Final draft ETSI ES 202 765-2 V1.1.3 (2010-03)

ETSI Standard

Speech and multimedia Transmission Quality (STQ); QoS and network performance metrics and measurement methods; Part 2: Transmission Quality Indicator combining Voice Quality Metrics

Reference
RES/STQ-00148-2

Keywords
performance, QoS, user, 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

Copyright Notification

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), and is now submitted for the ETSI standards Membership Approval Procedure.

1 Scope

The present document aims at identifying and defining indicators and methodologies for a use in a context of end-user quality characterization and supervision of voice telephony services.

In this context the measurements and metric determinations are perform by analysing signals accessible on user-end services and not on the network. In order to mirror the reality in terms of access to the services at the user-end measurements and analysis are perform on electrical signal that exclude the electro-acoustic part of the end equipment but the probe adaptation to electric interface of the end user equipment much take into account the electro-acoustic characteristics of this terminal.

All the indicators presented and defined in the present document are objective indicators obtained by instrumental measurement methods.

2 References

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.
- Non-specific reference may be made only to a complete document or a part thereof and only in the following cases:
  - if it is accepted that it will be possible to use all future changes of the referenced document for the purposes of the referring document;
  - for informative references.

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 indispensable for the application of the present document. For dated references, only the edition cited applies. For non-specific references, the latest edition of the referenced document (including any amendments) applies.

Not applicable

2.2 Informative references

The following referenced documents are not essential to the use of the present document but they assist the user with regard to a particular subject area. For non-specific references, the latest version of the referenced document (including any amendments) applies.

[i.1] ITU-T Recommendation P.800: "Methods for subjective determination of transmission quality".

[i.2] 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.3] ITU-T Recommendation P.862.1: "Mapping function for transforming P.862 raw result scores to MOS-LQO".

[i.4] ITU-T Recommendation P.862.2: "Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs".

[i.5] 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.6] ITU-T Recommendation P.800.1: "Mean Opinion Score (MOS) terminology".

[i.7] ITU-T Recommendation E.800: "Terms and definitions related to quality of service and network performance including dependability".

[i.8] ITU-T Recommendation E.845: "Connection accessibility objective for the international telephone service".

[i.9] ETSI EG 201 769: "Speech Processing, Transmission and Quality Aspects (STQ); QoS parameter definitions and measurements; Parameters for voice telephony service required under the ONP Voice Telephony Directive 98/10/EC".

[i.10] ITU-T Recommendation P.56: "Objective measurement of active speech level".

[i.11] ITU-T Recommendation O.41: "Psophometer for use on telephone-type circuits".

[i.12] ITU-T Recommendation G.131: "Talker echo and its control".

[i.13] ITU-T Recommendation G.168: "Digital network echo cancellers".

[i.14] ITU-T Recommendation G.114: "One-way transmission time".

[i.15] ITU-T Recommendation P.505: "One-view visualization of speech quality measurement results".

[i.16] ETSI EG 201 377 (all parts): "Speech Processing, Transmission and Quality Aspects (STQ); Specification and measurement of speech transmission quality".

[i.17] ITU-T Recommendation H.323: "Packet-based multimedia communications systems".

[i.18] ITU-T Recommendation H.225.0: "Call signalling protocols and media stream packetization for packet-based multimedia communication systems".

[i.19] ITU-Recommendation P.50: "Artificial voices".

[i.20] ITU-Recommendation P.501: "Test signals for use in telephonometry".

NOTE: This Recommendation includes an electronic attachment containing test signals for telephonometry applications.

[i.21] ETSI TR 102 506: "Speech Processing, Transmission and Quality Aspects (STQ); Estimating Speech Quality per Call".

