ETSI EG 201 474 V1.1.1 (2000-04)

ETSI Guide

Speech Processing, Transmission and Quality Aspect (STQ); Future approaches to speech transmission quality across multiple interconnected networks

Reference

DEG/STQ-00004

Keywords

PSTN, transmission, interworking, quality, speech

ETSI

Postal address

F-06921 Sophia Antipolis Cedex - FRANCE

Office address

650 Route des Lucioles - Sophia Antipolis
Valbonne - 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

Internet

secretariat@etsi.fr
Individual copies of this ETSI deliverable
can be downloaded from
http://www.etsi.org
If you find errors in the present document, send your
comment to: editor@etsi.fr

Important notice

This ETSI deliverable 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 this 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://www.etsi.org/tb/status/

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 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://www.etsi.org/ipr).

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 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 Guide (EG) has been produced by ETSI Technical Committee Speech processing, Transmission and Quality aspects (STQ).

The present document is a natural progression from an earlier TM3/BTC2/TE4 study (ETR 275 [2]) into the main delay contributions caused by the various network elements in an end to end connection. It also incorporates inputs derived from the work on transmission planning in corporate networks carried out by CN7.

Introduction

The field of telecommunications is currently evolving faster than at any time in its history due to rapid advancements in technological progress and regulatory changes, and this potentially has a considerable effect on the end to end speech transmission quality of telephone calls. It may not be possible in future to guarantee specific voice transmission quality levels to users by relying just on the traditional network planning guides and codes of practice. The present document explores the possibility of introducing a dynamic mechanism, based on enhancements to the existing network signalling systems for assuring specific levels of voice transmission quality are achieved on individual calls.

1 Scope

The present document identifies trends in the design of networks and choices of technology that affect speech transmission quality and are expected to lead to a wider range of quality levels that will be experienced by users. The present document reviews the implications of the ITU definitions of speech transmission quality categories and identifies the main impairment parameters that need to be managed in order to provide specified levels of quality. The present document explores the options for exchanging information between interconnected networks by possible additions to the existing signalling mechanisms, but does not cover the decision/control functions which will be needed to provide the management of the transmission quality. The decision/control functions will be subjects for further study.

The present document contains an example of the possible content of a "Stage 1" type of description for the exchange of quality parameter information about individual calls.

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, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.
- A non-specific reference to an ETS shall also be taken to refer to later versions published as an EN with the same number.

[1] ETSI ETR 250: "Transmission and Multiplexing (TM); Speech communication quality from mouth to ear for 3,1 kHz handset telephony across networks".

[2] ETSI ETR 275: "Transmission and Multiplexing (TM); Considerations on transmission delay and transmission delay values for components on connections supporting speech communication over evolving digital networks".

[3] ETSI ETR 283: "Equipment Engineering (EE); Transient voltages at Interface A on telecommunications direct current (dc) power distributions".

[4] ETSI EG 201 050: "Speech Processing, Transmission and Quality Aspects (STQ); Overall Transmission Plan Aspects for Telephony in a Private Network".

[5] ETSI EG 202 086: "Speech Processing, Transmission and Quality Aspects (STQ); Objectives and principles for the transmission performance of multiple interconnected networks that aim to provide "traditional quality" telephony services".

[6] ETSI TBR 38: "Public Switched Telephone Network (PSTN); Attachment requirements for a terminal equipment incorporating an analogue handset function capable of supporting the justified case service when connected to the analogue interface of the PSTN in Europe".

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

[8] ITU-T Recommendation P.311: "Transmission characteristics for wideband (150-7000 Hz) digital handset telephones".

[9] ITU-T Recommendation G.107 : "The E-Model, a computational model for use in transmission planning".

[10] ITU-T Recommendation G.109: "Definition of Categories of Speech Transmission Quality".

[11] ITU-T Recommendation G.113: "Transmission impairments - Appendix I: Provisional planning values for the equipment impairment factor le".

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

[13] ITU-T Recommendation G.116: "Transmission Performance Objectives Applicable To End-To-End International Connections".

[14] ITU-T Recommendation G.131: "Control of talker echo".

[15] ITU-T Recommendation G.136: "Application rules for automatic level control devices".

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

[17] CCITT Recommendation G.711: "Pulse code modulation (PCM) of voice frequencies".

[18] ITU-T Recommendation Q.115: "Logic for the control of echo control devices".

[19] 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".

[20] ETSI ETR 004: "Business Telecommunications (BT); Overall transmission plan aspects of a private branch network for voice connections with access to the public network".

[21] ETSI ETS 300 283: "Business TeleCommunications (BTC); Planning of loudness rating and echo values for private networks digitally connected to the public network".

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

[23] ITU-T Recommendation P.561: "In-service non-intrusive measurement device - Voice service measurements".

[24] ITU-T Recommendation G.101: "The transmission plan".

[25] ITU-T Recommendation G.108.01: "Conversational impacts on end-to-end speech transmission quality - Evaluation of effects not covered by the E-model".

[26] ETSI ETS 300 415: "Private Integrated Services Network (PISN); Terms and definitions".

[27] ISO/IEC TR 14475: "Information technology - Telecommunications and information exchange between systems - Private Integrated Services Network - Architecture and Scenarios for Private Integrated Services Networking".

[28] ITU-T Recommendation P.861: "Test vectors for implementations of Perceptual Speech Quality Measure (PSQM)".

3 Definitions and abbreviations

3.1 Definitions

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

Private Network: in the present document the term "Private Network" is used to describe a network which provides features only to a restricted user group in contrast to a network available to the general public. In general, a private network is a terminating network and consists of several interconnected nodes (e.g. PBXs), with interconnections to other (mainly public) networks. A "private Network" is sometimes referred to as a "Corporate Network"

Public Network: in the present document the term "Public Network" is used for any network providing transmission and routing functions as well as features which are available to the general public, not restricted to a specific user group. In this context, the word "Public" does not imply any relation to the legal status of the network operator

From the point of view of an end-to-end connection, a public network can function either as a "Transit Network" (a link between two other networks) or as a combination of "Transit- and Terminating Network" e.g. where the public network provides connections to terminal equipment such as telephone sets (terminating network), or PBXs (transit network)

Voice Quality of Service: measure of the end to end quality of speech transmission of a connection. Strictly speaking it should be a statistical value representing the aggregate performance of a number of connections, but within the present document it is often used in relation to the performance of a single connection. The voice quality literally applies to an end to end connection including the terminals i.e. from mouth to ear, but since the terminals are often no longer the responsibility of the network operators, the term VQoS is often used to refer to the performance of a connection from NTP to NTP

3.2 Abbreviations

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

ALC	Automatic Level Control
ANF	Additional Network Feature
ATM	Asynchronous Transfer Mode
CNG	Comfort Noise Generator
DCME	Digital Circuit Multiplication Equipment
EC	Echo Canceller
ISDN	Integrated Services Digital Network
ISUP	ISDN Signalling User Part
NTP	Network Termination Point
PDV	Packet Delay Variation
QoS	Quality of Service
VAD	Voice Activity Device (or Detector)
VQoS	Voice Quality of Service

4 Background

The changing face of telecommunications resulting from the twin driving influences of technical progress and regulatory change is capable of exerting a considerable effect on the end to end speech transmission quality of telephone calls. Traditionally the impairments affecting speech transmission quality were dominated by loss, delay, echo and distortion. With the introduction of digital techniques into switching and transmission the impairments associated with loss and noise were significantly reduced, and were replaced by other impairments such as quantisation distortion together with increased delay. The main impairment became transmission delay and the associated echo. This was reviewed in ETR 275 [2] which identified the delay contributions of the various network elements and also gave references to sources of guidance on delay allocation and echo control. ETR 275 [2] was seen as a start to the process of maintaining satisfactory speech transmission quality over evolving networks. This process is summarized in figure 1 which is taken from ETR 275 [2].

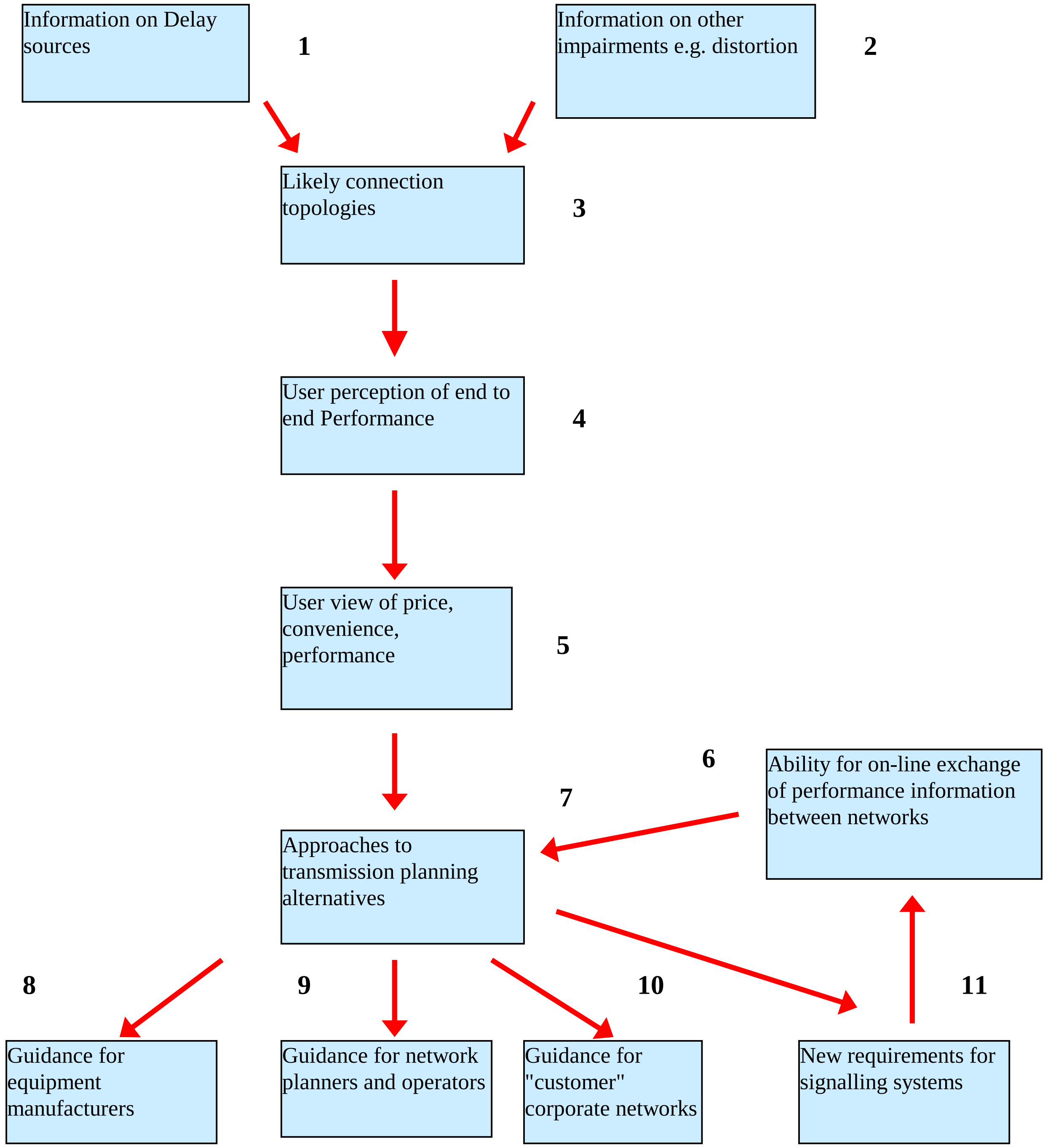


Figure 1: Transmission planning factors and their inter-relationships

1 The provision of an accurate and comprehensive set of information on the delay contribution of various technologies was the main purpose of ETR 275 [2].

