SLOVENSKI STANDARD
SIST ES 202 740 V1.3.2:2016

01-julij-2016

Kakovost prenosa govora in večpredstavnih vsebin (STQ) - Prenosne zahteve za širokopasovne zvočniške in prostoročne terminale VoIP glede na kakovost storitev (QoS), kot jih dojema uporabnik

Speech and multimedia Transmission Quality (STQ) - Transmission requirements for wideband VoIP loudspeaking and handsfree terminals from a QoS perspective as perceived by the user

Ta slovenski standard je istoveten z: ETSI ES 202 740 V1.3.2 (2010-09)

ICS:

33.050.01	Telekomunikacijska terminalska oprema na splošno	Telecommunication terminal equipment in general

SIST ES 202 740 V1.3.2:2016 en

ETSI ES 202 740 V1.3.2 (2010-09)

ETSI Standard

Speech and multimedia Transmission Quality (STQ); Transmission requirements for wideband VoIP loudspeaking and handsfree terminals from a QoS perspective as perceived by the user

Reference

RES/STQ-00165

Keywords

terminal, handsfree, loudspeaking, VoIP, quality

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 ETSI Standard (ES) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ).

Introduction

Traditionally, the analogue and digital telephones were interfacing switched-circuit 64 kbit/s PCM networks. With the fast growth of IP networks, wideband terminals providing higher audio-bandwidth and directly interfacing packet-switched networks (VoIP) are being rapidly introduced. Such IP network edge devices may include gateways, specifically designed IP phones, soft phones or other devices connected to the IP based networks and providing telephony service. Since the IP networks will be in many cases interworking with the traditional PSTN and private networks, many of the basic transmission requirements have to be harmonized with specifications for traditional digital terminals. However, due to the unique characteristics of the IP networks including packet loss, delay, etc. new performance specification, as well as appropriate measuring methods, will have to be developed. Terminals are getting increasingly complex, advanced signal processing is used to address the IP specific issues.

NOTE: Requirement limits are given in tables, the associated curve when provided is given for illustration.

1 Scope

The present document provides speech transmission performance requirements for 8 kHz wideband VoIP loudspeaking and hands-free terminals; it addresses all types of IP based terminals, including wireless, softphones and group audio terminals.

In contrast to other standards which define minimum performance requirements it is the intention of the present document to specify terminal equipment requirements which enable manufacturers and service providers to enable good quality end-to-end speech performance as perceived by the user.

In addition to basic testing procedures, the present document describes advanced testing procedures taking into account further quality parameters as perceived by the user.

NOTE: The present document does not concern headset terminals.

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 reference 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.

[1] ETSI I-ETS 300 245-6: "Integrated Services Digital Network (ISDN); Technical characteristics of telephony terminals; Part 6: Wideband (7 kHz), loudspeaking and hands free telephony".

[2] ETSI TS 126 171: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); AMR speech codec, wideband; General description (3GPP TS 26.171 version 6.0.0 Release 6)".

[3] ITU-T Recommendation G.108: "Application of the E-model: A planning guide".

[4] ITU-T Recommendation G.109: "Definition of categories of speech transmission quality".

[5] ITU-T Recommendation G.122: "Influence of national systems on stability and talker echo in international connections".

[6] ITU-T Recommendation G.131: "Talker echo and its control".

[7] ITU-T Recommendation G.711: "Pulse code modulation (PCM) of voice frequencies".

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

[9] ITU-T Recommendation G.722.1: "Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss".

[10] ITU-T Recommendation G.729.1: "G.729 based Embedded Variable bit-rate coder: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729".

[11] ITU-T Recommendation G.1020: "Performance parameter definitions for quality of speech and other voiceband applications utilizing IP networks".

[12] ITU-T Recommendation P.50: "Artificial voices".

[13] ITU-T Recommendation P.56: "Objective measurement of active speech level".

[14] ITU-T Recommendation P.58: "Head and torso simulator for telephonometry".

[15] ITU-T Recommendation P.79: "Calculation of loudness ratings for telephone sets".

[16] ITU-T Recommendation P.310: "Transmission characteristics for telephone band (300-3400 Hz) digital telephones".

[17] ITU-T Recommendation P.340: "Transmission characteristics and speech quality parameters of hands-free terminals".

[18] ITU-T Recommendation P.341: "Transmission characteristics for wideband (150-7000 Hz) digital hands-free telephony terminals".