3 Abbreviations

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

ADSL	Asymmetrical Digital Subscriber Line
ATA	Analog Telephone Adapter
IP	Internet Protocol
ISDN	Integrated Services Digital Network
ITU-T	International Telecommunication Union - Telecommunication standardization sector
GPS	Global Positioning System
GSM	Global System for Mobile communications
HATS	Head And Torso Simulator
MGCP	Media Gateway Control Protocol
MOS	Mean Opinion Score
MOS-LQOM	Mean Opinion Store-Listening Quality Objective Mixed bandwidths
PDD	Post Dialling Delay

PESQ	Perceptual Evaluation of Speech Quality
PSTN	Public Switched Telephone Network
RTP	Real Time Protocol
SIP	Session Initiation Protocol
UMTS	Universal Mobile Telecommunications Service
VoIP	Voice over Internet Protocol

4 Introduction

The assessment of transmission quality based on voice quality metrics is already addressed in several standards at ETSI (e.g. EG 201 377 [i.16] series) and elsewhere (mostly ITU-T recommendations from the P and G series). These different documents are addressing the measurement methodologies in terms of metrics, threshold, data acquisition or modelling of subjective opinion.

The objective of the present document is to complement this material with practical requirements of use in the context of service verification and benchmark on a large and representative scale from the point of view of the end-users or of the regulatory authorities. This has been made necessary by the current or recent evolutions of the telecommunication sector:

- the competitive environment, in particular in voice services, where public protocols with high quality services have been replaced by a multitude of service providers with less guarantees, and where clients can very easily change their service providers;
- the development of time varying quality in telecommunications, first in mobile offers (due to mobility and irregular network coverage), but now also for fix services (mostly VoIP);
- the cohabitation, interaction and competition between services based on different technologies.

Voice transmission quality is now recognized as a differentiating factor, but it remains very difficult to quantify.

To achieve the goal mentioned beforehand, there are several existing possibilities, not fully satisfying:

- Customer surveys. This is by far the cheapest way to assess the perception of end users. But the bias introduced by the other factors like price, as well as the fact that voice quality itself is rarely questioned as itself or in a satisfactory way (one never knows before a survey what are the problems encountered by end users), makes this source not really reliable.
- Pseudo-subjective tests, with a few human testers assessing the quality of real links in several situations. This method has the major drawback of its lack of reproducibility, and is often applied without using the standard metrics and quality scales that can be found in standards like ITU-T Recommendation P.800 [i.1]. It is also very long to run and not really cheap in the current competitive context where so many offers have to be assessed. And it is not easily applicable in a context of quality changing over time.
- Objective tests. This is the most reliable way, although it is also based on sampling and can cost a lot of money in the case of a large deployment of probes or robots.

The present document assumes that this last family of methodology answers the needs of a reliable comparison of telephony offers and is applied without combination with other methods.

What definitely matters is the point of view of the end-users. What they perceive is not only the result of the transmission of a signal across a network; the processing of this signal at the sending and at the receiving sides has also a big importance. Therefore, it seems obvious not to use passive network monitoring systems to assess end-to-end voice quality, but rather active systems simulating the behaviour of the end users, including the terminal. A big advantage of such an approach is that it is highly technical and protocol agnostic, and therefore compliant with the expectations of users, which are not judging voice quality of PSTN, GSM or VoIP services following different criteria.

Last important aspect that is addressed in the present document is the practical organization of measurement campaigns in order to get a realistic and reliable vision of the services as perceived by the end-users. In particular, the questions of the periodicity of measurement and of the geographical coverage (i.e. more generally the sampling approach).

In order to mirror the reality in terms of access to the services, a reliable measurement or supervision system should provide the possibility to collect information from probes or robots adapted to the most common interfaces available. This includes:

- analogue access (for the simulation of PSTN or of analogue phones behind an ATA box or an ADSL modem);
- ISDN access;
- handset (for any wireline terminal);
- electrical input and output (for PC soundcards of for any wireless terminal);
- GSM;
- UMTS;
- ethernet with IP phone termination (SIP, ITU-T Recommendation H.323 [i.17], MGCP, etc.).

