated media for QoS calibration

The purpose of this measurement is to allow intrusive testing of network performance for determining network class.

A measurement is made by transmitting a simulated VoIP media stream between two measurement hosts. It is important that the simulated stream represents the in-service VoIP traffic for which the network will transport. The measurement is round-trip (e.g. MH1 - MH2 - MH1) to emulate a full-duplex VoIP call, but one-way measurements (MH1-MH2 and MH2-MH1) can be performed as a part of the round-trip measurement. The measurement hosts can be located within the same domain, such as the core network as shown in figure C.1, or different domains. The measurement methodology is independent of where in the end-to-end path the measurement hosts are located. Locating the measurement hosts in different parts of the network allows for various parts of a network to be qualified.

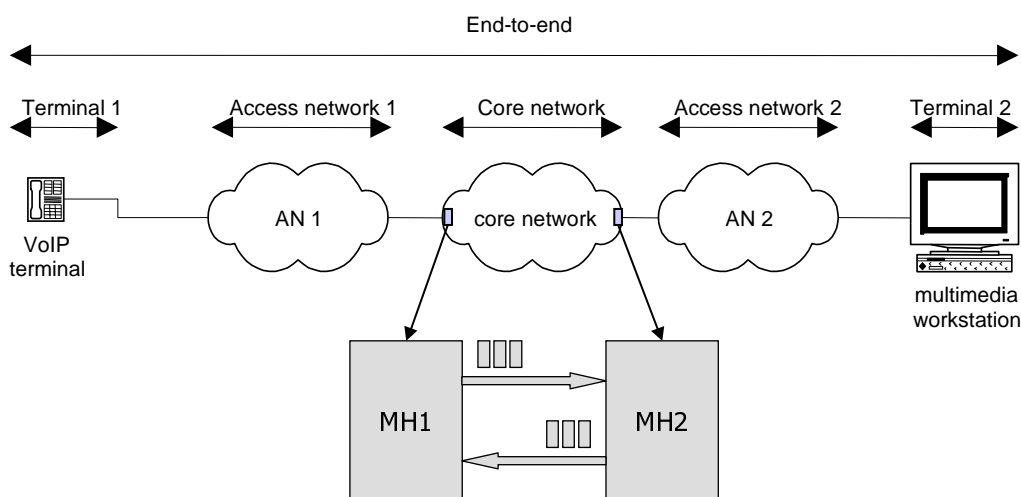


Figure C.1: A test scenario for determining core-network QoS level

The measurement is performed in an application-level fashion using the same protocol stack as a VoIP application (figure C.2). The streams are transmitted through the UDP interfaces in both MH1 and MH2.

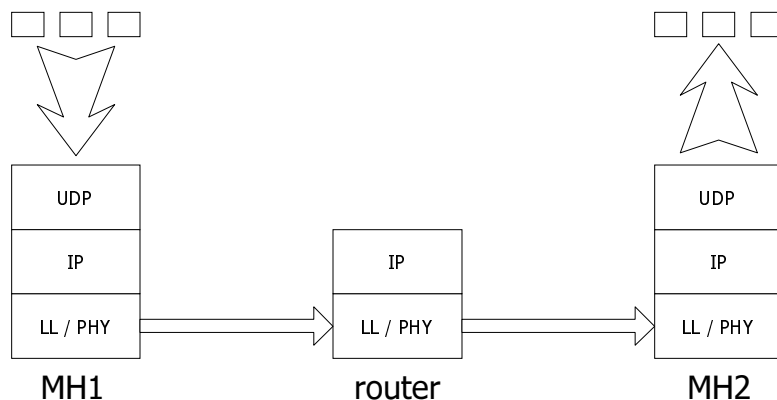


Figure C.2: Protocol stack view of the application-level measurement principal

Simulated Stream Definition and Measurement Process

The simulated media stream needs to be representative of packet sizes, transmission intervals, codec type and more. For a constant bit-rate traffic stream the following parameters need defining:

- Packet size.
- Packet transmission interval.
- Number of packets transmitted.

If discontinuous transmission (DTX) is to be simulated then talkspurt/silence alternation characteristics must be included in the stream description.

