INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.1010

(07/2004)

SERIES P: TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

Transmission performance and QoS aspects of IP end-points

Fundamental voice transmission objectives for VoIP terminals and gateways

ITU-T Recommendation P.1010

ITU-T P-SERIES RECOMMENDATIONS

TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series	P.10
Subscribers' lines and sets	Series	P.30
		P.300
Transmission standards	Series	P.40
Objective measuring apparatus	Series	P.50
		P.500
Objective electro-acoustical measurements	Series	P.60
Measurements related to speech loudness	Series	P.70
Methods for objective and subjective assessment of quality	Series	P.80
		P.800
Audiovisual quality in multimedia services	Series	P.900
Transmission performance and QoS aspects of IP end-points	**Series**	**P.1000**

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation P.1010

Fundamental voice transmission objectives for VoIP terminals and gateways

Summary

This Recommendation provides 3.1 kHz telephony speech transmission performance requirements for the whole range of packet-based gateways and terminals, including wireless and softphones. Measurement methodologies are not covered by this Recommendation, however, work on this topic is under way in Study Group 12 and is planned to be incorporated in a future revision or, alternatively, in a separate new Recommendation. Also, requirements for wideband telephony may be added in a future version of this Recommendation.

Source

ITU-T Recommendation P.1010 was approved on 7 July 2004 by ITU-T Study Group 12 (2001-2004) under the ITU-T Recommendation A.8 procedure.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

ITU-T Recommendation P.1010

Fundamental voice transmission objectives for VoIP terminals and gateways

Scope

This Recommendation provides fundamental speech transmission performance requirements for 3.1 kHz handset telephony of packet-based terminals and gateways; it applies to both access gateways and trunking gateways. Requirements are given for handset operation, with the assumption that the corresponding requirements for headset and handsfree are obtained via appropriate conversions to take into account the different acoustic reference points.

This Recommendation addresses the whole range of IP-based gateways and terminals, including wireless and softphones.

Measurement methodologies are not covered by this Recommendation; however, work on this topic is under way in Study Group 12 and is planned to be incorporated in a future revision or alternatively in a separate new Recommendation.

1 Introduction

This Recommendation limits its scope to the fundamental speech transmission performance requirements for VoIP terminals and VoIP gateways with respect to the interface with the packet-based network. Thus, a number of traditional parameters have not been considered in this Recommendation. This is mainly because it is anticipated that the respective requirements from existing Recommendations, e.g., P.310 [11] or in the Q.55x series of Recommendations, can be applied to VoIP terminals and VoIP gateways in an appropriate way. Attention should also be paid to the end-to-end considerations provided in clause 5.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[1] ITU-T Recommendation G.100 (2001), *Definitions used in Recommendations on general characteristics of international telephone connections and circuits*.

[2] ITU-T Recommendation G.107 (2003), *The E-model, a computational model for use in transmission planning*.

[3] ITU-T Recommendation G.108 (1999), *Application of the E-model: A planning guide*, plus Amendment 1 (2003): *New Appendix I – The relationship between and interaction of talker echo and absolute delay*, and Amendment 2 (2004): *New Appendix II – Planning examples regarding delay in packet-based networks*.

[4] ITU-T Recommendation G.108.2 (2003), *Transmission planning aspects of echo cancellers*.

[5] ITU-T Recommendation G.109 (1999), *Definition of categories of speech transmission quality*.

[6] ITU-T Recommendation G.168 (2004), *Digital network echo cancellers*.

[7] ITU-T Recommendation G.711 (1988), *Pulse code modulation (PCM) of voice frequencies*.

[8] ITU-T Recommendation G.799.1/Y.1451.1 (2004), *Functionality and interface specifications for GSTN transport network equipment for interconnecting GSTN and IP networks*.

[9] ITU-T Recommendation G.1020 (2003), *Performance parameter definitions for quality of speech and other voiceband applications utilizing IP network,* plus Amendment 1.