2 No work on distortion is currently in progress, although it is addressed to some extent in ETR 250 [1].

3 No document has yet been produced. New reference models are needed. This work is linked to item 7.

4 This item has been addressed fully in ETR 250 [1].

5 No coherent set of information is currently available, although sources of information are increasing through marketing surveys of user views of some mobile and corporate systems. The model contained in ETR 250 [1] includes an expectation factor that could be used.

6 and 11 The ability of the currently defined signalling systems is limited, new requirements for new facilities may be identified. This is investigated further in the present document.

7 This activity is covered in the present document.

8 Some guidance is provided in various ETSs.

9 Guidance on User Network Interface-User Network Interface (UNI-UNI) performance and interconnection arrangements is urgently needed. Guidance hitherto has been based on the public-private model and the 25 ms Rule M of ITU-T Recommendation G.131 [14].

10 Some guidance based on the old public-private network model has been produced in ETR 004 [20] and ETS 300 283 [21]. ETR 004 [20] was reviewed initially by BTC 2, and the task completed by CN7 as EG 201 050 [4] (May 1998 and reviewed by STQ in February 1999).

5 General

5.1 Factors affecting speech transmission quality

- **Evolving technology:** as new technology is introduced into the network elements it often brings with it new types of impairment or new variants of existing types of transmission impairment. This can have the effect of diluting or rendering less significant some of the more traditional impairments such as loss and noise, and thus making existing planning guidance less effective or even obsolete. Another important issue is the speed of introduction of new technology into the telecommunications field. The increasing pace of technological advance shows no signs of slowing down, and in the telecommunications field is spurred on by the increased competition resulting from recent regulatory changes - particularly in Europe.

- **More complex topology:** the increased competition in provision of telecoms is evident in the proliferation of public network operators and service providers. These new operators are providing alternative facilities in both the local and trunk networks. Similarly, the private networks have also been evolving, often leading the introduction of new technology, sometimes changing from data only networks to carry both voice and data, and sometimes international in reach. When considering the growing use of virtual private networks and features such as intelligent (low cost) routing together with the proliferation of networks, it becomes virtually impossible to predict the route and hence performance of an end to end connection. *In such an environment, the responsibility for the quality of service of any particular connection becomes increasingly obscure.*

- **Less control:** the process of liberalization whilst providing the spur to the competition which is providing greater choice in the provision on networks and services also tends to weaken control of speech transmission quality by reducing regulatory requirements and minimizing standards. This situation is exacerbated by the fact that many of the new network operators lack experience and expertise in network performance planning compared to the established ex-national monopoly operators.

5.2 Assessment of speech transmission quality

- **Subjective methods:** in the final analysis, the speech transmission quality of a telephone connection is a subjective matter and can only be measured by subjective methods. These can be used to determine the performance of different aspects of speech transmission quality such as conversation flow, or articulation and may be performed by expert or naïve users according to the requirements of each test. When carried out in a proper manner according to well established principles such as those defined in ITU-T Recommendation P.800 [22], they are capable of providing reasonably consistent results. They are generally more time consuming and more expensive to perform than objective tests.

- **Objective methods:** objective methods of assessing speech transmission quality are potentially capable of providing less expensive and more immediate results than subjective methods. There are no direct objective measurements of speech quality. Two main approaches are used but both rely on objective measurement followed by interpretation by some kind of model. The first method measures the objective performance of (the transmission impairments of) the various elements in an end-to-end connection and combines them in a transmission model such as the E model defined in ETR 250 [1] and incorporated in ITU-T Recommendation G.107 [9]. The second, newer method objectively measures speech samples entering and leaving the voice telephony connection and processes them in a perceptual model of human hearing. Both methods ultimately need to be validated/calibrated against the more traditional subjective methods. More information on these methods may be found in EG 201 377-1 [19].

- **Application to automatic systems:** for any future system where the end-to-end speech transmission quality across multiple networks is controlled dynamically (by flexible distribution/allocation of impairments), the first objective method described above which relies on a transmission model would appear to be the more promising candidate and can certainly be linked more closely into the network planning process.

5.3 Control of speech transmission quality

- **Traditional experts:** until comparatively recently the technical responsibility for the performance of the worlds telephone network has rested almost entirely on the national monopoly network operators. The very success of the world-wide telephone network is largely due to the knowledge, skills, and experience of these operators, and was achieved virtually without regulatory standards or requirements because of this sound technical base.

- **Impact of liberalization:** the advent of liberalization has had a major impact on the telecoms scenario. Many new network operators have been set up and have been the major influence in driving down costs and providing a wider choice of services. Some of these new network operators (and/or their suppliers) have limited knowledge or experience in the issues relating to network planning, or are working to short term commercial pressures which has resulted in their networks not performing as well as those of the more established operators. In order to ensure satisfactory operation it has been found necessary to produce a wide range of guides, voluntary standards, codes of practice, and interconnect agreements.

- **Current inadequacies:** we have seen that the introduction of many more network operators and a changing regulatory environment has resulted in a scenario where the possible route taken by a single call has a multiplicity of options possibly with different speech performance quality. Similarly the introduction of new technology (which is probably accelerated by the forces of liberalization) is modifying the distribution of transmission impairments or introducing new impairments. This in turn requires revision of the various guides, voluntary standards, codes of practice, and interconnect agreements etc. to ensure some kind of performance is maintained. It is likely that a point will be reached where it is not possible to keep control by this approach, and some automatic system, possibly based on enhancements to the network signalling systems, will be required.

5.4 Voice band traffic

The worlds telephone networks have traditionally been designed for voice telephony and have been optimized for speech performance. With the rapid growth in voice band data traffic it is essential that the requirements for end-to-end data performance as well as speech performance are fully considered.

Items to be considered for further study include:

- by-passing of fax and modem traffic;

- re-modulation of voice band data.

5.5 Automatic speech transmission quality system

An automatic voice quality system may be defined as "a system which uses (enhanced) signalling means to provide a required level of end-to-end speech transmission quality of a connection or to provide a measure of the actual end-to-end speech transmission quality level of a connection".

An automatic system for handling the transmission impairments could provide some or all of the following benefits:

- Dynamic optimization of transmission quality during call set-up;

- Choice of speech transmission quality to the user;

- Choice of speech transmission quality provision to the supplier;

- Billing information based on quality of service of individual calls;

- Better tracking of the effect of technological changes on network performance;

- Less impact on speech transmission quality of network topological changes;

- Less documentation maintenance required e.g. guides, codes of practice, interconnect agreements etc. do not have to be revised each time some new codec type or other new impairment source is introduced into the network infrastructure.

5.6 Transmission impairments

5.6.1 Types of transmission impairment

There are numerous factors which impact the overall speech quality of a connection due to imperfections in the transmission chain. These factors are known as transmission impairments. Considerable information exists on impairment factors in such references as ETR 250 [1], ITU-T Recommendations G.108.01 [25], G.113 [11] etc.

5.6.2 Trends in transmission impairments

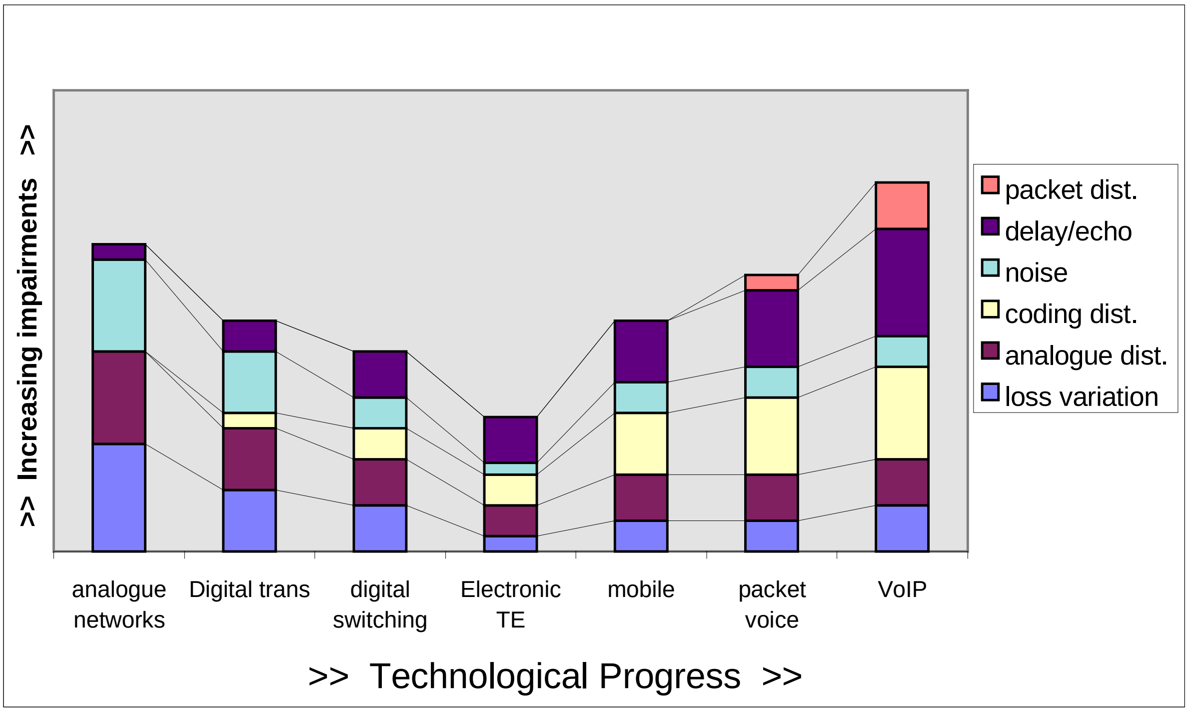


Figure 2: Trends in end-to-end transmission impairments

Figure 2 shows the trends in the various transmission impairments which affect the end to end speech transmission quality of a telephone call. The diagram is intended to be illustrative rather than quantitative so the vertical axis should not be assumed to be a direct measure of the end to end quality.