An example test stream is a 20 ms framed, continuously transmitted, ITU-T Recommendation G.711 [25] stream. This can be simulated by transmitting 200-byte IP packets (the size includes payload and protocol headers) every 20 ms.

- Payload (20 ms of speech) = 160 bytes
- RTP = 12 bytes
- UDP + IP headers = 28 bytes

Timing and sequence information is then inserted into each packet as the packets are transmitted between the two hosts as shown in figure C.3.

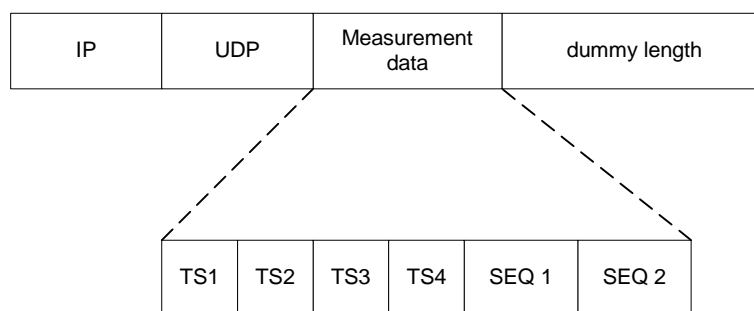


Figure C.3: Structure of a measurement packet

The appropriate protocol layers add IP and UDP headers and padding is used to make the measurement packet correspond to requirements of the simulated stream specifications. The TS and SEQ fields are filled as the packet is transmitted between MH1 and MH2. Here is the process by which these are filled.

Field	Description	Set
TS1	Transmission timestamp at host 1	Set immediately before sending packet to host 2 through UDP interface
TS2	Reception timestamp at host 2	Set immediately after receiving packet at host 2 through UDP interface
TS3	Transmission timestamp at host 2	Set immediately before sending packet to host 1 through UDP interface
TS4	Reception timestamp at host 1	Set immediately after receiving packet to host 1 through UDP interface
SEQ 1	Transmission sequence number at host 1	Set in host 1
SEQ 2	Received sequence number at host 2	Set in host 2

The first packet that MH1 transmits in a measurement is assigned SEQ1 = 0. For subsequent packets, SEQ1 is incremented by one for each transmitted packet. The first packet that MH2 receives is assigned SEQ2 = 0. For each subsequent packet, SEQ2 is incremented by one. Packets are transmitted from host2 to host1 in the order of arrival.

Metric Calculations

Round-trip transmission time, for the route from MH1 - MH2 - MH1, is given by:

$$d_{RTT} = (TS4 - TS3 - TS2 + TS1)$$

Packet delay variation is given by:

$$j = [TSX_N - TSX_{N-1} - (SEQY_N - SEQY_{N-1})\Delta t]$$

where TSX_N is the reception timestamp of the N^{th} received measurement packet and $SEQY_N$ is the transmission timestamp of the N^{th} received measurement packet.

Packet loss ratio is given by

$$l = \frac{M}{N_{pkt}}$$

where M is the number of packets received at the application level and N_{pkt} is the total number of packets transmitted.

Packet loss correlation is illustrated below. In the figure, there are three different loss sequences LS1, LS2, and LS3. The length of the loss sequences indicates the number of adjacent packets lost, and is 1, 2, and 3 respectively.

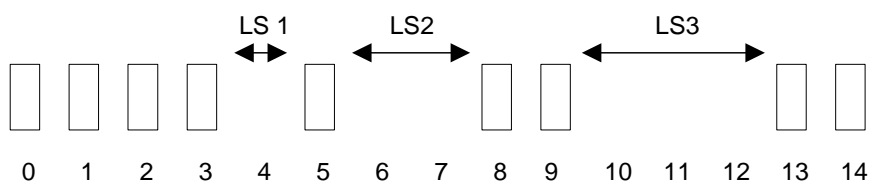


Figure C.4: Illustration of packet loss correlation

The *packet loss correlation* is the number of lost packets divided by the number of loss sequences.