[19] ITU-T Recommendation P.501: "Test signals for use in telephonometry".

[20] ITU-T Recommendation P.502: "Objective test methods for speech communication systems using complex test signals".

[21] ITU-T Recommendation P.581: "Use of head and torso simulator (HATS) for hands-free terminal testing".

[22] 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".

[23] ISO 3 (1973): "Preferred numbers - Series of preferred numbers".

[24] ITU-T Recommendation P.800.1: "Mean Opinion Score (MOS) terminology".

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] ETSI EG 202 396-1: "Speech Processing, Transmission and Quality Aspects (STQ); Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database".

[i.2] ETSI EG 202 425: "Speech Processing, Transmission and Quality Aspects (STQ); Definition and implementation of VoIP reference point".

[i.3] ETSI EG 202 396-3: "Speech Processing, Transmission and Quality Aspects (STQ); Speech quality performance in the presence of background noise; Part 3: Background noise transmission - objective model".

[i.4] ETSI TR 102 648-1: "Speech Processing, Transmission and Quality Aspects (STQ); Test Methodologies for ETSI Test Events and Results; Part 1: VoIP Speech Quality Testing".

[i.5] NIST net.

NOTE: Available at http://snad.ncsl.nist.gov/itg/nistnet/.

[i.6] Netem.

NOTE: Available at http://www.linuxfoundation.org/en/Net:Netem.

3 Definitions and abbreviations

3.1 Definitions

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

artificial ear: device for the calibration of earphones incorporating an acoustic coupler and a calibrated microphone for the measurement of the sound pressure and having an overall acoustic impedance similar to that of the median adult human ear over a given frequency band

codec: combination of an analogue-to-digital encoder and a digital-to-analogue decoder operating in opposite directions of transmission in the same equipment

ear-Drum Reference Point (DRP): point located at the end of the ear canal, corresponding to the ear-drum position

freefield equalization: artificial head is equalized in such a way that for frontal sound incidence in anechoic conditions the frequency response of the artificial head is flat

freefield reference point: point located in the free sound field, at least in 1,5 m distance from a sound source radiating in free air

NOTE: In case of a head and torso simulator (HATS) in the centre of the artificial head with no artificial head present.

group-audio terminal: handsfree terminal primarily designed for use by several users which will not be equipped with a handset

handsfree telephony terminal: telephony terminal using a loudspeaker associated with an amplifier as a telephone receiver and which can be used without a handset

HATS Hands-Free Reference Point (HATS HFRP): corresponds to a reference point "n" from ITU-T Recommendation P.58 [14] "n" is one of the points numbered from 11 to 17 and defined in table 6a of ITU-T Recommendation P.58 [14] (coordinates of far field front point)

NOTE: The HATS HFRP depends on the location(s) of the microphones of the terminal under test: the appropriate axis lip-ring/HATS HFRP is to be as close as possible to the axis lip-ring/HFT microphone under test.

Head And Torso Simulator (HATS) for telephonometry: manikin extending downward from the top of the head to the waist, designed to simulate the sound pick-up characteristics and the acoustic diffraction produced by a median human adult and to reproduce the acoustic field generated by the human mouth

loudspeaking function: function of a handset telephone using a loudspeaker associated with an amplifier as a telephone receiver

Mouth Reference Point (MRP): is located on axis and 25 mm in front of the lip plane of a mouth simulator

nominal setting of the volume control: setting which is closest to the nominal RLR

softphone: speech communication system based upon a computer

3.2 Abbreviations

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

CSS	Composite Source Signal
DRP	ear Drum Reference Point
EL	Echo Loss
ERP	Ear Reference Point
HATS	Head And Torso Simulator
HFRP	Hands Free Reference Point
L_E	Earphone coupling Loss

MOS-LQOy	Mean Opinion Score - Listening Quality Objective, y being n for narrow-band, w for wideband, and M for mixed

NOTE: See ITU-T Recommendation P.800.1 [24].

MRP	Mouth Reference Point
NLP	Non Linear Processor
PCM	Pulse Code Modulation
PLC	Packet Loss Concealment
POI	Point Of Interconnection
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RLR	Receive Loudness Rating
RLRmax	Receive Loudness Rating corresponding to the maximum setting of the volume control
RLRmin	Receive Loudness Rating corresponding to the minimum setting of the volume control
SLR	Send Loudness Rating
TCLw	Terminal Coupling Loss (weighted)
TCN	Trace Control for Netem
TELR	Talker Echo Loudness Rating
VoIP	Voice over Internet Protocol

4 General considerations

4.1 Coding Algorithm

The assumed coding algorithm is according to ITU-T Recommendation G.722 [8]. VoIP terminals may support other coding algorithms.