NOTE 1: The introduction of new technology into a network does not necessarily replace earlier technology, but for economic reasons may coexist or interwork with it for some considerable time.

NOTE 2: The column shown as "Electronic TE" represents the current yardstick or benchmark for calls via the fixed PSTN.

NOTE 3: The term "packet distortion" is intended to be a generic term covering cell and packet loss and jitter.

The first column is indicative of the situation where trunk transmission was analogue, switching was mostly electromechanical in nature e.g. Strowger, and where telephones were passive devices using carbon microphone transmitters, and consequently there were reasonably high levels of analogue noise and distortion and loss variation. The next column shows the effect of introduction of digital transmission which largely removes the loss variation and noise due to analogue transmission but introduces some coding distortion and some additional delay. The introduction of digital switching significantly reduces the noise due to electromechanical (analogue) switching, but adds additional delay.

The next column shows the performance improvement resulting from the introduction of electronic telephones. Instead of passive transmission circuits these use active circuits where the impedance, send and receive gains, etc. are more accurately controlled by feedback. Furthermore the carbon microphone was replaced by a dynamic or electret device further reducing noise and variation in sending sensitivity. The combination of electronic telephone with digital transmission and switching represents the benchmark for current toll quality end-to-end speech transmission quality.

The introduction of digital mobile significantly increases both delay and coding distortion compared to the benchmark, whilst the (often) far from optimum mobile handset geometry further increases the noise and loss variation. The introduction of packet voice technology such as ATM in the networks further increases delay due to packetisation delay and jitter removal. Similarly the introduction of Voice over the Internet (VoIP) is likely to increase impairments such as delay, (particularly if software encoding/decoding is performed e.g. by a personal computer), packet distortion, and possibly loss variation since there are currently no voice performance standards for VoIP terminals.

It can be clearly seen that there has been a steady reduction of the typical "analogue" impairments of noise, distortion and loss variation. This has been largely due to the introduction of digital transmission systems, digital switching, and electronic subsets. Conversely it can be seen that there has more recently been a steady increase in "digital" impairments such as quantisation distortion, delay, and jitter so that the end to end quality of service is actually beginning to worsen. This trend will continue at a more rapid pace as new technologies and services are introduced, spurred on by the pressures of liberalization. The effects of these new impairments are not fully known and their effect in tandem which can occur in a multiple network environment is often unpredictable.

6 Signalling requirements to support end-to-end speech transmission quality

It is a tacit assumption that any "automatic" system for handling the distribution of transmission impairments will be based upon additional requirements on the network signalling systems. In order to define these additional requirements a number of issues need to be resolved. A number of the key issues are discussed in this section.

6.1 Network configuration

All networks may be regarded as either terminating networks or transit networks for the purposes of assessing the end to end voice quality of a connection. A single network may simultaneously provide the functionality for both terminating and transit networks for a multiplicity of connections, but may only have the characteristics of a transit or terminating network for each individual connection.

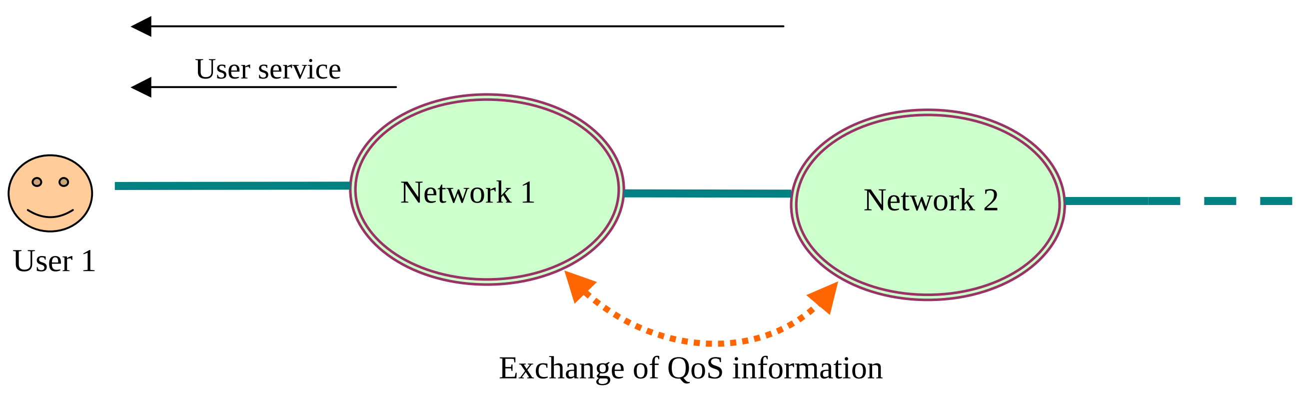


Figure 3: Basic network interconnection

Figure 3 shows User 1 connected to Network 1 (a terminating network), which is interconnected to Network 2 (a transit network) which is in turn interconnected to another terminating network either directly or via one or more additional transit networks to complete a connection.

In order for the end to end speech transmission quality of the overall connection to be predicted, it is necessary for interconnected networks to exchange information on the transmission impairments introduced within the associated networks. This exchange of information can be either dynamic (which makes use of the networks signalling capabilities), or static (which requires the inclusion of transmission impairment information into the network operators interconnect agreements). The exchange of information may also include forwarding of information received from another interconnected network e.g. type of codec used so as to enable the use of tandem free operation, or the accumulated delay time.

The different networks are likely to use different technologies such as circuit – switched, radio, and VoIP, and different signalling systems such as ISUP, QSIG and DSS1.

The networks may need to exert some control on the (potential) connection depending on the received information e.g. to prevent unnecessary tandeming of Voice Activity Detection (VAD) devices.

The above scenario is independent of the ownership of the networks i.e. it does not distinguish between private (corporate) and public networks.

6.2 Use of VQoS information

Besides the obvious interest of the designers and manufacturers of telecommunications equipment, there are two main parties concerned with the use of speech transmission quality information i.e. the end user (subscriber) and the network operator.

6.2.1 Impact on network operators

It has been established (EG 202 086 [5]) as a general principle that the operator who is charging for the call has overall responsibility for the quality of service of a call. This operator will suffer the time and expense of dealing with any complaints even if the cause of poor speech transmission quality is external to his network. Consequently network operators will only want to interconnect with other networks which can reliably provide connections with acceptable performance. This acceptable performance may encompass a range of speech transmission quality levels. It is envisaged that the QoS parameters, including those relating to speech transmission quality, will be incorporated into Interconnect Agreements. These agreements may specify the control of speech transmission quality dynamically by means of signalling, statically by defining statistical limits for the relevant parameters e.g. packet loss will be below 2 % for 98 % of the time, or by a combination of both dynamic and static means. Any statistical limits could be objectively monitored e.g. according to ITU-T Recommendation P.861 [28] or non-intrusive methods e.g. to ITU-T Recommendation P.561 [23], to assure performance.

NOTE: In Europe the operator has as a rule no choice whether to interconnect or not. Due to the ONP-Framework, public telecommunications network operators are forced to interconnect by law. There are no strict regulations concerning quality of service agreements.

6.2.2 Impact on end users

The other main reason for generating and signalling speech transmission quality information is to provide the user with a choice of speech transmission quality categories, most probably with different costs. This can be regarded as providing a supplementary service to the user. Selection of the VQoS category could be either on a subscription basis or call by call selection. The latter approach is a supplementary service and will require the generation of a service description.

NOTE: This implies there is a "basic" service with a guaranteed minimum quality of service.

In order to provide universally agreed levels of speech transmission performance, it is necessary to have three key components, namely; information on the accumulated impairments of a connection; some control mechanism which uses the impairment information; and well defined voice VQoS categories.

6.3 Speech transmission quality categories

In order to ensure consistent speech transmission quality categories for connections traversing multiple networks it is essential to have universally accepted quality categories which are well defined and can be both measured and predicted. Such categories for narrow band (3,1kHz) speech are now available in ITU-T Recommendation G.109 [10]. The categories are defined in terms of 'R-values', details of which may be found in annex B.

The speech transmission quality categories may be provided by a network operator either as guaranteed speech transmission quality levels or on a "best effort" basis. The main differences are shown in the following table. It should be understood that the 'guaranteed' speech transmission quality levels are not absolute guarantees of quality, but statistical probabilities (which may have very high values). It may prove necessary to monitor the speech transmission quality levels realized in practice in order to substantiate the guaranteed speech transmission quality levels.

Table 1: Provision of speech transmission quality categories

Best efforts basis	Guaranteed quality levels
• Estimate of speech transmission quality of each call is not essential • User choice of intermediate network operators taking into account cost and quality • Minimum speech transmission quality design target should be provided (R > 50 recommended) • Signalling of speech transmission quality information not essential	• Estimate of speech transmission quality of each call is essential • Automatic selection of intermediate network operators to achieve requested quality • Quality limits for each speech transmission quality category defined • requires signalling of quality/impairment information • requires decision making/control mechanisms within network

6.4 User strategies and issues

For users connected to networks offering guaranteed quality levels an increased range of choices become available. For example:

- Should the user subscribe to a specific speech transmission quality category for all his calls, or should the user select the category on a call by call basis?
- How many categories should be offered to a user; and
- How many can he be expected to cope with?
- Should selection be based on the speech transmission quality or the cost?
- What happens if these parameters change with time?

It is outside the scope of the present document to attempt to answer most of these questions. In fact most of these issues should be left to individual network operators to decide so as to encourage originality and competition in the provision of these features. One point becomes clear however, and that is that a new supplementary service will be required to allow the user to select outgoing calls based on a combination of quality and cost.

ETSI TIPHON currently specify four different quality categories. There is an alternative view that probably two speech transmission quality categories are sufficient and anymore would confuse most users. This is substantiated to some degree by some recent field trials of VoIP systems which indicated that there seems to be a minimum quality threshold below which people will not use however low the cost. It is recommended that no call attempt should be aborted because the requested quality level can not be achieved. Such calls should be completed on a "best efforts" basis.

6.5 Signalling parameters

Table 2 summarizes which transmission impairments need associated signalling parameters in order to support end to end speech transmission quality. It is based upon annex A which contains an evaluation of the various transmission impairments affecting end to end speech transmission quality. The annex also identifies for each parameter, whether any signalling parameters are needed to support a dynamic VQoS system or whether the parameter is likely to be adequately controlled by ITU-T Recommendations, or Codes of Practice.