Annex D (normative): Testing procedures for TIPHON terminals and systems

D.1 Introduction

Typical TIPHON terminals and systems may be very likely not of the traditional type using handset telephony. It may be expected, that the typical terminals are more hands-free type telephones. Such the acoustical test set-up should take into account the typical use conditions, test procedures are to be defined to take into account the advanced signal processing expected in such terminals. TIPHON system components may introduce similar complex signal processing. Many of the subjective relevant parameters can be adopted from the intensive investigations of hands-free telephony terminals. In general the same procedures can be used for testing complete TIPHON systems. This annex concentrates on the terminal and system aspects. The general aspects have been introduced already in ITU-T Contribution COM12-42 [24]. A brief overview introducing the general concepts for subjective (auditory) evaluation and extracting the relevant objective quality parameter is given in clause D.2.

D.2 Parameters, defining speech transmission quality including terminals

When evaluation the overall speech transmission quality, networks and terminals may influence quite significantly the speech quality of a connection: Coding, delay and processing techniques like speech echo cancellers or DCME are mainly introduced by the network(s) but similar signal processing can be found in terminals as well. The transfer functions and loudness ratings of a connection are mainly determined by the terminals, the background noise and the background noise transmission are highly influenced by the terminal and the acoustical environment the terminal is exposed to. The conversational properties which are the most important ones in a conversation are determined by the terminal in combination with the network: double talk capability, switching characteristics and delay are dominant impairments often introduced.

In order to find the determining factors a set of subjective test procedures have been developed allowing to extract the dominant quality aspects: Conversational test, talking and listening tests, double talk tests and listening only tests as described in ITU-T Recommendation P.832 [17], ITU-T Contribution COM12-42 [24], ITU-T Recommendation P.831 [16] are the basis of the parameter extraction procedure.

The subjectively relevant parameters determining the "speech transmission quality" are as follows:

The overall quality is determined by:

- sound quality;
- quality of background noise transmission at idle, in single talk and double talk conditions;
- speech level variations during single talk and double talk;
- disturbances caused by switching during single talk and double talk (completeness of speech transmission);
- disturbances caused by echoes during single talk and double talk.

Consequently the objective evaluation needs to be divided into single talk measurements and double talk evaluations. In addition evaluations are required during periods of silence where only background noise is present.

Since the typical test set-up should include all components involved in the mouth to ear transmission a test arrangement should include the terminals "attached" to a realistic substitution of a user and his typical environment. Figure D.1 illustrates how a test set-up from end to end may look like typically.



Figure D.1: Typical test set-up for determining the speech transmission quality from end to end by subjective evaluation of the speech quality relevant parameters (example for handset/hands-free communication)

Test set-ups as shown in figure D.1 are used in auditory (subjective) tests to determine the quality aspects subjectively (see ITU-T Recommendation P.832 [17] and ITU-T Contribution COM12-42 [24]). From these, evaluations procedures have been derived which allow the objective testing of the relevant parameters of terminals (or even end to end scenarios).

D.3 Test set-up for terminals and systems including terminals

See clause 4.

D.3.1 Test Signals

Due to the speech signal based signal processing, standard test signals are not applicable for the tests, appropriate test signals (general description) are defined in ITU-T Recommendations P.50 [22] and P.501 [13].

For narrow band terminals the test signal used shall be bandfiltered between 200 Hz and 4 kHz with a bandpass filter providing a minimum of 24 dB/Oct. filter steepness, when feeding into the receiving direction.

The test signal levels are referred to the average level of the (band filtered in receiving direction) test signal, averaged over the complete test signal length.

D.4 Measurement of standard parameters

The standard parameter and the according measurements to be included here are the following:

Frequency Response in Sending and Receiving Direction;

SLR Sending Loudness Rating;

RLR Receiving Loudness Rating;

(OLR Overall Loudness Rating);

STMR Sidetone Masking Rating;

STR Listener Sidetone Rating;

D D-Value of Terminal;

TCLw Terminal Coupling Loss (weighted);

WEPL Weighted Echo Path Loss;

TELR Talker Echo Loudness Rating;

(qdu	Number of Quantizing Distortion Units);
Nc	Circuit Noise referred to the 0 dBr-point; Distortion in Sending and Receiving Direction; Out of Band Signals in Sending and Receiving Direction.