Any combination of end-to-end connection between the types of access mentioned here have to be considered when a measurement campaign is scheduled. Nevertheless, of course, there are practical limitations:

- the number of measurements for a given type of access should be in proportion with its level of use in the real life;
- the number of probes and of measurement results available will be adapted to the real needs as well as to the capacity (mostly in terms of cost and of processing capability) of the entity running these measurements.

Figure 4.1 shows these different configurations and interfaces.

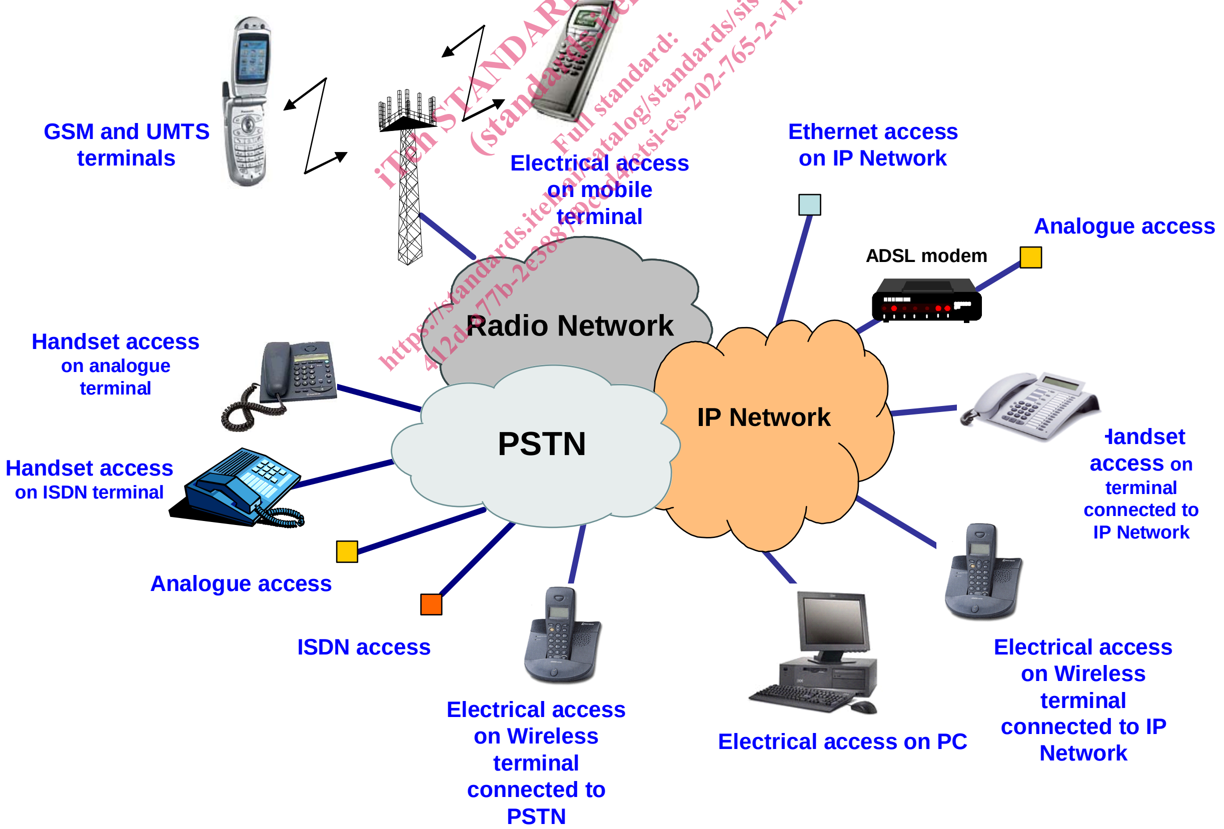

Figure 4.1: Possible configurations and interfaces in context of user characterization

5 Measurement type

To perform quality services assessments, there are two different methods: intrusive and non intrusive measurements.

The non intrusive measurements are not really adapted to end user surveys because it requires to install probes at the user's terminals.