NOTE: Associated Packet Loss Concealment, e.g. as defined in ITU-T Recommendation G.722 [8], appendixes 3 and 4, should be used.

4.2 End-to-end considerations

In order to achieve a desired end-to-end speech transmission performance (mouth-to-ear) it is recommended that general rules of transmission planning tasks are carried out with the E-model taking into account that E-model does not directly address handsfree or loudspeaking terminals; this includes the a-priori determination of the desired category of speech transmission quality as defined in ITU-T Recommendation G.109 [4].

While, in general, the transmission characteristics of single circuit-oriented network elements, such as switches or terminals can be assumed to have a single input value for the planning tasks of ITU-T Recommendation G.108 [3], this approach is not applicable in packet based systems and thus there is a need for the transmission planner's specific attention.

In particular the decision as to which delay measured according to the present document should is acceptable or representative for the specific configuration is the responsibility of the individual transmission planner.

ITU-T Recommendation G.108 [3] with its amendments provides further guidance on this important issue.

The following optimum terminal parameters from a users' perspective need to be considered:

- Minimized delay in send and receive direction.
- Optimum loudness Rating (RLR, SLR).
- Compensation for network delay variation.
- Packet loss recovery performance.
- Maximized terminal coupling loss.

- Some more basic (I-ETS 300 245-6 [1]) parameters are applicable, if ITU-T Recommendation G.722 [8] is used.

4.3 Parameters to be investigated

4.3.1 Basic parameters

The basic parameters are given in I-ETS 300 245-6 [1], ITU-T Recommendation P.340 [17] and ITU-T Recommendation P.341 [18].

4.3.2 Further Parameters with respect to Speech Processing Devices

For VoIP terminals that contain non-linear speech processing devices, the following parameters require additional attention in the context of the present document.

The measurements for further parameters with respect to speech processing devices which are novelties to terminal requirement standards, have been successfully used in the ETSI Speech Quality Test Events (see TR 102 648-1 [i.4]):

- Objective evaluation of speech quality for VoIP terminals.
- Minimum activation level and sensitivity in Receive direction.
- Automatic Level Control in Receive.
- Double Talk Performance.
- Minimum activation level and sensitivity of double talk detection.
- Switching characteristics.
- Quality of echo cancellation.
- Variant Impairments; Network dependant.
- Etc.

5 Test equipment

5.1 IP half channel measurement adaptor

The IP half channel measurement adaptor is described in EG 202 425 [i.2].

5.2 Environmental conditions for tests

The following conditions shall apply for the testing environment:

a) ambient temperature: 15 °C to 35 °C (inclusive);

b) relative humidity: 5 % to 85 %;

c) air pressure: 86 kPa to 106 kPa (860 mbar to 1 060 mbar).

5.3 Accuracy of measurements and test signal generation

Unless specified otherwise, the accuracy of measurements made by test equipment shall be equal to or better than:

Table 1: Measurement Accuracy

Item	Accuracy
Electrical signal level	±0,2 dB for levels ≥ -50 dBV ±0,4 dB for levels < -50 dBV
Sound pressure	±0,7 dB
Frequency	±0,2 %
Time	±0,2 %
Application force	±2 Newton
Measured maximum frequency	10 kHz

NOTE: The measured maximum frequency is due to P. 58 limitations [14].

Unless specified otherwise, the accuracy of the signals generated by the test equipment shall be better than:

Table 2: Accuracy of test signal generation

Quantity	Accuracy
Sound pressure level at Mouth Reference Point (MRP)	±3 dB for frequencies from 100 Hz to 200 Hz ±1 dB for frequencies from 200 Hz to 4 000 Hz ±3 dB for frequencies from 4 000 Hz to 8 000 Hz
Electrical excitation levels	±0,4 dB across the whole frequency range
Frequency generation	±2 % (see note)
Time	±0,2 %
Specified component values	±1 %
NOTE: This tolerance may be used to avoid measurements at critical frequencies, e.g. those due to sampling operations within the terminal under test.	