A more detailed description is found e.g. in TBR 8 [23]. This description of measurement parameters and measurement procedures need to be adapted to the TIPHON terminal/system situation.

In general the measurement principles are the same but special consideration needs to be given to the following points:

- The appropriate measurement signal (especially with respect to the codec) needs to be chosen. It is recommended to use a speech-like test signal. Speech like test stimuli can be found in ITU-T Recommendation P.50 [22] and P.501 [13].
- The averaging times used to determine the transfer characteristics need to be adapted to the measurement signal chosen.
- Instead of level measurements using sine wave signal excitation, typically Fourier transformation is used for calculation /estimation of the output spectra.
- The measured output spectrum is always referred to the signal spectrum in order to determine transfer functions, loudness ratings, etc.
- New procedures need to be established in order to determine distortion, especially to determine the parameter "speech sound quality" in combination with the acoustical interface. The objective speech quality measure described in Draft ITU-T Recommendation P.861 [4] may be used for this purpose.

NOTE: When using the E-model for predicting the speech quality in terms of R-values the following, additional parameter need to be determined:

I_e	Equipment Impairment Factor (low bit-rate Codecs), see ITU-T Recommendations G.107 [9] and ITU-T Recommendation G.113 [27];
Nfor	Noise Floor at the Receive-side;
Ps	Room Noise at the Send-side;
Pr	Room Noise at the Receive-side;

D.5 Advanced Measurements, Taking Into Account the Conversational Situation

In cases where the terminal signal processing cannot be assumed to be linear and time invariant (except the codec), signal processing may influence the speech transmission quite substantially, especially in the conversational situation. The signal processing procedures to be expected are the ones found in hands-free terminals: voice activated switching and amplification, echo cancellation (acoustic and electric), noise reduction, etc. Based on results of subjective tests the ITU-T Recommendation P.340 [6] defines different categories of hands-free telephones. The relevant classification parameter is the objectively measured attenuation range [6]. This parameter highly influences the double talk performance and the quality of background noise transmission. The importance of double talk performance and background noise transmission was derived by conversational tests and investigated more in detail by using specific double talk tests and listening only tests. Table 1 gives an overview of the parameters subjectively relevant for the perceived speech quality.

In the left column of table 1 some of the subjectively most relevant parameters are highlighted as they were identified in subjective tests. The two other columns give a more detailed description and indicate some of the correlating objective parameters. Note, that the given parameters and its combination depend on the technical implementation and they are not necessarily implemented together.

**Table 1: Subjectively relevant parameters and correlating objective parameters
(SND: send direction, RCV: receive direction)**

No.	Subjectively relevant parameter	More detailed description	Correlating Objective parameter
1	quality of background noise transmission	typically the transmission in SND direction <ul style="list-style-type: none"> • at idle mode • with far end speech • with near end speech 	<ul style="list-style-type: none"> • attenuation range • attenuation in SND direction • switching characteristics • minimum activation level in SND direction • frequency response • design of NLP or center clippers in conjunction with EC's • design of noise reduction systems • sensitivity of background noise detection (activation level, absolute level, level fluctuations)
2	double talk performance	typically in SND and RCV direction <ul style="list-style-type: none"> • loudness variation between single and double talk periods • loudness variation during double talk • echo disturbances • occurrence of speech gaps 	attenuation range <ul style="list-style-type: none"> • attenuation in SND/RCV direction during double talk • switching characteristics • minimum activation level to switch over from RCV to SND direction and from SND to RCV direction • echo attenuation • spectral and time dependent echo characteristics • design of NLP or center clippers in conjunction with EC's
3	echo disturbances under single talk conditions	measured between RCV and SND direction	<ul style="list-style-type: none"> • echo level • echo level fluctuation vs. time • spectral echo attenuation
4	speech sound quality	in SND and RCV direction	<ul style="list-style-type: none"> • frequency responses • distortions
5	loudness	in SND and RCV direction	loudness ratings in SND and RCV
6	noise	in SND and RCV direction	<ul style="list-style-type: none"> • noise level • level fluctuations • spectral characteristics

NOTE 1: The relevant subjective and the correlating objective parameters may be incomplete as future technologies may possibly introduce additional parameters.