Table 2: Transmission impairments needing signalling parameters

Category	Impairment type	Comment
No signalling parameters required	• loss • attenuation distortion • group delay distortion • circuit noise • Automatic level control	Can be expected to be adequately controlled by correct adherence to existing transmission plans (e.g. ITU-T Recommendation G.101 [24]) and codes of practice.
Signalling parameters required	• Mean one way transmission delay • Codec type (identifier) • Codec impairment factor (I_e) • Packet loss • Packet delay variation • VAD control • EC control • User choice of speech transmission quality category 'R' • Expectation factor	(cumulative counter required) (for tandem free operation TFO) (cumulative counter required) (could be included in Ie) (parameter to be defined) (simple counter required) (see ITU-T Recommendation Q.115 [18] for parameters) (numeric value) (could be included as an offset to 'R')

In order to ensure network signalling systems will be able to support end to end speech transmission quality by providing transfer of performance parameters on a per call basis, the signalling parameters identified in the above table will need to be included in a service description for an Additional Network Feature (ANF). An example of a possible draft ANF is included as annex C.

In order to provide the required speech transmission quality for individual calls, a mechanism will be required for using the various signalling parameters to predict the call quality in order to make any appropriate control decisions. A method based on a simplified version of the "E-model" basic equation may prove suitable.

NOTE: It is envisaged that the definition of dynamic speech quality control mechanisms and algorithms will be the subject of further work items.

6.6 Interim issues

Few networks will initially be capable of signalling the necessary parameters needed to support dynamic voice quality establishment. It seems logical to focus initially on those types of new networks which currently have no established VQoS signalling capability e.g. IP based networks. It seems most appropriate that these are the very networks which are driving the concept of provision of multiple speech quality levels.

Where not all the networks involved in a single call can support VQoS signalling, use should be made of "default" parameters to represent the attributes of the non conformant networks.

Annex A (informative): Transmission impairment

A.1 Transmission impairment types

The transmission impairments identified in this section all have an audible effect on end to end speech transmission quality. Some of the impairments are generated by analogue networks or sections of a network; some are introduced by digital network components, and some by both analogue and digital. With the evolution towards all digital networks some impairments associated with analogue network components need no longer be considered for inclusion in any automatic speech quality system.

For each of the transmission impairments identified in this annex a judgment is required as to whether any signalling information needs to be transported for (dynamic) QoS call control during call set-up. For those impairments where signalling is needed, it is also necessary to identify whether the signalled parameter is needed

a) on a cumulative basis e.g. to establish the end to end delay of a connection by summing the delay of each element in the connection; or

b) on a transparent end to end basis e.g. to avoid unnecessary transcoding when the same type of codec is used at both ends of a connection.

A.1.1 Loss

The loss (attenuation) of a mouth to ear connection is a critical factor for satisfactory speech communication., and is normally taken care of by existing transmission planning guides such as G.108.01 [25]. As most new networks employ digital transmission throughout it should be relatively straightforward to maintain a zero loss throughout the digital domain. The overall loss is then determined almost entirely by the sending and receiving terminals. The loudness ratings of most wired and mobile terminals are generally acceptable as they meet standards such as TBR 38 [6], P.310 [7] etc. However no standards currently exist for the electroacoustic performance of VoIP terminals and this should be addressed. No new signalling parameters are required to transmit loss information about the terminating and transit networks.

A.1.2 Attenuation Distortion

The overall attenuation distortion (frequency response) of a connection is a function of the number of translations to and from voiceband that occur within the end-to-end connection, together with the frequency shaping introduced by analogue (baseband) transmission sections. It may be safely assumed that for predominantly digital connections the attenuation distortion characteristics will meet the Q.550 series Recommendations (or equivalent) and consequently require no specific signalling parameter to be introduced for automatic speech transmission quality use.

The attenuation distortion characteristics of terminals are not generally the responsibility of the network operator, but terminals are normally designed to acceptable standards such as TBR 38 [6], ITU-T Recommendations P.310 [7], P.311 [8] etc. The one area which is not currently covered is VoIP terminals, and this needs to be addressed.

A.1.3 Group delay distortion

Like the attenuation distortion, the overall group delay distortion of a connection is a function of the number of translations to and from voiceband that occur within the end-to-end connection, and similarly will require no specific signalling parameter to be introduced for dynamic control of speech transmission quality.

A.1.4 Circuit noise

All digital connections are likely to have an acceptable noise performance which is substantially independent of connection length and therefore requires no specific signalling parameter for dynamic control of speech transmission quality.

A.1.5 One-way transmission time

The greater use of digital transmission and digital speech processing in modern networks results in higher values of one-way transmission times. When echo is present, this delay has a major effect on the speech transmission quality, and consequently a significant impact on the performance requirements and positioning of echo cancellers. ITU-T Recommendation G.114 [12] provides guidance in this matter.

A.1.6 Echo

When the one-way transmission time exceeds 25 ms it is generally recommended that an echo canceller is provided. ITU-T Recommendation G.168 [16] gives information on the performance requirements for echo cancellers, while ITU-T Recommendation G.131 [14] provides guidance on the deployment of echo cancellers in the networks.

Traditionally, in medium sized countries, where the mean one-way delay did not exceed the 25 ms, echo cancellers have been inserted in the International Switching Center for international connections, only. As long as one administration had the control of the whole public network in that country it was relatively straightforward to ensure that the parameters of the echo cancellers in the International Switching Center were properly adjusted to the echo tail path in the national network.

By contrast, in a liberalized environment, where multiple network operators may provide their services in a single country, the mean one-way delay might exceed the 25 ms in certain cases and hence it is necessary to check the properties of the echo canceller with the parameters of the echo tail path for each connection. This may require a parameter to represent the sum of the delay in the tail from the last inserted forward Half Echo Canceller (HEC).

In general there are currently no limits to the number of echo cancellers which may be connected in tandem in a single connection, but there are recommended limits for echo suppressors.

Furthermore, some echo cancellers also use a center clipper to remove any residual echo and this is often accompanied by a comfort noise generator to reduce the noise contrast of switching between residual echo and silence. There are often subjective impairments introduced by such devices particularly when connected in tandem, and some signalling parameter associated with echo control devices may be needed to support dynamic control of this aspect of speech quality. More recently, results presented in ITU-T Study Group 12 have shown degradation in connections equipped with tandemed ECs (source: Deutsche telekom et. al.), and echo suppressers are no longer recommended by the ITU. Further study on the use of echo cancellers in tandemed situations may be required in order to optimize the transmission performance of such connections, and this may result in the need for additional signalling parameters to provide effective monitoring and controlling of the echo control process.

A.1.7 Speech clipping

This impairment relates to temporal clipping introduced as an unwanted side effect by the voice switching function of devices such as DCME equipment or mobile radio access such as GSM. The primary purpose of such devices is to save transmission bandwidth. With the continuing reduction in the intrinsic cost of provision of transmission bandwidth e.g. by optical fiber it might be expected that the use of voice switching in the fixed (wired) networks would also be reducing. However due to the current high price of leased transmission bandwidth, many of the networks currently being installed still make significant use of voice switching to reduce leasing costs. With mobile, the reason for voice switching to reduce bandwidth utilization is the finite limitation of available bandwidth in the radio spectrum. This pressure is most unlikely to diminish. Voice switching is also likely to be widely used for transmission of voice over IP networks. Temporal clipping is an inherent problem with voice switching and is due to the finite time it takes to distinguish the start of a talk burst from the ambient background noise. This process is repeated in tandemed links and can lead to complete loss of the first part of the speech burst with subsequent poor subjective performance. Provision of a signalling parameter associated with the speech clipping function should be considered. Guidance on the use of speech clipping can be found in ITU-T Recommendation G.116 [13], (and further study is currently being undertaken by Study Group 11).

A.1.8 Comfort noise

The use of voice switching generally introduces unpleasant noise contrast effects to the listener since the ambient background noise transmitted during the speech burst disappears during the gaps between speech bursts. This unpleasant noise contrast effect can be reduced to some extent by introducing comfort noise into the receiving transmission path after the voice switching equipment during the gaps between speech bursts. However there are still audible differences between the real background noise and the injected comfort noise. These differences may be of level, or of spectral content. If there is a difference in level between the real background noise and the injected comfort noise then the noise contrast can increase on tandemed connections. This may require the provision of a signalling parameter to control.

A.1.9 Codecs

All digital codecs introduce some impairment to the transmission path, the actual amount being dependent on the coding algorithm used.. In general the more complex the algorithm used, the lower the bandwidth required, and the greater the impairment introduced. Until recently the impairment was expressed in quantisation distortion units (qdu) – different values being assigned to each code type, with the individual qdus being summed to provide a rating for the total end-to-end connection. With the introduction of lower bit rate codecs this method was found unacceptable and a system of equipment impairment factors (I_e) introduced. The quantisation distortion units were retained for ITU-T Recommendation G.711 [17] pcm codecs and digital pads. Provisional equipment impairment factors for various codecs are given in ITU-T Recommendation G.113 [11] Appendix 1. In addition, codecs introduce delay into the transmission path. which depends primarily on the frame length and look-ahead time as described in clause 7 of ITU-T Recommendation G.108.01 [25].

Codecs may be used in the terminal equipment as in ISDN or GSM terminals and also within the network itself e.g. ADPCM used in DCME equipment. At best the impairments introduced by tandemed codecs are additive, but for some combinations of codec types additional interactions may occur resulting in greater impairment than expected. If both the sending and receiving terminals use the same codec type, there is no intrinsic need to perform intermediate transcoding and a straight through digital connection can be made with a minimum of impairments. This tandem free operation (TFO) requires specific capabilities from the end-to-end signalling to implement. Where tandem free operation is not possible, some enhancement of the signalling is required in order to accumulate information on the accumulated codec distortion of the network elements making up the end-to-end connection.

A.1.10 Packet delay variation

For IP networks the main impairments to speech quality are packet delay, packet delay variation (PDV) and packet loss. The packet loss and delay variation may be reduced by increasing the size of dejitter buffer, but at the expense of increased delay. Some designs of equipment dynamically change the size of the dejitter buffer according to the percentage packet loss detected in order to minimize the delay. The subjective effects of such a technique are not yet understood.

Recent investigations on IP network characteristics has shown that a small percentage of the total packets may be delayed significantly greater than the mean delay by a single network device such as a firewall, with a significant adverse effect on speech quality. It is recommended that each IP network element should have a well controlled PDV defined by both maximum packet delay and percentage PDV greater than packet size (see annex D).

