INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

P.561

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

(02/96)

TELEPHONE TRANSMISSION QUALITY

OBJECTIVE MEASURING APPARATUS

IN-SERVICE, NON-INTRUSIVE MEASUREMENT DEVICE – VOICE SERVICE MEASUREMENTS

ITU-T Recommendation P.561

(Previously "CCITT Recommendation")

FOREWORD

The ITU-T (Telecommunication Standardization Sector) is a permanent organ of the International Telecommunication Union (ITU). The 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 Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1 (Helsinki, March 1-12, 1993).

ITU-T Recommendation P.561 was prepared by ITU-T Study Group 12 (1993-1996) and was approved under the WTSC Resolution No. 1 procedure on the 6th of February 1996.

NOTE

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

© ITU 1996

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the ITU.

CONTENTS

SUMMARY

This Recommendation provides specifications for voice service transmission measurement devices that provide measurements on an in-service and non-intrusive basis. These In-Service, Non-intrusive Measurement Devices (INMDs) are utilised primarily for the measurement of voice-grade parameters such as speech level, noise level, echo loss and speech echo path delay.

INMDs may also be used to measure parameters associated with digital transmission systems that impact the performance of the voice-grade channels being transported. This Recommendation specifies interface, measurement range, and accuracy requirements for measuring voice-grade transmission parameters as well as descriptions of optional functions associated with these parameters.

IN-SERVICE, NON-INTRUSIVE MEASUREMENT DEVICE – VOICE SERVICE MEASUREMENTS

(Geneva, 1996)

1 Scope, purpose, and application

1.1 Scope

This Recommendation provides specifications for voice service transmission measurement devices that provide measurements on an in-service and non-intrusive basis. These In-Service, Non-intrusive Measurement Devices (INMDs) are utilized primarily for the measurement of voice-grade parameters such as speech level, noise level, echo loss and speech echo path delay. INMDs may also be used to measure parameters associated with digital transmission systems that impact the performance of the voice-grade channels being transported. This Recommendation specifies interface, measurement range, and accuracy requirements for measuring voice-grade transmission parameters as well as descriptions of optional functions associated with these parameters. It does not specify the measurement algorithms to be used, nor the application of the resulting measurement.

1.2 Purpose

The INMD in this Recommendation is intended for in-service (maintenance) application, to detect network anomalies affecting transmission performance on voice services.

1.3 Application

The INMD is used as a stand-alone device or can be used as part of a network element. They may be deployed at selected switch and facility nodes in telecommunications networks to measure the in-service performance parameters of voice grade services, and to locate and analyse network anomalies. For the switched network, analysis of network anomalies is made easier when the connection information such as calling and called address digits, circuit assignments involved, etc., are known, together with the measured performance. Recording of such information does not constitute intrusion of privacy, since speech intelligence is not monitored. Other optional functions may be added to make the INMD more useful.

The INMD can only be used at a four-wire point. In order to study conditions on the two-wire part of a subscriber line, the INMD must be connected via a four-wire trunk on the network element (that connects to the subscriber's line under study); thus, to isolate a problem to the particular subscriber's line, some means of conveying connection information from the network element to the INMD must be employed. For facility access, the INMD monitors the MF (Multi-Frequency) or DTMF (Dual-Tone, Multi-Frequency) signalling; however, in cases such as SS7 (Signalling System No. 7), a new standard will be required to define the interface protocol. The INMD measures transmission on the path including the customer provided equipment to the point of INMD measurement access. In this way the INMD can detect transmission anomalies on the built-up connection. These anomalies can be caused by the customer environment, subscriber line, switches and trunks, including anomalies at interfaces between these network elements. In particular, the INMD can, potentially, observe anomalies that are not detectable by traditional out-of-service tests. Examples of these difficult-to-detect anomalies are:

- intermittent fading;
- acoustical feedback;
- room noise;
- defective customer equipment;
- intermittent leakage on metallic circuits;
- intermittent noise;
- design violations;

- pair gain system problems;
- digital switch level and echo control problems;
- echo problems at the line-to-trunk interface;
- tones and announcement level control problems; and
- switch translation problems that cause violations in network loss plans.