NOTE 2: The behaviour of subjects during subjective tests clearly demonstrate that the individual speech levels on both sides of the connection highly influence the transmission performance of the HFT under test. Consequently the measurement levels must be adapted in an appropriate way to represent the possible level variations at the RCV input port and the microphone of the HFT.

NOTE 3: Additionally the room characteristics highly influence the transmission quality [24]. This leads to the demand of testing in an appropriate environment.

NOTE 4: Recent subjective tests emphasize the importance of speech clipping and echo disturbances during double talk. It was found, that echo during double talk may influence the speech quality more than expected in the past. The impact on the speech quality, test procedures and limits can be found in ITU-T Recommendations P.340 [6], P.502 [14] and Appendix II to ITU-T Recommendation G.131 [3].

D.5.1 Measurement set-up for objective tests

The measurement set-up should be chosen as described above. In addition some changes are suggested in order to improve the accuracy of the measurements:

- A possible delay in SND or RCV direction of the HFT-type terminals and the acoustical propagation delay between artificial mouth and the HFT microphone should be considered. This is especially important for analyses, where the measured signals are referred to the original test signals. These analyses require an exact synchronization of both signals.
- When analysing hands-free type telephones, in addition a microphone should be positioned very close to the HFT loudspeaker in order to record the RCV signal. This recorded signal can easier be distinguished from the signal introduced by the artificial mouth under double talk measurement conditions.

D.5.2 Practical realization of test signals

Objective parameters like frequency responses, distortions, loudness ratings or noise can be determined with test signals like the Composite Source Signal, artificial voice or others as they are given in ITU-T recommendation P.501 [13]. The signals described in the following are suggested to determine additional parameters as given in table 1. In case background noise simulation is required, care should be taken when choosing the type of background noise. Time and frequency characteristics should be chosen to simulate the typical environment the system is supposed to operate, e.g. car type noise for HFT's in car type environments. In case of specific acoustic pre-processing provided by the terminal, the spatial characteristics of the background noise may be important. In order to simplify the description, long term power density spectrum and long term level density during the measurement should be noted.

D.5.3 Test procedures

Depending on the system under evaluation the following test should be conducted:

- Convergence parameters of echo cancellers (section 4 of ITU-T Recommendation P.502 [14]);
- Speech quality parameters during double talk (section 5 of ITU-T Recommendation P.502 [14]);
- Companding and AGC characteristics (section 6 of ITU-T Recommendation P.502 [14]);
- Quality of background noise transmission (section 7 of ITU-T Recommendation P.502 [14]);
- Switching parameters (section 8 of ITU-T Recommendation P.502 [14]).

When measuring hands-free type of telephones or including them in an overall connection measurement all relevant information can be found in ITU-T Recommendations P.340 [6] and P.581 [15].

Annex E (informative): Method for determining an equipment impairment factor using passive monitoring

E.1 Introduction

This annex describes a method of passively monitoring a Voice over IP stream that produces an Equipment Impairment (I_e) factor that may be used with the E-Model to calculate an R factor. The I_e factor determined using this approach incorporates the effects of the Voice over IP CODEC, packet loss, packet loss distribution, jitter and rency. This process can be applied on a per-call or continuous session basis.

Passive monitoring comprises the extraction of performance metrics from an in-process call or session. This process is non-intrusive, i.e. does not interfere with the data stream, and adds little or no overhead to network traffic. The real time elements of the algorithms described below have been designed to be computationally efficient.

The preferred location of the monitoring point is at the Voice over IP CODEC, at which the monitoring function has access to information such as CODEC type, jitter buffer size and post-jitter buffer packet loss events. The monitoring point may be placed at other locations however may not have access to all the necessary information.

NOTE: The methodology described in this annex requires further validation by subjective testing. Based on these tests numerical constants used in this annex may be modified. After validation of the methodology and some more practical experience annex E should become normative.