A.1.11 Packet loss

The equipment impairment factors assigned to codecs are initially based on their performance in the wired network. When used in IP networks the codecs are likely to be operated under conditions where packet loss is present and this will increase the amount of impairment introduced. Furthermore different codec types will introduce different amounts of impairment for the same degree of packet loss. Many of the codecs will therefore require further characterization to establish their performance under packet loss conditions.

The widely used G.711 [17] codec is a low complexity device which introduces minimum delay and distortion under normal conditions. However due to its real time nature (i.e. the value of each sample is independent of the preceding or succeeding samples), it is particularly prone to generating unpleasant sounds as a result of any unwanted packet loss introduced when used over IP networks. New recovery algorithms are currently appearing which minimize the subjective effects of lost packets. These Packet Loss Concealment Algorithms (PLCA) act only on the receive side of the codec pair and hence do not have to be standardized. It would be useful to have a flag of some kind supported by the signalling systems to indicate the availability or unavailability of any Packet Loss Concealment Algorithm. Provisional information on the Equipment Impairment Factors of some codecs under various packet loss conditions, with and without PLCA are given in revised Appendix 1 of G.113 [11] September 1999.

A.1.12 Automatic level control

Current ITU-T thinking is that properly designed networks (e.g. those in accordance with ITU-T Recommendation G.101 [24] The Transmission Plan) do not require the use of ALC devices, and that such devices should not be used in lieu of proper network design and implementations. If an ALC is used however, it should only be used on connections equipped with active echo cancellers, but should not be placed in the tail path of the echo canceller. ALC devices should not be used in tandem.

Rules for the application of ALC devices may be found in ITU-T Recommendation G.136 [15].

A.1.13 Expectation factor

Although the speech transmission quality provided by any telephone connection is determined by the totality of the impairments introduced by the individual components of the connection, the judgment of acceptability of the speech transmission quality by the users may be influenced by other factors. For example, a user may be prepared to accept a lower quality to gain the advantage of mobility, or of paying a lower charge. In order to take account of external influences when predicting the user perceived acceptability of a voice connection, an additional factor has to be introduced. This is commonly known as the "Expectation Factor". This factor has been incorporated in the E-model defined in ITU-T Recommendation G.107 [9]. It should be noted that the effect of the Expectation Factor can be positive or negative, and that it may vary with time.

Table A.1 summarizes the signalling requirements for the various types of impairment.

Table A.1: Impact of impairments on signalling requirements

Impairment type	Comment	Cumulative	end-to-end
Loss	Counter for digital pads? – not required if preferred loss plans are used.	-	-
Attenuation Distortion	Primarily a TE attribute – no signalling parameter required	-	-
Group delay distortion	Not a major issue for speech – no signalling parameter required	-	-
Circuit noise	Not a major issue – no signalling parameter required	-	-
One-way transmission time	Sum of delay from call originating point (terminal) to actual point	✓	-
Echo	Indicator of ECANs already deployed	✓	
Speech clipping (temporal)	(excessive) tandeming of VAD devices should be avoided. Provide VAD counter.	✓	-
Comfort noise	This parameter may be linked to VAD	✓	
Codec (I_e)	Counter for accumulation of Ie-values of equipment used in the transmission path	✓	-
Codec (TFO)	TE codec identifier needed to permit tandem free operation	-	✓
Packet loss	Likely to be a significant factor in IP based networks	✓	
Packet Loss Concealment Algorithm	Indicator for the availability of PLCA is likely to be required where G.711 [17] is transmitted via IP networks.	✓	
Packet delay variation	Impairment effects not yet fully appreciated.	✓	-
Automatic level control	Counter for ALC devices? – not required if preferred loss plans are used.	-	-
Expectation factor		-	✓

Annex B (informative): Speech transmission quality ratings

The impairments identified in the previous section all contribute in varying degrees to the end-to-end speech transmission quality of a connection. In any automatic (dynamic) system for controlling speech quality it is necessary to have some universally acceptable criteria for various quality levels. A classification based on the speech transmission quality perceived by the user, which has been incorporated into ITU-T Recommendation G.109 [10] and has now been adopted by ETSI TIPHON is summarized in the following table.

Table B.1: Definition of categories of speech transmission quality

R-Value Range (see note 1)	Speech Transmission Quality Category	User satisfaction
90 < R < 100	Best	Very satisfied
80 < R < 90	High	Satisfied
70 < R < 80	Medium	Some users dissatisfied
60 < R < 70	Low	Many users dissatisfied
50 < R < 60	Poor	Nearly all users dissatisfied

NOTE 1: The overall rating "R-Value" is a rating of the mean speech transmission quality calculated using the E-Model described in G.107 [9]. It has the advantage of providing a linear scale of perceived quality.
NOTE 2: Connections with R-values below 50 are not recommended in G.109 [10].
NOTE 3: The "R-Value" is a rating of the unidirectional speech path from the talkers mouth to the listeners ear. It does not include the effects of impairments associated with 2-way conversational speech such as the double-talk performance of an echo canceller as such types of impairment have yet to be fully characterized.

Annex C (informative): Example ANF service description

Title: **Transfer of performance parameters on a per call basis**
End-to end speech transmission performance
Additional Network Feature
Service Description

C.1 Scope

Due to the fact of increasing numbers of consecutive interconnected networks, the aim of the present document is to help to ensure end-to-end performance of a call. In a first step the scope of the present document is limited to end-to-end speech transmission performance, only.

The present document specifies the Additional Network Feature (ANF) which enables the transfer of performance parameters between networks (private or public) on a per call basis in order to activate or deactivate specific network elements or to influence the routing of a particular call. This ANF is applicable to any call between any type of circuit or packet switching digital network like ISDN, PISN or digital mobile network as well as to multimedia, B-ISDN and Internet.

Any user related actions and any post processing of the values are outside the scope of the present document. However, the present document may form the basis for further supplementary services (SS).

The present document covers the stage 1 description of the ANF-PERF according to ITU-T Recommendation I.130 [6].

C.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, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.
- A non-specific reference to an ETS shall also be taken to refer to later versions published as an EN with the same number.

[1] ITU-T Recommendation E.438: "Performance parameters and measurement methods to assess N-ISDN 64 kbit/s circuit switched bearer service UDI in operation".

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

[3] ITU-T Recommendation G.113: "General Characteristics of International Telephone Connections and International Telephone Circuits Transmission Impairments".

[4] ITU-T Recommendation H.323: "Multiplexing protocol for low bit rate multimedia communication".

[5] ITU-T Recommendation H.245: "Control protocol for multimedia communication".

[6] CCITT Recommendation I.130: "Method for the characterization of telecommunication services supported by an ISDN and network capabilities of an ISDN".

[7] ITU-T Recommendation I.350: "General aspects of quality of service and network performance in digital networks, including ISDNs".

[8] ITU-T Recommendation I.380: "Internet protocol data communication service - IP packet transfer and availability performance parameters".

[9] ITU-T Recommendation I.570: "Public/private ISDN interworking".

[10] ITU-T Recommendation Q.115: "Logic for the control of echo control devices".

[11] ITU-T Recommendation Q.762: "Signalling System No. 7 - ISDN user part general functions of messages and signals".

[12] ETSI EG 201 050: "Speech Processing, Transmission and Quality Aspects (STQ); Overall Transmission Plan Aspects for Telephony in a Private Network".

[13] ETSI EN 300 171: "Private Integrated Services Network (PISN); Specification, functional models and information flows; Control aspects of circuit-mode basic services [ISO/IEC 11574 (1994) modified]".

[14] ETSI ETR 250: "Transmission and Multiplexing (TM); Speech communication quality from mouth to ear for 3,1 kHz handset telephony across networks".

[15] ISO/IEC 11579-1: "Information technology - Telecommunications and information exchange between systems - Private integrated services network - Part 1: Reference configuration for PISN Exchanges (PINX)".

[16] ETSI EG 201 474: "Speech Processing, Transmission and Quality Aspect (STQ); Future approaches to speech transmission quality across multiple interconnected networks".

[17] ETSI ETR 076: "Integrated Services Digital Network (ISDN); Standards guide".

[18] ETSI ETS 300 415: "Private Integrated Services Network (PISN); Terms and definitions".

[19] ISO/IEC TR 14475: "Information technology - Telecommunications and information exchange between systems - Private Integrated Services Network - Architecture and Scenarios for Private Integrated Services Networking".

C.3 Definitions and abbreviations

C.3.1 Definitions

For the purpose of the present document the following terms and definitions apply:

Additional Network Feature (ANF): see ETR 076 [17]

An Additional Network Feature (ANF) is a capability, over and above that of a basic service, provided by an ISDN, but not directly to an ISDN user

Intervening Network (IVN): see ISO/IEC 11579-1 [15]

The generic term for any real type of network which is employed for the provision of inter-PINX connections

Public Network: see EG 201 050 [12]:

The term "Public Network" is used in the present document for any network providing transmission and switching functions as well as features available to the general public, not just to a specific user group. In this context, the word "Public" does not imply any relation to the legal status of the network operator.

In some cases, a public network may provide only a limited set of features. In a competitive environment, a public network may be restricted to serve a limited number of customers, or restricted to specific features or functions. Generally, public networks provide access points to other networks or terminals only within a specific geographical area.

From the point of view of an end-to-end connection, a public network can function either as a "Transit Network"(a link between two other networks) or as a combination of "Transit- and Terminating Network" in case where the public network provides connections to terminal equipment such as telephone sets, or PBXs. In North America, inter-exchange carriers (IC) generally function as transit networks while the functions of a transit- and terminating- network are assigned to local exchange carriers (LEC).

Private Network: see EG 201 050 [12]:

In the present document the term "Private Network" is used to describe a network which provides features only to a restricted user group in contrast to the public network (PSTN) available to the general public. In general, a private network consists of several interconnected nodes (i.e., PBXs), with interconnections to other (mainly public) networks. The following is a list of private network characteristics which may impact the overall transmission quality of an end-to-end connection. A Private Network is characterized as follows:

1) It consists normally of more than one element of switching equipment (PBX or Key Telephone System KTS), connected via tie trunks or leased lines or via a Virtual Private Network (VPN). Network functionality is independent of its structure and hierarchy. Switching equipment and links can be either analogue or digital.

2) It provides switching functions and all other features only to a single customer or a group of customers, and is not accessible to the general public.

3) It is not limited by geographical size or to a specific national area or region.

4) It has no limitation with regard to the number of extensions and access points to other networks.

Supplementary Services: any services provided by a network in addition to its basic service or services

Terminal Equipment (TE): see ETS 300 415 [18]

An item of equipment attached to a telecommunication network to provide access for a user to one or more services.