Although the INMD is potentially able to detect such anomalies, it should be noted that the INMD cannot specifically separate combined signals, e.g. room noise and cable noise, or acoustical feedback and hybrid mismatch. Similarly, the INMD cannot distinguish between a trouble on a trunk, on a switch, on the terminating subscriber line, or on the terminal equipment. Network anomalies detected by the INMD may be subsequently isolated to the affecting network element, by applicable out-of-service tests, or facility performance monitoring. Alternatively, the use of advanced signal processing techniques, such as pattern recognition, may enable the INMD to infer the source of an anomaly, e.g. by recognizing particular types of noise.

As an alternative to routine testing of network elements, for the detection of network anomalies, the INMD is effective when used in call sampling mode. In this way, there is no need to maintain large routine testing databases. For post active detection, the INMD is most effective when used as a portable device for detecting intermittent problems. Stand-alone, facility access INMD systems, complete with proprietary interfaces to data collection and analysis software, currently exist and are being used by service providers.

2 Normative references

The following 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: all 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.

[1] ITU-T Recommendation P.10 (1993), *Vocabulary of terms on telephone transmission quality and telephone sets*.

[2] ITU-T Recommendation G.100 (1993), *Definitions used in Recommendations on general characteristics of international telephone connections and circuits*.

[3] ITU-T Recommendation P.56 (1993), *Objective measurement of active speech level*.

[4] ITU-T Recommendation G.122 (1993), *Influence of national systems on stability and talker echo in international connections*.

[5] CCITT Recommendation G.703 (1991), *Physical/electrical characteristics of hierarchical digital interfaces.*

[6] ITU-T Recommendation G.772 (1993), *Protected monitoring points provided on digital transmission systems*.

[7] CCITT Recommendation M.3010 (1992), *Principles for a telecommunications management network*.

[8] ITU-T Recommendation G.763 (1994), *Digital circuit multiplication equipment using (Recommendation G.726) ADPCM and digital speech interpolation*.

[9] CCITT Recommendation G.223 (1988), *Assumptions for the calculation of noise on hypothetical reference circuits for telephony*.

[10] CCITT Recommendation G.212 (1988), *Hypothetical reference circuits for analogue systems*.

[11] ANSI/IEEE 743-1984: *Standard methods and equipment for measuring the transmission characteristics of analogue voice frequency circuits*.

[12] CCITT Recommendation G.131 (1988), *Stability and echo*.

[13] CCITT Recommendation P.48 (1988), *Specification for an intermediate reference system*.

[14] CCITT Recommendation G.712 (1992), *Transmission performance characteristics of pulse code modulation*.

[15] ITU-T Recommendation O.41 (1994), *Psophometer for use on telephone-type circuits*.

[16] ITU-T Recommendation G.165 (1993), *Echo cancellers*.

[17] CCITT Recommendation G.711 (1988), *Pulse Code Modulation (PCM) of voice frequencies*.

[18] ITU-T Recommendation G.191 (1993), *Software tools for speech and audio coding standardization*.

[19] ITU-T Recommendation P.80 (1993), *Methods for subjective determination of transmission quality*.

[20] ITU-T, *Handbook on Telephonometry (1993)*.

[21] ITU-T Recommendation P.830 (1996), *Subjective performance assessment of telephone-band and wideband digital codecs*.

[22] ITU-T Recommendation P.310 (1996), *Transmission characteristics for telephone band (300 - 3400 Hz) digital telephones*.

3 Abbreviations and definitions

3.1 Abbreviations

Relevant abbreviations in ITU-T Recommendation P.10 [1] will apply.

ADPCM – Adaptive Differential Pulse Code Modulation

DCME – Digital Circuit Multiplexing Equipment

DSI – Digital Speech Interpolation

DTMF – Dual-Tone Multi Frequency

ENR(p) – Echo to Noise (psophometrically weighted) Ratio

INMD – In-service Non-intrusive Measurement Device

IRS – Intermediate Reference System

LPC – Linear Predictive coding

LSTR – Listener Sidetone Rating

MF – Multi-Frequency

OLR – Overall Loudness Rating

PCM – Pulse Code Modulation

REL – Reference Echo Loss

REPL – Reference Echo Path Loss

RLR – Receiving Loudness Rating

r.m.s. – root mean square

RNL – Reference Noise Level

RSAF – Reference Speech Activity Factor

RSEPD – Reference Speech Echo Path Delay

RSEPL – Reference Speech Echo Path Loss

RSL – Reference Speech Level

SLR – Sending Loudness Rating

SNR(p) – Active speech level to noise (psophometrically weighted) ratio

STMR – Sidetone Masking Rating

3.2 Definition

For the purposes of this Recommendation, the following definitions apply.