[10] ITU-T Recommendation P.10 (1998), *Vocabulary of terms on telephone transmission quality and telephone sets*, plus Amendment 1 (2003): *New Annex A – List of psychoacoustic parameters*.

[11] ITU-T Recommendation P.310 (2003), *Transmission characteristics for telephone band (300-3400 Hz) digital telephones*.

[12] ITU-T Recommendation P.330 (2003), *Speech processing devices for acoustic enhancement*.

[13] ITU-T Recommendation P.340 (2000), *Transmission characteristics and speech quality parameters of hands-free terminals*.

[14] ITU-T Recommendation P.380 (2003), *Electro-acoustic measurements on headsets*.

[15] ITU-T Recommendation P.501 (2000), *Test signals for use in telephonometry*.

[16] ITU-T Recommendation P.502 (2000), *Objective test methods for speech communication systems using complex test signals*.

[17] ITU-T Recommendation Q.115.1 (2002), *Logic for the control of echo control devices and functions*.

[18] ITU-T Recommendation Y.1541 (2002), *Network performance objectives for IP-based services*.

[19] ANSI/TIA/EIA-810-A (2000), *Transmission Requirements for Narrowband Voice over IP and Voice over PCM Digital Wireline Telephones*.

3 Terms and definitions

Beside those in ITU-T Recs P.10 [10] and G.100 [1] this Recommendation defines the following terms:

3.1 send loudness rating (SLR): $\mathrm{SLR} = \mathrm{SLR(set)} + \sum_{i=1}^{n} \mathrm{CLR}_i$

3.2 receive loudness rating (RLR): $\mathrm{RLR} = \mathrm{RLR(set)} + \sum_{i=n+1}^{N} \mathrm{CLR}_i$

3.3 overall loudness rating (OLR): OLR = SLR + RLR

3.4 circuit loudness rating (CLR): The loudness loss between two electrical interfaces in a connection or circuit.

3.5 softphone: A softphone is an implementation of a VoIP terminal which is realized by a software program enabling a personal computer (PC) or personal digital assistant (PDA) equipped with an acoustic interface (microphone and earpiece/loudspeaker) to be used for two-way voice communications.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations:

CLR	Circuit Loudness Rating
IP	Internet Protocol
IPDV	IP Packet Delay Variation
OLR	Overall Loudness Rating
RLR	Receive Loudness Rating
RLR(set)	RLR of a telephone set
SCN	Switched-Circuit Network
SLR	Send Loudness Rating
SLR(set)	SLR of a telephone set
T	Mean One-Way Delay
TCLw	Weighted Terminal Coupling Loss
UNI	User-Network Interface
VAD	Voice Activity Detection
VoIP	Voice over Internet Protocol

5 End-to-end considerations

In order to achieve a desired end-to-end speech transmission performance (mouth-to-ear) it is recommended that transmission planning tasks be carried out with the E-model of ITU-T Rec. G.107 [2] as described in ITU-T Rec. G.108 [3]; this includes the *a priori* determination of the desired category of speech transmission quality as defined in ITU-T Rec. G.109 [5].

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 Rec. 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 category given in this Recommendation should be chosen to represent the specific configuration is the responsibility of the individual transmission planner.

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

6 Transmission characteristics

The requirements specified in this clause apply to both VoIP terminals and VoIP gateways with respect to the interface with the packet-based network as illustrated in Figure 1.

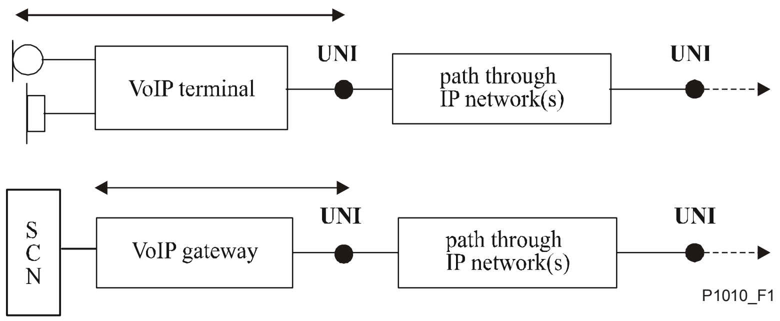

Figure 1/P.1010 – Area of applicability of this Recommendation

6.1 Default coding algorithm

VoIP terminals and VoIP gateways shall support the coding algorithm according to ITU-T Rec. G.711 [7] (both μ-law and A-law). VoIP terminals and VoIP gateways may support other coding algorithms.