E.2 Passive QoS Monitor Framework

The Equipment Impairment factor determined using the methodology described in this annex contains the following elements:

i) I_e (packet loss)

The distribution of packets lost and received is measured from observation of the received packet stream and modelled using a Markov process. The parameters of the Markov process are mapped onto an I_e factor using the CODEC specific curves described in ITU-T Recommendation G.113 [27].

ii) I_e (PDV)

The packet delay variation level is determined during the call and is assumed to be constant throughout the call and bounded by the jitter buffer level and discard thresholds. In many implementations the jitter buffer is of sufficient depth that received packets are either properly aligned in time or discarded, which would render this step un-necessary.

iii) I_e (CODEC)

The CODEC type is assumed to be constant throughout the call and is mapped to an I_e value using the parameters specified in ITU-T Recommendation G.113 [27], Appendix 1.

iv) Delay

The estimated one-way delay, including transmission, jitter buffer and CODEC related delays, is estimated.

E.3 Determining equipment impairment factor for packet loss

IP network packet loss distribution can be modelled using a Markov process. The resulting model can be used in both analytical and numerical performance estimation and has well known and understood properties.

In typical Voice over IP implementations packet loss can occur if packets are excessively delayed. It is therefore preferable to measure packet loss after the receive jitter buffer or with prior knowledge of the packet delay which would cause packets to be discarded. If packet loss is measured before the jitter buffer then it is preferable to measure the per-packet jitter and to assume that any packets that are delayed by more than the jitter buffer level are discarded.

The channel is assumed to have high packet loss (burst) and low packet loss (gap) conditions. During the Voice over IP call packet loss events and inter-loss gaps are counted. At the end of the call, or on request from a service management system, the transition probabilities of the Markov model are determined and used to compute an R factor for the call.

If the number of packets received between two successive lost packets is less than a minimum value g_{\min} then the sequence of the two lost packets and the intervening received packets is regarded as part of a burst. If a sequence of g_{\min} or more packets are correctly received the sequence is regarded as being part of a gap.

The Markov model is defined as having the following states and associated transitions:

State 1 - gap - no loss

p_{11} - packet received

p_{13} - packet loss (start of burst)

p_{14} - isolated packet loss

State 2 - burst - no loss

p_{22} - packet received within burst

p_{23} - packet lost within burst

State 3 - burst - packet loss

p_{31} - packet received (end of burst)

p_{32} - packet received within burst

p_{33} - packet lost

State 4 - gap - packet loss

p_{41} - packet received

This model can be constructed either by accumulating packet loss information during fixed sampling intervals or at packet loss events. An example of a computationally efficient method for determining the parameters of the Markov model is given hereafter.

Assume a counter pkt tracks the number of received packets, $lost$ tracks the number of lost packets in a burst, g_{min} is the minimum gap size, and that an event can be generated if a packet loss is detected:

Packet loss event->

$$c_5 = c_5 + pkt$$

if $pkt \geq g_{min}$ *then*

if $lost = 1$ *then*

$$c_{14} = c_{14} + 1$$

else

$$c_{13} = c_{13} + 1$$

$lost = 1$

$$c_{11} = c_{11} + pkt$$

else

$$lost = lost + 1$$

if $lost > 8$ *then* $c_5 = 0$

if $pkt = 0$ *then*

$$c_{33} = c_{33} + 1$$

else

$$c_{23} = c_{23} + 1$$

$$c_{22} = c_{22} + pkt$$

$$pkt = 0$$

The series of counters c_{11} to c_{14} are used to determine the corresponding Markov model transition probabilities (i.e. c_{11} is used to calculate p_{11}). Counter c_5 is used to measure the delay since the last "significant" burst of lost packets. Parameter g_{min} , the minimum gap size, is typically 16.