3.2.1 double talk interval: The interval during which both directions of transmission are experiencing incident speech spurts.

At the INMD monitoring point, this will be different from the double talk experienced by both parties due to the delay between the termination points and the measurement equipment.

3.2.2 echo: Echo is defined in Recommendation G.100 [2] as unwanted signal delayed and insufficiently weakened to such a degree that, for instance, in telephony, it is perceived as distinct from the wanted signal (i.e. the signal directly transmitted).

3.2.3 echo Loss: The echo loss (Recommendation G.122 [4]) is derived from the integral of the power transfer characteristic weighted by a negative slope of 3 dB/octave starting at 300 Hz and extending to 3400 Hz. The echo loss should be calculated with the speech echo path delay removed. This echo loss figure has been found to give better agreement with subjective opinion for individual connections than an unweighted echo path loss.

For a flat echo path frequency response, echo loss is equal to speech echo path loss and echo path loss.

3.2.4 echo path: The round trip electrical path starting from the point of incident speech measurement and ending at the point where the correlated reflected speech is measured.

3.2.5 echo path loss: The echo path has a unique impulse response. The Echo Path Loss is the integral of the impulse response (in the frequency domain). Echo Path Loss is not dependent on the speaker.

3.2.6 far-to-near (FN) transmission direction: A distinctly specified direction of transmission with one end of a circuit labelled far (F) and the other end labelled near (N), where transmission is in the direction from far to near.

3.2.7 hangover time: A specified duration of time beginning with the end of a speech spurt and ending at the beginning of noise measurements, if no new speech spurts occur in the mean time.

3.2.8 incident speech: Speech whose source is talker energy.

3.2.9 near-to-far (NF) transmission direction: A distinctly specified direction of transmission with one end of a circuit labelled near (N) and the other end labelled far (F), where transmission is in the direction from near to far.

3.2.10 noise level: The electrical energy (measured in dBmp) caused by spurious signals.

Spurious signals, i.e. noise, can be generated internally to the circuit or may be the result of interference from external sources. These sources can be classified as environmental, circuit and impulse noise. Noise level measurements should only be applied to steady noise during speech pause intervals.

Measurements of impulsive noise signals is an area for further study.

3.2.11 reflected speech: Speech whose direction of transmission and level has been altered by network discontinuities such as four-to-two wire conversions.

This definition is consistent with the definition of Talker echo but includes any reflected speech which may not be classified as echo.

Talker echo is defined in Recommendation G.100 [2] as echo produced by reflection near the listener’s end of a connection, and affecting the talker.

3.2.12 speech activity factor: The ratio of active speech to total measurement interval.

3.2.13 speech echo path delay: It is the period (in ms) between the detection of an incident signal at a zero reference point, on a four-wire point, to the detection of its corresponding reflected signal at the same four-wire point (on the opposite direction).

For multiple echo path reflections the speech echo path delay should be calculated for each detection of the corresponding reflected signal.

3.2.14 speech echo path loss: It is the ratio of the r.m.s. values of the incident to reflected speech signals with the speech echo path delay removed. The Speech Echo Path Loss is highly dependent on the speaker.

3.2.15 speech level: Active Speech level is defined in Recommendation P.56 [3]. In general, speech level is the electrical energy generated by the conversion of acoustical talker energy excluding any noise that is not part of the speech (such as impulses, reflected speech, and steady noise during periods of silence), but including inter syllabic pauses (brief periods of low or zero power that are not perceived as interruptions in the flow of speech).

3.2.16 speech pause interval (or quiet interval): A period of time during which speech levels are absent due to inter syllabic and conversational pauses.

Inter syllabic pauses are the gaps inherent in the articulation process. Such gaps are short approximately up to 350 ms and are not noticed as such by the listener. These pauses should be considered as part of the utterance and therefore included in a measurement of speech.

Conversational pauses are generally longer. They will be noticed by the listener, either consciously or subconsciously, and should be excluded from the a measurement of speech level since they do not contribute to the subjective loudness of the speech. When these pauses are excluded the measurement is said to be made when the talker is “active”.

3.2.17 speech present: It is a period of time where speech is considered present. Speech is not a continuous stream of sound, as it may seem when listening to it, but contains many pauses. Hence the speech present contains speech spurt intervals and speech pause intervals.