The intrusive measurements are more adapted to end user surveys because probe connection with end user terminals is easier. Compared to non intrusive measurements, the intrusive methods have an advantage: the opportunity for voice quality assessment to use models with references such as ITU-T Recommendation P.862 [i.2] (see also ITU-T recommendations P.862.1 [i.3], P.862.2 [i.4] and P.862.3 [i.5] concerning mapping functions and application guide) which give results close to subjective perception of the speech quality.

In this context, the intrusive measurements using models working with references for speech quality assessment will be perform for end user survey.

6 Voice quality scale

It is important to consider that nowadays telephony has entered an era where traditional narrowband services will cohabit with new services offering wideband audio capacities. For end-users, these are not separated kinds of services. Therefore, the assessment of transmission quality of voice should now be based on common metrics and objective quality levels and scales, in replacement of the existing narrow-band only ones. In this context, it is appropriate to use the MOS-LQOM scale to characterize voice quality of narrow-band services and wideband services. See ITU-T Recommendation P.800.1 [i.6] for more information on MOS terminology.

7 List of indicators

The indicators for the context of end-user quality survey of voice services are given in the following sub-clauses.

It should be noted that the indicators defined below do not completely cover the conversational quality. Additional indicators may be:

- double talk performance parameters;
- switching characteristics parameters.

7.1 Post Dialling Delay

Definition	Post Dialling Delay (PDD) evaluates service availability to set up calls in an acceptable delay. It is linked to the service architecture complexity, and to the performance of the constituting network elements. Post Dialling Delay is the time interval between the end of dialling by the caller and the reception back by him of the appropriate ringing tone or recorded announcement. Metric determines on one of the two access of the communication.
Assessment method	Indicator determines sequentially from the two access of call configuration. This indicator characterizes only the caller part of the configuration.
Unit	Millisecond with an integer value.
Standardization reference	
Significant	Mandatory.
Comment	This indicator has to be separated between call types (IP to IP, IP to PSTN, IP to mobile, etc.) for a detailed analysis. The objective set up in for universal telephony service has been set up to 2 900 ms in the French regulator recommendation.

7.2 Media establishment delay

Definition	Time determines on one of the two access of the communication, between off hock of the called and the beginning of voice signal receive.
Assessment method	Indicator determines sequentially from the two access of call configuration. On an IP access this indicator may be assessed by using a non-intrusive probe, such as a protocol analyser. Media establishment delay may be evaluated through the analysis of media flows and signalling. For ITU-T Recommendation H.323 protocol [i.17] the flow establishment delay corresponds to the time elapsed between the emission of the ITU-T Recommendation H.225.0 [i.18] "CONNECT" message and the arrival of the first IP packet including speech signal.
Unit	Millisecond with an integer value.
Standardization reference	
Significant	Optional.
Comment	This indicator has to be separated between caller and called site for a detailed analysis.

7.3 Unsuccessful call ratio

Definition	Ratio of unsuccessful calls to the total number of call attempts in a specified time period. An unsuccessful call is a call attempt to a valid number, properly dialled following dial tone, where neither called party busy tone, nor ringing tone, nor answer signal, is recognized on the access line of the calling user within 30 seconds from the instant when the address information required for setting up a call is received by the network.
Assessment method	Indicator determines sequentially from the two access of call configuration.
Unit	% with the resolution of 1 digit after the decimal point.
Standardization reference	ITU-T Recommendation E.800 [i.7], ITU-T Recommendation E.845 [i.8], EG 201 769 [i.9].
Significant	Mandatory.
Comment	The limit of 30 seconds is the default set-up of a timer in SS7 protocol.

7.4 Premature release probability

Definition	This indicator characterizes the ability to release a service. It is based on the measurement of the number of released communications in comparison with the number of established communications. Released communications are defined as communications released before voluntary action from one of the ends of the transmission.
Assessment method	
Unit	% with the resolution of 1 digit after the decimal point.
Standardization reference	ITU-T Recommendation E.800 [i.7].
Significant	Optional.