6.2 Delay

While delay incurred from traditional terminals was significantly lower than the delay possibly incurred from the network, this is no longer true for VoIP terminals and for VoIP gateways.

Hence, for traditional terminals, the appropriate Recommendations (e.g., ITU-T Rec. P.310) were able to specify one single requirement for delay.

With the introduction of packet-based transmission, however, the increased delays introduced by the terminals and the gateways can have a tremendous impact on the end-to-end delay and, thus, need proper consideration in transmission planning. For example, depending on the delay variation accumulated by packets traversing an IP network, the design of the de-jitter buffer of an IP gateway or terminal is extremely critical and has to consider a trade-off between the permissible number of dropped packets and the maximum delay to be incurred (see Figure 3/G.1020 [9]).

Accordingly, for the parameter of delay, IP terminals and gateways may be designed to operate in a number of modes, with the mode selection being based on end-to-end considerations, especially desired speech quality level and likely network configurations.

For IP networks which comply with or which perform better than Class 0 of ITU-T Rec. Y.1541 [18], terminals or gateways that comply with any of the three delay categories given in this Recommendation will be a starting point to achieve acceptable speech quality.

For other kinds of networks, for example, networks with larger delay variations than any class of Y.1541 limits, only delay categories A or B may be a valid choice. Furthermore, there are cases such as the interconnection of two local IP islands via an intercontinental path through the PSTN, where only delay category A will be a valid basis for achieving a reasonable end-to-end speech transmission performance (due to the long delay of the PSTN path).

Irrespective of network configurations encountered, the deployment of some low bit rate codecs may further decrease the perceived speech quality and narrow the valid choice of delay categories.

The following subclauses specify three delay categories for either send or receive directions. It is the common understanding that, in order to comply with a specific delay category, a VoIP terminal or a VoIP gateway has to comply with both send and receive requirements of this category.

NOTE – For some VoIP terminals, e.g., softphones, it is well recognized that, under certain circumstances, technical and practical considerations may not permit any of the above categories to be met at this time. Therefore, this issue, as well as the introduction of a fourth delay category, are for further study.

6.2.1 Send delay categories (provisional values)

For a VoIP terminal, send delay is defined as the one-way delay from the acoustical input (mouthpiece) of this VoIP terminal to its interface to the packet-based network. VoIP gateway send delay is defined as the one-way delay from the electrical input (TDM or analogue) of this VoIP gateway to the interface to the packet-based network. In each case, the total send delay is the upper bound on the mean delay and takes into account the delay contributions of all of the elements shown in Figures 2 and A.1/G.1020 [9], respectively.

- Category A: Ts ≤ 20 ms;
- Category B: Ts ≤ 35 ms;
- Category C: Ts ≤ 50 ms.

6.2.2 Receive delay categories (provisional values)

For a VoIP terminal, receive delay is defined as the one-way delay from the interface to the packet-based network of this VoIP terminal to its acoustical output (earpiece). VoIP gateway receive delay is defined as the one-way delay from the interface to the packet-based network to the electrical output (TDM or analogue) of this VoIP gateway. In each case, the total receive delay is the upper bound on the mean delay and takes into account the delay contributions of all of the elements shown in Figures 3 and A.2/G.1020 [9], respectively.

- Category A: Tr ≤ 30 ms;
- Category B: Tr ≤ 65 ms;
- Category C: Tr ≤ 100 ms.