C.3.2 Abbreviations

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

ECD	Echo Control Device
PISN	Private Integrated Services Network, see EN 300 171 [13]
PINX	Private Integrated Services network eXchange, see EN 300 171 [13]
CN	Corporate Network, see ISO/IEC TR 14475 [19]

C.4 Specification of the ANF

C.4.1 Description of the ANF

C.4.1.1 General

ANF-PERF is an additional network feature which enables the exchange and the update of performance parameters of consecutive connection parts of the specific call.

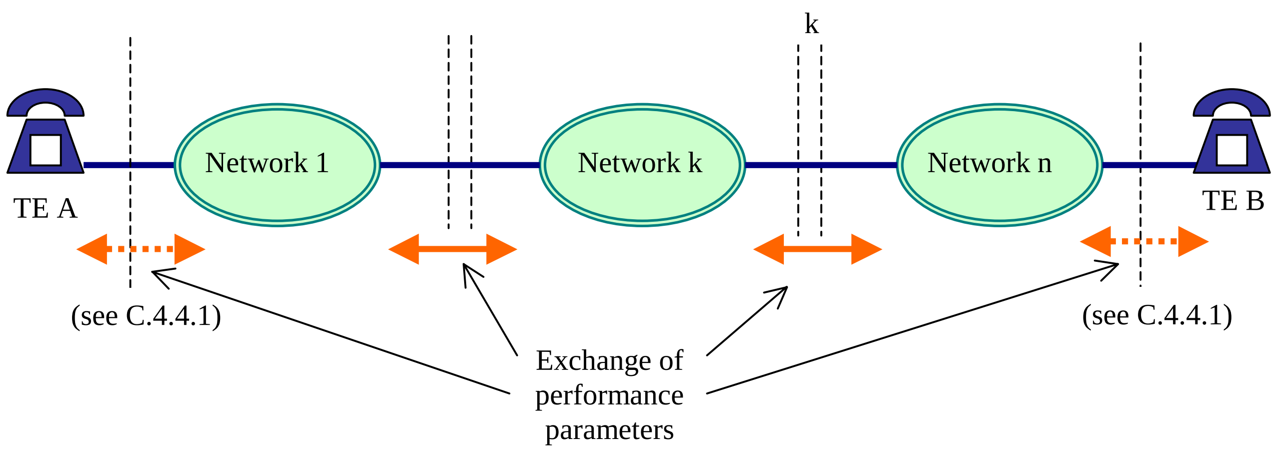


Figure C.1: Exchange of parameters across network

The performance parameter values are set to an initial value by the call originating network (private or public) that initiates the call. It takes into account the values of TE A.

The values are updated upon each network boundary crossed depending on the transmission facilities used for that connection.

In the backward direction, the performance parameters are set to an initial value by the call terminating network (private or public) considering the values of TE B. These parameter values are updated in the same way as above.

C.4.1.2 Applicability

This ANF is focused on all connections that are used for voice communication, which are usually used in a conversational (two way) mode.
Other applications may need only one way communication in which some parameters may be less relevant (e.g. delay). As networks are not aware of the detailed mode, all voice calls are treated in the same manner.

A call should be identifiable by each interconnecting network as a voice call.

The applicability to other type of calls (e.g. data or multimedia) is for further study.

C.4.2 Parameters

All parameters which are relevant for the user information transfer function (see ITU-T Recommendation I.350 [7]) should be considered.

As a first step the following performance parameters are exchanged (as indicated in EG 201 474 [16]):

- Accumulated one-way transmission time at the interface k in the forward direction Daccf(k)
This parameter is used in ITU-T Recommendation Q.762 [11] as propagation delay counter.
It is also used in ITU-T Recommendation E.438 [1] as OWPD (mean One Way Propagation Delay).

- Accumulated one-way transmission time at the interface k in the backward direction Daccb(k)
This parameter is also accumulated in the forward direction.

NOTE 1: It is for further study if it may be assumed, that the one way transmission time in both directions is equal.

NOTE 2: With the parameters above the network k may calculate the Tail delay in the backward direction:
Dtailb(k) = Daccf(k)+Daccb(k).

- Total one-way transmission time in the forward direction Daccf(tot)
This parameter will be used by the originating network for the estimation of the ear-to-ear quality.

NOTE 3: With the parameters above the network k may calculate the Tail delay in the forward direction:
Dtailf(k) = Daccf(tot)-Daccf(k)+Daccb(k).

- Echo control device indicator (ecdi) (as defined in ITU-T Recommendation Q.115 [10]).
- Incoming half echo control device request indicator (as defined in ITU-T Recommendation Q.115 [10]).
- Incoming half echo control device response indicator (as defined in ITU-T Recommendation Q.115 [10]).
- Outgoing half echo control device request indicator (as defined in ITU-T Recommendation Q.115 [10]).
- Outgoing half echo control device response indicator (as defined in ITU-T Recommendation Q.115 [10]).
- Equipment Impairment factor (as defined in ITU-T Recommendation G.113 [3], App.I and ETSI ETR 250 [14]).
- Voice activity detectors (VAD).
- Codec type (necessary for tandem free operation).
- Counter for ALC-devices and/or digital pads inserted.

Possible parameters for further applications:

- charge information;
- QoS, categories of speech transmission quality (ITU-T Recommendation G.109 [2]);
- comfort noise level;
- expectation factor;
- G.711 packet loss concealment capability.

C.4.3 Parameter Values

The values of the exchanged parameters are assumed as mean values. The networks involved in a call are responsible for the validity of the calculation, measurement or estimation of the updated values of performance parameters.

The values may be time-dependent, e.g. different in the busy hours.

If appropriate, the mean values may be added, i.e. the parameters may be assumed as statistically independent.

The value of the parameters may be different for both directions of the connection.

C.4.4 Procedures

C.4.4.1 General

The ANF is always activated and used with every call set-up.

The impairments of the transmission line between two networks are taken into account by the originating network.

The updated values are stored at least for the duration of a call.
In addition they may be kept for any post processing by separate feature request. This action will be done by management e.g. for statistics or required by a regulation. It is out of the scope of the present document.

The recipient network during call establishment shall accept the incoming call and handle it in a best effort way.

The attempt to provide guaranteed transmission quality levels is for further study.
In such a case, detailed negotiation procedures would be required. Calls might even have to be released when the required quality values can not be achieved. In addition, charging information depending on the resulting level would also be very desirable.

Changes of the parameter values during the call are not considered.

This service may also be offered to the terminal equipment. This may be realized on a subscription basis and could be considered as a supplementary service.

C.4.4.2 Procedures

The procedures depend on the parameters:

- Propagation delay counter:
  Accumulation of the received value
  In addition the routing may be dependent on the received value, e.g. by choosing preferred paths with low delay if the received value is relatively high.

- VAD:
  The value of this parameter is binary. It is set if VAD is used.

- Type of codecs:
  This parameter is used to support tandem free operation (TFO).
  It shall be used in a compatible way to H.323 [4] respectively H.245 [5].

- Parameters for Echo control:
  The logic is based on ITU-T Recommendation Q.115 [10]. Examples of the signalling procedures see ISUP99.
  In addition the tail delay shall be taken into account as described in EG 201 474 [16].

C.4.5 Interaction with other Supplementary Services and ANFs

Only the supplementary services or ANFs with interactions are indicated, in particular where a new destination network is involved:

C.4.5.1 Call diversion services (SS-CFU, SS-CFB, SS-CFNR)

If a call is forwarded to another destination, the values are updated according to the new transmission path. Further study is required.

C.4.5.2 Call Transfer (SS-CT)

The values are recalculated according to the new transmission path. Further study is required.

C.4.5.3 Path Replacement (ANF-PR)

The values may need to be recalculated according to the new transmission path. Further study is required.

C.4.6 Interworking considerations

C.4.6.1 Intervening Network

A part of a connection may be an intervening network. The intervening network may be a leased line, a satellite link, a part of an ATM network with a PVC, an IP based network as examples.
The network that is responsible for the routing of a call to an intervening network should have knowledge of the performance values for that part of a connection. This will enable the network to update the performance values and to send them to the next network that handles the signalling.

C.4.6.2 Interworking with an analogue network

An indication shall be inserted, that no end-to-end digital path for user information is available.

C.4.6.3 Interworking with a digital network, that does not support this ANF

An indication shall be inserted, that the ANF is not supported all the way.

C.4.6.4 Interworking with a mobile network

No special requirement.
The mobile network should indicate that an echo control device is included (see ITU-T Recommendation Q.115 [10]), and should insert impairment factors according to the codec type used.

C.4.7 Overall SDL diagram

Figure C.2 shows an overall SDL diagram of ANF-PERF. Input/output symbols represent stimuli from/to basic call control.

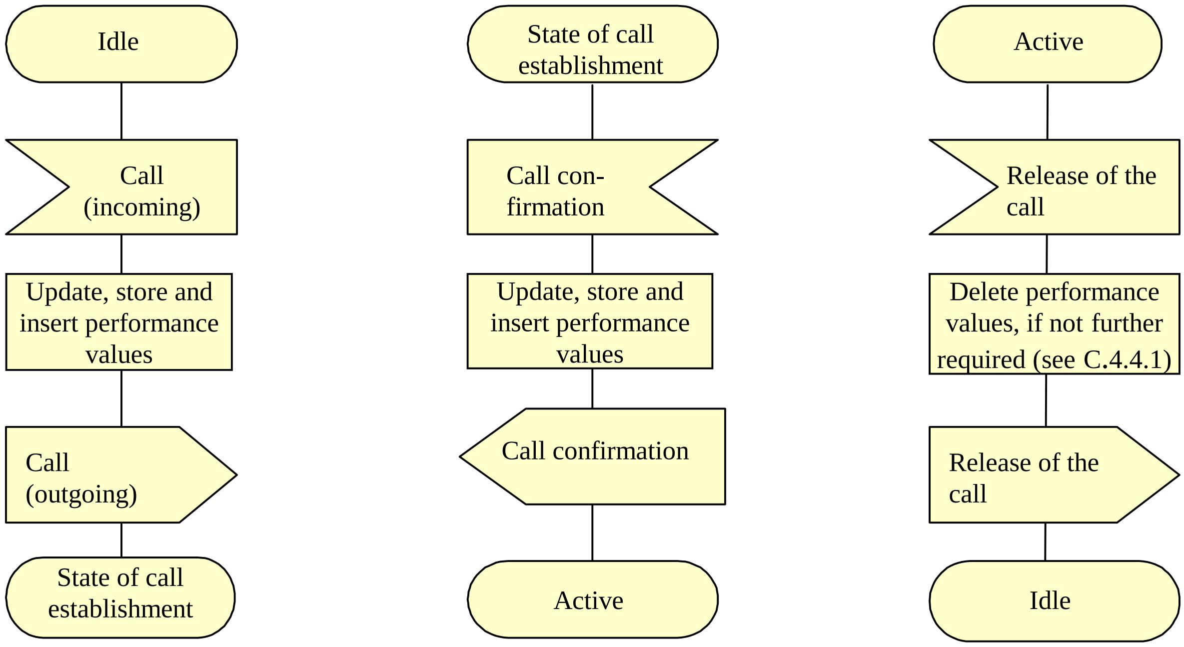


Figure C.2: Overall SDL diagram

C.4.8 Interworking with different signalling protocols

Figure C.3 shows an example of how some of the impairment parameters may be signalled /used during set-up of a call which spans more than one network domain.