For terminal equipment which is directly powered from the mains supply, all tests shall be carried out within ±5 % of the rated voltage of that supply. If the equipment is powered by other means and those means are not supplied as part of the apparatus, all tests shall be carried out within the power supply limit declared by the supplier. If the power supply is alternate current, the test shall be conducted within ±4 % of the rated frequency.

5.4 Network impairment simulation

At least one set of requirements is based on the assumption of an error free packet network, and at least one other set of requirements is based on a defined simulated loss of performance of the packet network.

An appropriate network simulator has to be used, for example NISTnet [i.5] (http://snad.ncsl.nist.gov/itg/nistnet/) or Netem [i.6].

Based on the positive experience, STQ have made during the ETSI Speech Quality Test Events with "NIST Net" this will be taken as a basis to express and describe the variations of packet network parameters for the appropriate tests.

Here is a brief blurb about NIST Net:

- The NIST Net network emulator is a general-purpose tool for emulating performance dynamics in IP networks. The tool is designed to allow controlled, reproducible experiments with network performance sensitive/adaptive applications and control protocols in a simple laboratory setting. By operating at the IP level, NIST Net can emulate the critical end-to-end performance characteristics imposed by various wide area network situations (e.g. congestion loss) or by various underlying subnetwork technologies (e.g. asymmetric bandwidth situations of xDSL and cable modems).

- NIST Net is implemented as a kernel module extension to the Linux operating system and an X Window System-based user interface application. In use, the tool allows an inexpensive PC-based router to emulate numerous complex performance scenarios, including: tuneable packet delay distributions, congestion and background loss, bandwidth limitation, and packet reordering/duplication. The X interface allows the user to select and monitor specific traffic streams passing through the router and to apply selected performance "effects" to the IP packets of the stream. In addition to the interactive interface, NIST Net can be driven by traces produced from measurements of actual network conditions. NIST Net also provides support for user defined packet handlers to be added to the system. Examples of the use of such packet handlers include: time stamping/data collection, interception and diversion of selected flows, generation of protocol responses from emulated clients.

The key points of Netem can be summarized as follows:

- Netem is nowadays part of most Linux distributions, it only has to be switched on, when compiling a kernel. With Netem, there are the same possibilities as with nistnet, there can be generated loss, duplication, delay and jitter (and the distribution can be chosen during runtime). Netem can be run on a Linux-PC running as a bridge or a router (Nistnet only runs on routers).
- With an amendment of Netem, TCN (Trace Control for Netem) which was developed by ETH Zurich, it is even possible, to control the behaviour of single packets via a trace file. So it is for example possible to generate a single packet loss, or a specific delay pattern. This amendment is planned to be included in new Linux kernels, nowadays it is available as a patch to a specific kernel and to the iproute2 tool (iproute2 contains Netem).
- It is not advised to define specific distortion patterns for testing in standards, because it will be easy to adapt devices to these patterns (as it is already done for test signals). But if a pattern is unknown to a manufacturer, the same pattern can be used by a test lab for different devices and gives comparable results. It is also possible to take a trace of Nistnet distortions, generate a file out of this and playback exactly the same distortions with Netem.

5.5 Acoustic environment

In general two possible approaches need to be taken into account: either room noise and background noise are an inherent part of the test environment or room noise and background noise shall be eliminated to such an extent that their influence on the test results can be neglected.

Unless stated otherwise measurements shall be conducted under quiet and "anechoic" conditions. Depending on the distance of the transducers from mouth and ear a quiet office room may be sufficient e.g. for handsets where artificial mouth and artificial ear are located close to the acoustical transducers. But this is not applicable for handsfree and loudspeaking terminals.

In cases where real or simulated background noise is used as part of the testing environment, the original background noise must not be noticeably influenced by the acoustical properties of the room.

In all cases where the performance of acoustic echo cancellers shall be tested, a realistic room, which represents the typical user environment for the terminal shall be used.

In case where an anechoic room is not available the test room has to be an acoustically treated room with few reflections and a low noise level.

Considering this, test laboratory, in the case where its test room does not conform to anechoic conditions as given in ITU-T Recommendation P.341 [18], has to present difference in results for measurements due to its test room.

5.6 Influence of terminal delay on measurements

As delay is introduced by the terminal, care shall be taken for all measurements where exact position of the analysis window is required. It shall be checked that the test is performed on the test signal and not on any other signal.