The key metrics needed for determining application performance are:

$$c_{31} = c_{13}$$

$$c_{32} = c_{23}$$

$$c_{11} = c_{11} + c_{14} \quad (\text{for simplicity - combine states 4 and 1})$$

$$p_{11} = c_{11} / (c_{11} + c_{13})$$

$$p_{13} = 1 - p_{11}$$

$$p_{31} = c_{31} / (c_{31} + c_{32} + c_{33})$$

$$p_{32} = c_{32} / (c_{31} + c_{32} + c_{33})$$

$$p_{33} = 1 - p_{31} - p_{32}$$

$$p_{22} = c_{22} / (c_{22} + c_{23})$$

$$p_{23} = 1 - p_{22}$$

$$d = (p_{23} p_{31} + p_{13} p_{32} + p_{13} p_{23})$$

$$p_1 = p_{31} p_{23}/d$$

$$p_2 = p_{13} p_{32}/d$$

$$p_3 = p_{13} p_{23}/d$$

frame size	$F =$ frame size	(in seconds)
average packet loss rate	$L = 100 p_3$	percent
gap length	$g = F/(1 - p_{11})$	seconds
gap loss density	$D_g = 100 c_{14}/c_{11}$	percent
burst length	$b = F(1 - p_1)/(p_1 + p_{13})$	seconds
burst loss density	$D_b = 100 p_{23}/(p_{23} + p_{32})$	percent

delay since last burst $y = Fc_5$ seconds

An estimate of the published "Provisional Planning values for the Equipment Impairment Factor" is given by the equation below:

$$I_e(\text{Loss}) = 0 \quad D < 0,5 \text{ percent}$$

$$I_e(\text{Loss}) = d_1 D \quad 0,5 < D < d_2 \text{ percent}$$

$$I_e(\text{Loss}) = d_3 + d_4 D \quad d_2 < D \text{ percent}$$

This can be separately applied to the packet loss rates for the gap and burst state (D_g and D_b), giving I_{eg} and I_{eb} .

CODEC	d_1	d_2	d_3	d_4
G.723.1+VAD 6,3k	4,25	4,8	12	1,75
G.729A				

E.4 Determining Equipment Impairment Factor for Packet Delay Variation

The packet delay variation is bounded by the jitter buffer, which removes small amounts of variation by increasing delay, and by the discard threshold. In many implementations the discard threshold is effectively equal to the jitter buffer delay, which would result in packets either be properly retimed or discarded - in this case the $I_e(\text{PDV})$ value would be 0.

The effect of low levels of packet delay variation on voice quality is substantially less than that of packet loss. Packet loss and packet delay variation are often correlated as high levels of packet delay variation will lead to an increased level of packet discard.

Let the packet inter-arrival time be t , the jitter buffer delay be denoted t_{jb} , the discard delay be denoted t_{discard} and the adjusted packet inter-arrival time be t' - all given in milliseconds.

$$t' = 0 \quad t < t_{jb}$$

$$t' = t - t_{jb} \quad t_{jb} < t < t_{\text{discard}}$$

$$\text{omit measurement} \quad t > t_{\text{discard}}$$

adjusted PDV = average(t')

$I_e(\text{PDV}) = 0,1 \text{ adjusted PDV}$

E.5 Measuring Delay

The monitoring function estimates the round trip transmission delay using an echo mechanism, for example RTCP. This value is divided by two to give the estimated one way transmission delay t_{owtd} .

NOTE: This assumes that the delay is symmetric.

The processing delay through the "transmitter" and "receiver" is:

CODEC encoder delay t_{enc}

Framing delay t_{frame}

Jitter buffer delay t_{jb}

Decoding delay t_{dec}

The values of these parameters for typical CODECs can be found in ITU-T Recommendation G.114 Appendix 1 [28].

The overall one-way delay is therefore

$$t_d = t_{\text{owtd}} + t_{\text{enc}} + t_{\text{frame}} + t_{\text{jb}} + t_{\text{dec}}$$

E.6 Determining Equipment Impairment Factor for CODEC

ITU Recommendation G.113 Appendix 1 [27] gives the following Equipment Impairment factors for certain CODECs.

CODEC	G.711 [25]	G.729A + VAD	G.723.1 + VAD 6,3 kbit/s
I_e (CODEC)	0	11	15

E.7 Determining Overall Equipment Impairment Factor

E.7.1 Determining Average Equipment Impairment Factor

It is generally accepted that perceived quality does not change abruptly but exponentially "decays" from one level to another. This is intuitively obvious, as a 100 ms burst of noise would be less annoying than a 10 s burst.