3.2.18 speech spurt (or utterance) interval: A period of time during which speech is present due to syllabic emphasis.

4 Interface requirements

The INMD has two types of external interfaces, electrical and data transport. The electrical interface is at a protected monitor digital access point where the signal is measured. The data transport interfaces are at the data port to network elements (for transport of connection information) and operations systems (for remote processing of INMD data).

Other additional interfaces may be supplied with the INMD, such as interfaces to data analysis systems, which can be proprietary or use open standards.

4.1 Electrical interfaces

The type of electrical interface is determined by the point of access, on a DS1 facility or at a digital switch. Analogue access is also possible but is not defined within this Recommendation.

4.1.1 DS1 digital interface

The DS1 signal format and facility protected monitor interface is described by Recommendations G.703 [5] (physical/electrical characteristics of hierarchical digital interfaces) and G.772 [6] (Digital Protected Monitoring Points).

4.2 TMN data transport interfaces

Recommendation M.3010 [7] specifies principles for telecommunications management networks, it may be desirable to provide interfaces that conform to it.

5 Functional requirements

The functional interface requirements depend on whether access to connection information is at a facility access point or at the switch network element when measuring on switched connections. For facility access, the facility may be the only data source, whereas for switch access, connection information is more conveniently obtained from the network element rather than the INMD. In fact, the network elements may be the only convenient source of connection information for common channel signalling trunk connections performed by signalling system seven. For voice grade INMDs there are recommended and optional functions. These functions are specified on a connection specific basis.

5.1 Required measurement functions

The required measurement functions are:

Speech and noise characterization

- active speech level;
- noise level (psophometric weighted);
- speech activity factor.

Echo characterization

- speech echo path delay (single or multiple reflection measurements);

and at least one of the following echo measurements:

- echo loss (single or multiple reflection measurements);
- echo path loss (single or multiple reflection measurements);
- speech echo path loss (single or multiple reflection measurements).

The above required functions are for voice connections only, and do not apply to connections of short duration, where there is insufficient time to make an accurate determination of the values, or where a connection is not completed. Measurement of these recommended functional values is described subsequently. These recommended functions are necessary for the in-service voice transmission application.

Three echo loss parameters are recommended as each gives different information concerning the connection. The echo loss conforms to Recommendation G.122 [4] and has been found to correlate better with subjective opinion than unweighted echo path loss measurements. The echo path loss provides an unweighted measurement of the echo power and can identify echo problems at high frequencies (>2000 Hz) which could impact voiceband data performance. The speech echo path loss reflects the actual amount of speech energy reflected and because of this, could be strongly correlated with the subjective opinion of callers. Further work is needed to determine if this is true. On some connections it may not be possible to resolve all three measurements. When comparing results, between INMDs, care should be taken that the correct echo measurement definition is used. For a thorough assessment of the echo path performance it is recommended that all three echo loss parameters are measured.

5.2 Optional functions

The optional functions (which is not an exhaustive list) are:

1) originating and terminating address digits;
2) facility or circuit identification;
3) time and duration of connection;
4) signal classification (voice/data/other);
5) customer identification (dedicated circuits only);
6) DS1 performance measurements;
7) 3 kHz flat noise level;
8) connection disposition measurements;
9) data analysis and reports;
10) saturation clipping;
11) measurement interval;
12) double talk;
13) front-end clipping;
14) one-way transmission;

15) crosstalk;

16) stability loss;

17) distortion.

Functions 1) through 6) are connection information that may be accessed via the facility or interface to network elements. The interface to network elements may be integrated to the INMD when used as a mediation device or separate operations systems interfacing with the INMD. Measurement or recording of these optional functional values is described in clause 7. These optional functions are useful for locating and analysing network anomalies.

6 Description of required measurement functions

The required measurements are performed on a circuit during the speech present state. These measurements should begin no sooner than two seconds after speech begins. This is to minimize the effects of transients caused by state changes and echo canceller convergence. The two basic parameters measured are speech and noise level. Echo loss and speech echo path delay are derived from incident speech in the near-to-far direction and its correlated reflected speech in the far-to-near direction of transmission and vice versa.

If it is determined that, during the measurement interval, signals other than speech are detected, the recommended measurements should not be reported as speech but may be reported as non-speech parameters.