NOTE: For simplicity, messages to control echo canceller devices are not included.

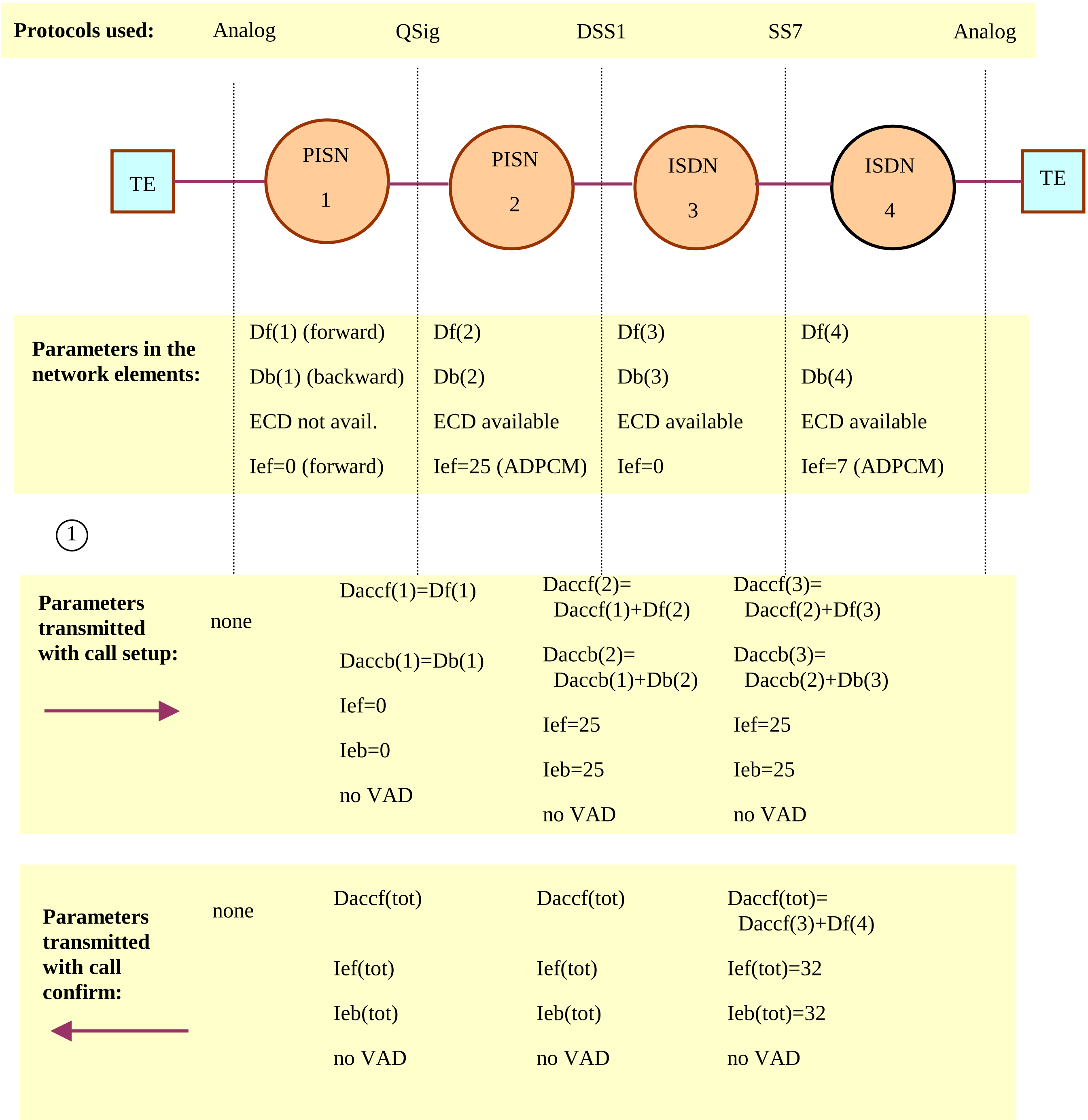


Figure C.3: Example of the use of transmission impairment parameter messages

Annex D (informative): QoS - IP performance vs. Speech Quality

D.1 Introduction

This annex describes a real effect of packet delay variation (PDV) on end-to-end speech quality, and illustrates that a mean PDV measure is inappropriate for defining a network quality of service (QoS) level. The annex describes the measurement process used to identify the effect and suggests more suitable PDV statistical measures.

The work described in this annex relates aspects of IP transmission performance to end-to-end speech quality for the purpose of VoIP network planning. The annex details the most significant findings and demonstrates how the use of multidisciplinary techniques has lead to faster analysis and clearer understanding of the problem.

The work has been based on a series of tests that were performed on VoIP technologies across a corporate Intranet – the scenario used for this report is shown in figure D.1. The system was tested at 5-minute intervals throughout the day, with each test lasting 115 seconds. The tests consisted of a number of measurements, these included:

- Echo Return Loss
- Round Trip Delay
- Loss
- Listening Quality Prediction

During these tests the IP traffic stream was captured at both ends of the network connection for post analysis.

The test scenario shown in figure D.1 included:

- 64kbit/s mu-Law PCM
- No silence suppression
- Continuous packets with no losses
- 30ms frames with one frame in a packet
- 10Mbit/s Ethernet LAN
- Two geographically remote sites

* PIG (PSTN to Internet Gateway)

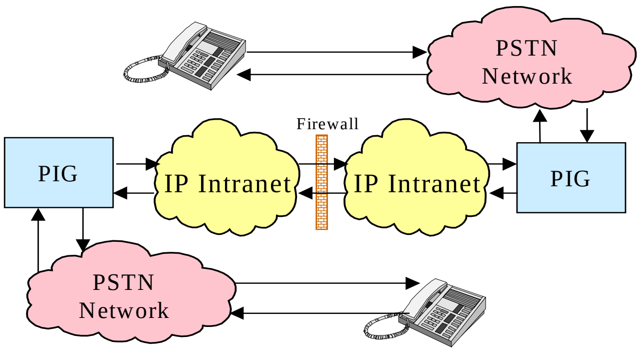


Figure D.1: Test Network Architecture

D.2 PDV effect on Speech Quality

For IP networks it is readily understood that the mechanisms effecting VoIP end-to-end conversational quality, in the absence of packet loss, are packet delay and packet delay variation. Packet delay affects the end-to-end transmission delay, while the PDV can affect the end-to-end delay and the perceived speech quality; since large packet delay variations can exceed jitter buffer dimensions and result in packets being dropped. Packet delay is easy to define for maintainable QoS but what parameters are required to define PDV to ensure a maintainable QoS.

To investigate the impact of PDV on perceived speech quality the captured IP traffic stream was analysed and correlated with the output of BT's Perceptual Analysis/Measurement System (PAMS).

D.2.1 IP Packet Analysis

Analysis of the traffic source showed that the PIG packets were generated constantly every 30 ms. Therefore any PDV can be associated purely with the IP network. As expected the packets arriving at the receiver exhibited packet delay variation.

Listed in table D.1 are simple measures used to statistically describe the PDV at the receiver.

Table D.1: PDV Measurements

PDV Measure	Value	PDV Measure	Value
Mean	30 ms	5 percentile	-8 ms
Maximum	107 ms	95 percentile	7 ms
Minimum	-31 ms	1 percentile	-29 ms
%age PDV greater than frame size	0,7 %	99 percentile	13 ms

Mostly these delay variations are small, however, the graph in figure D.2 clearly shows that a small percentage of the total packets have been delayed approximately 100 ms greater than the mean.

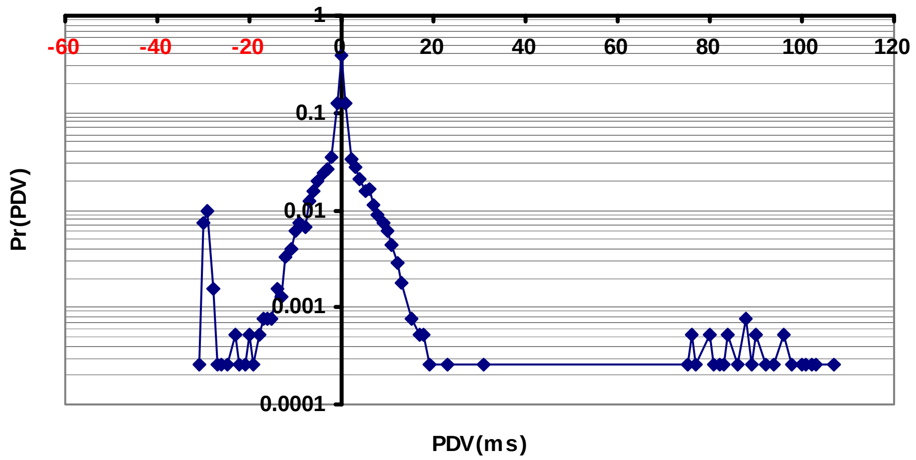


Figure D.2: PDF of packet delay variation

Further analysis shows that the excessively delayed packets occurred periodically throughout the test, as shown in figure D.3.