Determine the Equipment Impairment value for the burst condition and the gap condition as:

$$I_{\text{eg}} = I_{\text{eg}}(\text{LOSS}) + I_e(\text{PDV}) + I_e(\text{CODEC})$$

$$I_{\text{eb}} = I_{\text{eg}}(\text{LOSS}) + I_e(\text{PDV}) + I_e(\text{CODEC})$$

Let I_1 be the quality level at the change from burst condition I_{eb} to gap condition I_{eg} and let I_2 be the quality level at the change from I_{eg} to I_{eb} .

$$I_1 = I_{\text{eb}} - (I_{\text{eg}} - I_2) e^{-b/t_1} \quad \text{where } t_1 \text{ typically equals } 5$$

$$I_2 = I_{\text{eg}} + (I_1 - I_{\text{eg}}) e^{-g/t_2} \quad \text{where } t_2 \text{ typically equals } 15$$

Combining these gives:

$$I_2 = (I_{eg} (1 - e^{-g/t_2}) + I_{eb} (1 - e^{-b/t_1}) e^{-g/t_2}) / (1 - e^{-b/t_1 - g/t_2})$$

Integrating the expressions for I_1 and I_2 to give a time average gives:

$$I_e(av) = (b I_{eb} + g I_{eg} - t_1 (I_{eb} - I_2) (1 - e^{-b/t_1}) + t_2 (I_1 - I_{eg}) (1 - e^{-g/t_2})) / (b + g)$$

E.7.2 Recency Effect

It has been noted by a number of researchers that the perceived quality of a call varies with the location of impairments. Impairments occurring late in a call have more effect than those occurring early in the call.

ANSI T1A1.7/98-031 [30] described an experiment in which both mutes and noise bursts were introduced at the beginning, middle and end of a 60 second call. For the "high burst" result given:

burst at start of call MOS = 3,82

burst at middle of call MOS = 3,28

burst at end of call MOS = 3,18

ITU-T Contribution COM.12-D.139 [29] conducted an experiment in which a burst of high packet loss of duration 15, 30 s or 60 s was introduced at the start, middle and end of a 180 second call and noted similar effects.

It is proposed that a simplified "adjustment" for recency be used, to minimize complexity. The delay since the last burst of packet loss is given above as y . It is assumed that the value of I_e at the end of the previous burst is given by I_1 and that the adjusted average quality approaches $I_e(av)$ exponentially.

$$I_e(\text{end of call}) = I_e(av) + (k(I_1 - I_e(av))) e^{-y/t_3}$$

where y is the delay to the previous burst, t_3 is a time constant (assumed to be 30) and k is a constant (assumed to be 0,7).

E.8 Use of the E Model

The Equipment Impairment value and the estimated one-way delay determined above may be used as inputs to the E-Model (ITU-T Recommendation G.107 [9]) in order to calculate an R factor for the call. If other parameters required for the E-Model are unavailable then they should be set to their default values.

Annex F (informative): Bibliography

Gierlich, H.W., Kettler, F., Diedrich, E.: "Speech Quality Evaluation of Hands-Free Telephones During Double talk: New Evaluation Methodologies; EUSIPCO" (1998), Proceedings, Vol. II.

Gierlich, H.W.: "The Auditory Perceived Quality of Hands-Free Telephones: Auditory Judgements, Instrumental Measurements and Their Relationship, Speech Communication 20" (1996), 241-254, October 1996.

ITU-T Recommendation P.810 (1996): "Modulated Noise Reference Unit (MNRU)".

ITU-T Recommendation G.168: "Digital Network Echo Cancellers".

ETSI EG 201 377-1: "Speech Processing, Transmission and Quality Aspects (STQ); Specification and measurement of speech transmission quality; Part 1: Introduction to objective comparison measurement methods for one-way speech quality across networks".

RFC 1889 RTP: "A Transport Protocol for Real-Time Applications".

ETSI TS 102 024 series: TIPHON Release 4 QoS documents.

ETSI TS 102 025 series: TIPHON Release 5 QoS documents.

History

Document history		
V1.1.1	November 2000	Publication as TS 101 329-5
V1.1.2	January 2002	Publication as TS 101 329-5
V4.1.1	September 2003	Publication