Measurements with insufficient samples or integration time, or beyond the range of the INMD should be reported as such. Tables 1 and 2 give the INMD measurement specifications for range, accuracy, and measurement resolution. These measurements are uniquely defined in this Recommendation and should not be confused with out-of-service measurements defined by other Recommendations.

6.1 Speech measurements

Once speech is detected the speech level shall be measured over speech spurt intervals.

6.1.1 Speech classification

At the beginning of a measurement, the signal needs to be classified to determine whether the circuit is in a speech present state. Signals that can be present on a circuit are speech, data, fax, idle circuit, music, reorder, ring back, no circuit tones, etc.

The requirements for speech classification in INMD is similar to that of DCME. Recommendation G.763 [8] has Recommendations for data/speech classification for DCME equipment. INMD equipment should adhere to these Recommendations.

Idle can be classified as a circuit that has a stream of ± 1 for A-law (alternate D5/55) and ± 0 for μ-law (alternate FF/7F) [17].

6.1.2 Active speech level measurement

The active speech level is the root mean square (r.m.s.) of the speech spurt amplitudes expressed in dBm. The active speech level shall be the speech level averaged over speech spurt intervals including any hangover time and does not include measurements during conversational pauses. By including hangover time, any inter syllabic pauses are included in the r.m.s. value. Typically, these measurements are estimated by sampling or integrating the active speech signal. This definition is consistent with active speech level as defined in Recommendation P.56 [3].

Recommendation P.56 [3] Method B does not give guidance for the measurement of speech with high noise or reflected speech present. Measurement under these conditions is for further study.

6.1.3 Speech level measurement interval

The speech level measurement interval shall be long enough to predict an accurate active speech level within one dB of a measurement made over the entire call duration. The measurement interval should include a minimum of 20 seconds of active speech.

If less than 20 seconds of active speech is measured the result should be reported separately as having a reduced reliability.

6.1.4 Speech activity factor

Speech activity factor is the ratio of active time to total time elapsed during a measurement, usually expressed as a percentage. The active time is the aggregate of all intervals of time when speech is deemed to be present.

This parameter is recommended because of its usefulness in helping in interpreting other measurements. However because speech activity factor measurements use hangover times, caution must be used in relating this measurement to the actual activity factor of the source speech material.

6.2 Noise measurement

Once speech is classified, the noise level shall be measured over measurement intervals taken when both directions of transmission are quiet (no incident or reflected speech). This is usually during speech pause intervals with corresponding listening intervals in the other direction of transmission.

Alternatively, measurements can be made when only the direction of transmission currently being measured is quiet (no incident speech), but it must be ensured that either reflected speech (echo) has been adequately cancelled to meet the noise measurement accuracy requirements or other techniques are used to remove the echo from the noise measurement value.

6.2.1 Noise level (psophometric weighting)

The psophometric weighting is specified by Recommendation G.223 [9][1)].

The noise level shall be the average root mean square (r.m.s.) of the weighted noise amplitude expressed in dBmp (Recommendation G.212). The noise level is the psophometric weighted noise level averaged over speech pause intervals excluding any hangover time. By excluding hangover time, r.m.s. noise measurements are not corrupted by any residual speech. r.m.s. averaging for noise should comply with Recommendation O.41 [15].

6.2.2 Noise level measurement interval

The noise level measurement interval shall be long enough to predict an active noise level within one dB of a measurement made over the entire call duration. It is currently assumed that typically, this measurement interval is one minute or longer, although this needs to be verified.

6.2.3 Non-stationary noise classification and measurement

Techniques for non-stationary noise classification and measurement are currently under study. Appendix II (Degradation of Telephone Transmission Quality due to non-stationary noise) provides an overview of the current work in this area.

6.3 Echo measurement

Once speech is detected, echo loss and delay are measured over measurement intervals taken during incident speech spurts in the transmit direction and reflected speech spurts during the corresponding listening intervals in the receive direction of transmission. Measurements shall not be made during speech pause intervals.

It should be noted that the echo loss and delay measurements are uniquely defined here as in-service measurements, and should not be confused with out-of-service measurements defined by other Recommendations.

The definitions given below are in terms of the impulse response or frequency response of the echo path. However the intention of these definitions is to provide a reference for the measurements, not to exclude alternative methods of calculating the echo loss or speech echo path delay.

6.3.1 Speech echo path delay measurement

The INMD should be able to detect single and/or multiple echo paths.