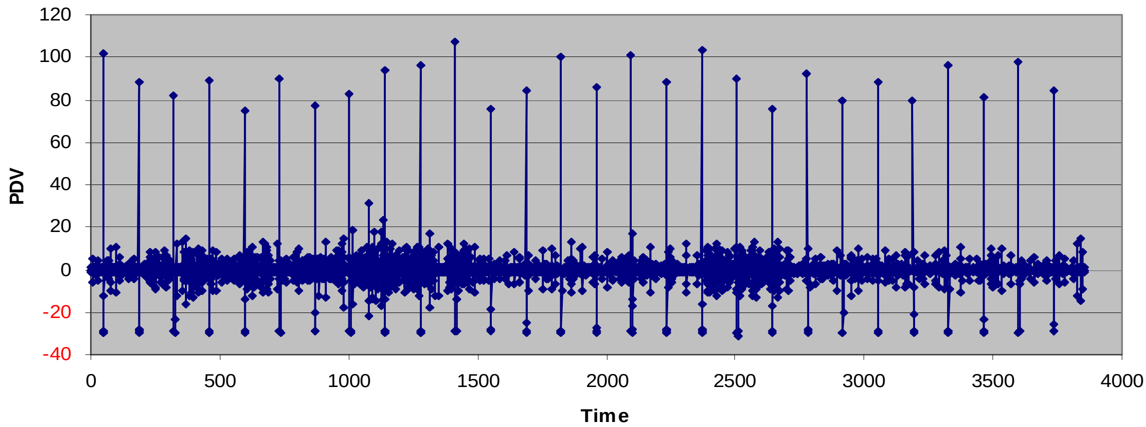
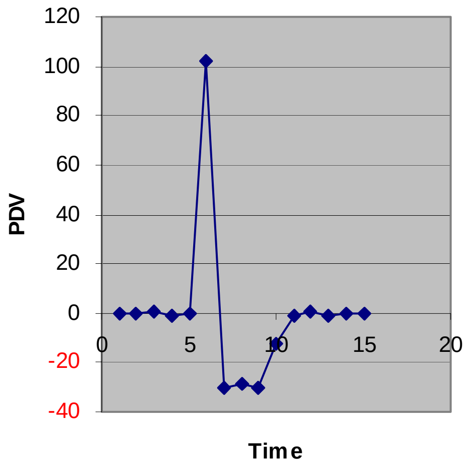


Figure D.3: Time Series Plot of Packet Delay Variation

The occurrence of the excessively late packets is periodic, every 130 packets or 4 seconds. Also three early packets follow each late packet, as shown in figure D.4.



This event equates to ~1% of packets being very late and approximately ~3% being grouped and filling up the buffer. The network is evidently buffering the packets while it performs some housekeeping task, as the delays are not due to a route change or congestion. The effect has been traced to the Firewall. To date Firewalls are typically tested for data throughput and not delay performance. The delay occurs while the Firewall verifies that the IP address to which the data is being transmitted is still a valid address and has not been hijacked. The Firewall stores the arriving packets while it performs the check and then transmits all stored packets in succession.

Figure D.4: Time Series Close Up

A question arises as to whether this effect is significant on end-to-end speech quality. In fact by using PAMS, as described in subclause D.2.2, this mechanism is shown to have a significantly adverse effect on speech quality. The mechanism could easily be avoided with minimal security risk for the protected network. The measures, in table D.1, that are most indicative of this problem are the "Maximum PDV" and "%age PDV greater than packet size".

D.2.2 PAMS Analysis

PAMS (Perceptual Analysis/Measurement System) is a speech quality assessment algorithm designed for end-to-end assessment of telephone-bandwidth voice transmission systems. The structure is shown in figure D.5 below.

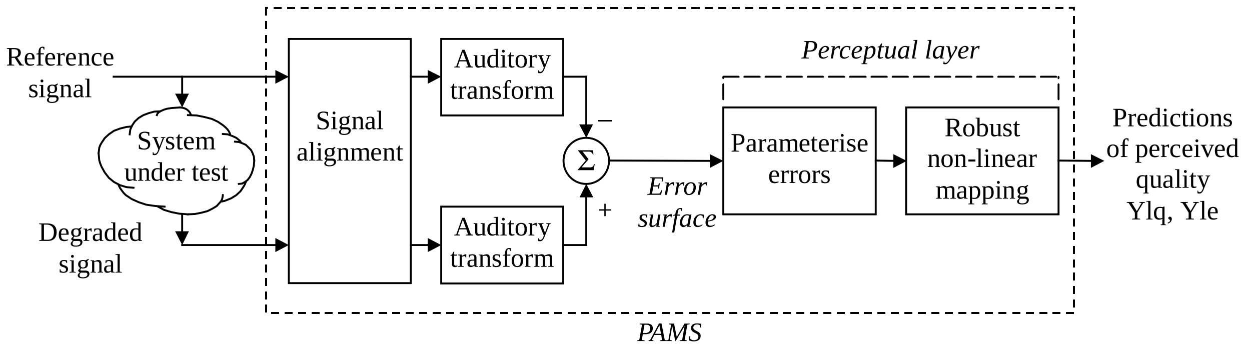


Figure D.5: PAMS Structure

It was anticipated that the excessively delayed packets identified in subclause D.2.1 would result in a reduction of speech quality since the application layer typically drops late packets. PAMS was used to quantify this effect and the study is summarized in the remainder of this section.

Of the 115-second test sequence, the PAMS assessment was performed for 30 seconds in each direction. The signal comprised of a pseudo conversational speech-like test sequence, consisting of twelve separate speech utterances, shown in the top-half of figure D.6.

For the complete days testing the mean listening quality prediction for the connection was 3,0, equating to fair on the standard ITU five point listening quality scale. The test speech in this investigation achieved a listening quality score of 3,1 and can therefore be regarded as representative of the connection.

As shown in figure D.5, PAMS subtracts the reference auditory transform from degraded auditory transform to produce an error surface; a two dimensional data structure describing the signal error in terms of time, frequency band and magnitude. Figure D.6 shows the error surface for the test case aligned with the degraded speech signal. Markers are placed at four-second intervals, the period identified in subclause D.2.1 for the excessively delayed packets, and can be seen to align with the major errored speech segments.

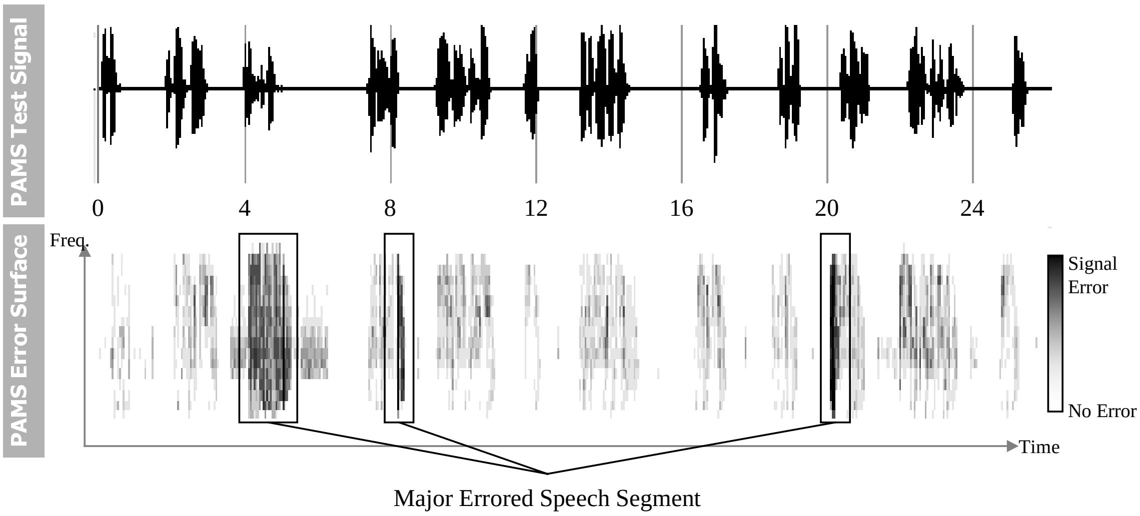


Figure D.6: PAMS Test Signal aligned with Error Surface Output

An alternative 3D representation (figure D.7), of the same error surface, clearly illustrates that the majority of error in the speech signal is associated with these 4-second intervals

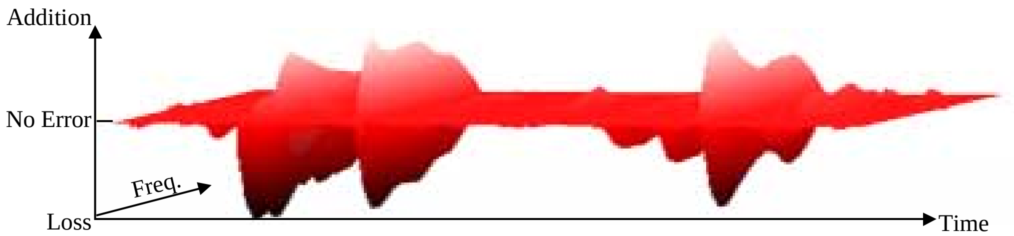


Figure D.7: 3D Error Representation - Illustrating Error Magnitude at 4s intervals

Figures D.6 and D.7 demonstrate that although small errors can be seen, due to coding and analogue stages, the dominant degradation is associated with the periodic late packets. The degradation is quantifiable by reconstructing the recorded IP packet stream, with a simulated "infinite" playout buffer and again assessing with PAMS. This allows for a comparison to be made between the best possible speech quality (no late/dropped packets) and the speech quality achieved by the tested system. This technique is currently being developed.

The degradation can be approximately quantified by removing the speech events suffering from the extreme PDV effect. This is practical because only three of the twelve utterances align with the four-second events and suffer from large errors. The remainder of the extreme PDV events occurs in silence intervals. When the three degraded utterances are no longer assessed the speech quality prediction increases from 3,1 to 4,1 – an improvement from fair to good.

D.3 Conclusion

An extensive series of performance test were conducted with VoIP technology across a corporate IP Intranet. This annex concentrated on end-to-end speech quality and PDV measures and, in particular, the relationship between them.

The IP packet analysis confirmed that the PIG packet generation was constant. Therefore the corporate IP network introduced all PDV. A small amount of PDV is expected with any IP network, due to queuing, however the network periodically exhibited excessively large PDV. This effect has been traced to a Firewall.

The PAMS analysis has shown that the large PDV significantly impacts on the speech quality. The connection quality drops from good to fair. It should be noted that this measure is of listening quality and does not include the effects of delay and echo, dominant factors in perceived conversational quality.

The work has clearly identified a strong benefit in using both IP packet based analysis and the emerging objective speech quality measures to assess VoIP technologies. The PAMS analysis highlighted and quantified the level of degradation introduced by the VoIP system and the IP packet analysis allowed the problem to be diagnosed.

For VoIP network planning, it has been demonstrated that there needs to be tightly defined PDV specifications for each network element. Although large PDV may occur with route changes, a single network element, such as a Firewall or Proxy Server, should not cause this effect. To ensure large PDVs are not introduced by single network elements, each element should be tested for "maximum packet delay" and "percentage PDV greater than packet size" and not just throughput.

History

Document history		
V1.1.1	February 2000	Membership Approval Procedure MV 200015: 2000-02-15 to 2000-04-14
V1.1.1	April 2000	Publication