6.2.2.1 IP Packet delay variation (IPDV)

VoIP terminals and gateways should accommodate network IPDV values of up to 50 ms; this is consistent with ITU-T Rec. Y.1541 Classes 0 and 1. It is left to the designer's discretion whether the VoIP terminal or gateway is capable of accommodation of IPDV values above this value, which may occur in other networks.

6.3 Send Loudness Rating

For VoIP terminals, the recommended nominal value of Send Loudness Rating (SLR) should be:

- SLR = SLR(set) = 8 dB.

For VoIP gateways, the recommended nominal value of Send Loudness Rating (SLR) should be:

- SLR = SLR(set) + CLR = 8 dB.

For the gateway at the ingress to the IP network, the nominal SLR has to be achieved solely by means of transmission planning; this requires a proper dimensioning of the CLR.

6.4 Receive Loudness Rating

For VoIP terminals, the recommended nominal value of Receive Loudness Rating (RLR) should be:

- RLR = RLR(set) = 2 dB.

For VoIP gateways, the recommended nominal value of Receive Loudness Rating (RLR) should be:

- RLR = RLR(set) + CLR = 2 dB.

For the gateway at the egress from the IP network, the nominal RLR has to be achieved solely by means of transmission planning; this requires a proper dimensioning of the CLR.

6.5 Weighted Terminal Coupling Loss (TCLw)

Due to the increased delay values introduced by VoIP terminals compared to traditional digital terminals as per ITU-T Rec. P.310 [11], the TCLw requirement needs to be increased (see also ANSI/TIA/EIA-810-A [19]).

For VoIP terminals the recommended nominal value of weighted Terminal Coupling Loss (TCLw) should be:

- TCLw ≥ 55 dB

For further details see clause 10/P.310 [11]; in addition, the TCLw has to be maintained under all operational conditions. Relaxation of the TCLw value by not more than 10 dB is only permissible during double-talk periods. Note that this value is provisional and for further study.

NOTE – TCLw is strongly influenced by the amount of acoustic coupling in the acoustic interface. In the absence of any standardized requirements for this interface, the recommended value of TCLw may not always be met. Nevertheless, in the interest of achieving satisfactory speech quality, designers of VoIP terminals should strive to meet this value as closely as possible, taking into account the characteristics of the acoustic interfaces to be used.

6.6 Residual Echo Level

For VoIP gateways, an echo canceller according to ITU-T Rec. G.168 [6], is recommended for cancelling echoes that may originate in the PSTN. For details regarding control of Echo Cancellers see Annex A/Q.115.1 [17]. Details on the residual echo level that may be achieved and on transmission planning aspects for echo cancelling devices, are discussed in ITU-T Rec. G.108.2 [4]. Details on functional aspects of trunking gateways can be found in ITU-T Rec. G.799.1/Y.1451.1 [8].

6.7 Further parameters with respect to speech processing devices

For VoIP terminals and gateways that contain non-linear speech processing devices, the following parameters require additional attention and are for further study in the context of this Recommendation:

- objective evaluation of speech quality for VoIP gateways and VoIP terminals;
- double-talk capability;
- switching behaviour;
- partial echo effects;
- occurrence of artefacts;
- time-variant impairments.

As is the case for traditional telephone sets, reduced performance for handsfree and headset mode operation is anticipated for VoIP terminals. Further guidance in this area is provided by ITU-T Recs P.340 [13] and P.380 [14].

7 Measurement considerations

Measurement methodologies are not covered by this Recommendation; however, work on this topic is under way in Study Group 12 and is planned to be included in a future revision or, alternatively, in a separate new Recommendation.

Nevertheless, it can be assumed that VoIP terminals and gateways contain non-linear and time-variant functions (e.g., automatic gain control, adaptive jitter buffer, VAD). This needs to be considered when approaching the measurement of transmission characteristics. ITU-T Recs P.330 [12], P.501 [15] and P.502 [16] provide further guidance in this respect.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	**Telephone transmission quality, telephone installations, local line networks**
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks and open system communications
Series Y	Global information infrastructure, Internet protocol aspects and Next Generation Networks
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2004