---

1) Recommendation G.223 [9] defines the psophometric weightings. Recommendation G.212 [10] defines psophometric noise level.

6.3.1.1 Single reflections

The speech echo path delay shall be the value from the zero reference at the point of measurement to the peak absolute amplitude of the time domain impulse response of the echo path.

6.3.1.2 Multiple reflections

The speech echo path delay, of any one reflection, shall be the value from the zero reference at the point of measurement to the peak absolute amplitude of the time domain impulse response of the echo path.

The INMD should be able to differentiate between reflections that are separated by 10 ms or more.

6.3.1.3 Methods of calculating speech echo path delay

For INMD, there are currently two known measurement techniques for speech echo path delay – correlation and adaptive filter analysis. It is likely that each technique will be appropriate to a different range of values. A brief introduction to the two techniques is given in Appendix I.

6.3.2 Echo Loss measurement

The INMD should be able to detect single and/or multiple echo paths.

Echo Loss (a-b) as defined in Recommendation G.122 [4] can be calculated from the frequency response of the echo path using a frequency weighting. This Echo Loss figure has been found to give better agreement with subjective opinion for individual connections than an unweighted echo path loss. However, for large samples of actual connections it has been found that the two methods give very similar means and standard deviations.

6.3.2.1 Single reflections

The Echo Loss shall be the integral of the weighted frequency response of the echo path.

6.3.2.2 Multiple reflections

The Echo Loss of any one reflection shall be the integral of the weighted frequency response for that reflection.

6.3.3 Echo path loss measurement

Echo path loss is derived from the impulse response of the echo path.

6.3.3.1 Single reflections

The echo path loss shall be the integral of the total impulse response (in the frequency domain) of the echo path. Using Parseval’s theorem this is equivalent to mean sum of squares of the impulse response (in the time domain).

6.3.3.2 Multiple reflections

The echo path loss of any one reflection shall be the integral of the unweighted frequency response for that reflection.

6.3.4 Speech echo path loss measurement

The speech echo path loss can only be calculated if the speech echo path delay is known. The speech echo path delay is used to determine when the r.m.s. of the reflected speech signal is calculated.

The speech echo path loss shall be the r.m.s. ratio of incident to reflected speech signals in dB.

7 Description of optional functions

The optional functions are performed on a circuit during the active state or (for facility access devices) during the signalling and supervision state. These functions should be performed once the circuit is seized for connection (off hook indication). Any or all of these parameters can be measured and reported per user options. Following are typical approaches and applications for these functions. Detailed function specifications may vary and are not part of this Recommendation.

7.1 Originating and terminating address digits

For facility access INMD the originating and terminating address are decoded from signalling information associated with the channel being observed. For switch access INMD, this information is obtained via data interface to the switch processor. This information is useful for identifying performance with a particular end user, access line, local switching office, or network route within the switched networks.

7.2 Facility or circuit identification

For facility access INMD the facility or circuit identification is usually the system group or digroup and channel codes (for multiplex facilities) or pair or channel code (for single channel facilities). The INMD can be primed with this information at the time of bridged access. For switch access INMD, the circuit termination information is obtained via data interface to the switch processor. This information is useful for identifying performance with a particular network element or circuit.

7.3 Time and duration of connection

The time of connection is usually the time at which the measured connection is placed in the active state. The duration of the connection is usually the interval between the connect and disconnect time. For facility access INMD, this information is usually decoded from the supervision information on the channel being observed. For switch access INMD, this information is obtained via data interface to the switch processor. This information is useful as traffic data. Additionally, correlation on connections with short duration and poor performance can indicate severe network failures.

7.4 Signal classification

The service classification (data, voice, etc.) can be determined. Also, more specific information such as data baud rate, analogue versus digital, signal format, etc. might be detectable by signal processing and pattern recognition techniques. This information is useful as traffic data. Furthermore, this classification will be necessary for measurement of connection disposition and data parameters to be covered by future INMD standards.

7.5 Customer identification

For dedicated circuits, the end user or inter-exchange customer identification is usually primed into the INMD at the time of bridged access. This information may be useful for reporting customer performance as well as sharing performance data with the customer.

7.6 DS1 performance measurements

For DS1 access, performance measurements such as framing bit error rate, loss of synchronization, frame loss and slips, jitter, etc., can be measured on the DS1 signal. These measurements and their performance criterion are described by other ITU-T Recommendations. These measurements are useful for reporting on digital facility performance. Additionally, they are useful for correlating the INMD performance measurements with the observed DS1 performance. In this way network problems and their causes can be more clearly identified.

7.7 3 kHz flat noise level

The 3 kHz flat noise level is measured in a way similar to the psophometric noise level in 6.2.1: except there, the psophometric weighting is changed to flat weighting. The noise level is weighted according to the ANSI/IEEE 743 [11] standard for measuring analogue voice frequency circuits. When digital signal processing techniques are used, any DC component shall be filtered out. The 3 kHz noise measurements are useful for correlating the psophometric noise level to determine the type and spectral content of the noise present on the circuit.

7.8 Connection disposition measurements

These measurements may be useful for determining circuit connection performance and points of abnormally high connection failures. Connection disposition could include, for example, “answered”, “busy”, and “ring no answer”.

7.9 Data analysis and reports

Data analysis and reports are useful for network characterization as well as network management and maintenance reports.

7.10 Saturation clipping

The percentage of active speech which has been saturated (amplitude “clipped”, causing non-linear distortion) can be detected by looking for occurrence of the maximum positive and negative PCM codes.

7.11 Measurement interval

Recording the interval over which the measurements have been taken allows monitoring of the consistency of the INMD measurements with different measurement intervals. The optimum measurement intervals can then be determined, although this would usually need to be done under controlled test conditions rather than on an “unknown” network.

7.12 Double talk

Double talk is a condition whereby, for whatever reason, one conversant in a telephone connection starts talking before the other has finished. This is a naturally occurring phenomenon – where one participant wishes to interrupt the flow of speech from the other but occurs more frequently the greater the propagation time of the connection becomes. Monitoring the occurrence of double talk provides a useful indicator linked to customer perception of performance.

Double talk may be reported as a percentage of the measurement interval. Other reporting metrics may be appropriate, these are for further study.

7.13 Front-end clipping

Front-end clipping occurs when the start of a speech burst is missed due to a failure in, for example, the speech detection and interpolation algorithms in digital circuit multiplication equipment or when the number of active signals temporarily exceeds the number of available channels in busy periods (freeze out).

7.14 One-way transmission

Temporary loss of one direction of the transmission has a severe effect on customer perception of quality. By an analysis of activity and noise levels, it may be possible to detect even short durations of one-way transmission. The ability to detect permanent one-way transmission will depend on whether the faulty circuit is monitored when the connection is made. This will depend upon the rate of scanning, since neither party will hold the circuit for very long. One-way transmission can occur in some genuinely connected calls (e.g. when a user is listening to a recording) and the probabilities of this situation will be taken into account in deciding whether to classify apparent one-way transmission as a fault or not.

7.15 Crosstalk

Crosstalk on a connection is said to occur when a talker participating in another conversation over a different circuit is heard on the connection. This occurs when there is excessive mutual coupling between circuits. “Near end” crosstalk is where the disturbing talker is located at the same end of the group of circuits as the disturbed listener. “Far end” crosstalk occurs when the disturbing talker and disturbed listener are at opposite ends of the group of circuits. It may be possible to detect crosstalk by an analysis of speech levels and speech statistics but it must be realized that there are distinct problems in differentiating far-end crosstalk from background talkers or from multiple talkers in three-way links.

7.16 Stability loss

The risk of the echo loss reaching low values at any frequency in the range 0-4 kHz should be as small as practicable. The stability loss is the lowest value of loss in the frequency band to be considered (Recommendation G.100 [2]).

7.17 Distortion

Classification and measurement of codec distortion is for further study.

8 Requirements for performance of measurements

The performance requirements which follow are the minimum required for verification, for the required measurement functions in clause 5. The requirements are split into four classes[2)] of INMD for operation on national, international and non-linear and/or time-invariant networks.

Class A: Limited to measure short delay routes containing analogue and 64 kbit/s PCM [17] components only (i.e. no low bit rate codecs) and no echo control devices.

Class B: Limited to measure moderate delay networks that include echo control devices.

Class C: For use in long delay networks that may include signal processing devices such as echo control and speech compression (e.g. DCME and ADPCM) which incorporate waveform encoding but do not include voice encoders (e.g. LPC).

Class D: For use on arbitrary, possibly non-linear and time-variant, networks which include processing devices such as LPC coders.

Tables 1 and 2 give the operating ranges and accuracy for each measurement.

NOTE – Values of speech level and noise level are quoted throughout this Recommendation in terms of absolute levels using dBm and dBmp units. These levels apply for measurements at points having 0 dBr relative levels, i.e. the levels are the same when quoted in terms of dBm0 and dBm0p, respectively. INMD devices can report values in these units if desired.

8.1 Description of different classes

8.1.1 Class A – Local (National for many countries) networks

A Class A device is specified for use on short delay routes with up to a roundtrip delay of 50 ms. The specification for measurement ranges is given in Table 1. The connections are expected to be formed by analogue and 64 kbit/s connections only (i.e. no low bit rate codecs) and no echo control devices.

Additional requirements for Class A device

- If a device can measure accurately beyond the specified ranges the result can be reported. A default code should be used when it is outside the prescribed accuracy.

Alternatively,

- If the monitored speech echo path delay is greater than 50 ms, the INMD should report a default code indicating the speech echo path delay is greater than 50 ms.
- If the monitored echo loss/echo path loss/speech echo loss is greater than 25 dB, the INMD should report a default code indicating echo loss/echo path loss/speech echo loss is greater than 25 dB.

[2)] INMDs designed for operation on international networks and networks that are non-linear and/or time invariant are likely to be much more complicated than those not so designed. It is felt that these devices need to meet separate more stringent requirements whilst not impeding the development of devices for national networks.

- If the monitored speech level is less than –35 dBm, the INMD should report a default code indicating speech level is less than –35 dBm.

- If the monitored noise level is greater than –40 dBmp, the INMD should report a default code indicating noise level is greater than –40 dBmp.

8.1.2 Class B – Medium delay networks

Class B INMDs are specified for use on medium delay routes, that have a roundtrip delay up to 150 ms. The specification for the measurement ranges is given in Table 1. It is expected that such routes will include echo control devices such as echo cancellers.

Additional requirements for Class B device

- If a device can measure accurately beyond the specified ranges the result can be reported. A default code should be used when it is outside the prescribed accuracy.

Alternatively,

- If the monitored speech echo path delay is greater than 150 ms, the INMD should report a default code indicating the speech echo path delay is greater than 150 ms.

- If the monitored echo loss/echo path loss/speech echo loss is greater than 35 dB, the INMD should report a default code indicating echo loss/echo path loss/speech echo loss is greater than 35 dB.

- If the monitored speech level is less than –35 dBm, the INMD should report a default code indicating speech level is less than –35 dBm.

- If the monitored noise level is greater than –40 dBmp, the INMD should report a default code indicating noise level is greater than –40 dBmp.

Specification for connections that contain echo cancellers

- If the echo canceller is working correctly, the INMD should report a default code indicating echo loss/echo path loss/speech echo loss is greater than 35 dB.

- If the echo canceller fails to fully remove the echo, the system should:

  i) measure the echo loss/echo path loss/speech echo loss and delay; or

  ii) indicate that echo is present.

Item i) is the most desirable; however, the returned signal may be sufficiently distorted such that an accurate measurement of echo and delay is no longer possible. In these cases item ii) is acceptable.

8.1.3 Class C – Long delay networks

Class C INMDs are specified for use on longer delay routes, that have a roundtrip delay up to 1000 ms. The specification for the measurement ranges is given in Table 1. It is expected that such routes will include signal processing devices such as echo cancellers and DCMEs. These devices may under some circumstances add additional degradation to the circuit's performance.

NOTE – Users should keep in mind that even if a Class C device can measure echo path delays up to 1000 ms (or more), according to Recommendation G.114, a round trip transmission delay exceeding 800 ms is not acceptable in international calls.

Additional requirements for Class C device

- If a device can measure accurately beyond the specified ranges the result can be reported. A default code should be used when it is outside the prescribed accuracy.

Alternatively,

- If the monitored speech echo path delay is greater than 1000 ms, the INMD should report a default code indicating the speech echo path delay is greater than 1000 ms.
- If the monitored echo loss/echo path loss/speech echo loss is greater than 45 dB, the INMD should report a default code indicating echo loss/echo path loss/speech echo loss is greater than 45 dB.
- If the monitored speech level is less than –35 dBm, the INMD should report a default code indicating speech level is less than –35 dBm.
- If the monitored noise level is greater than –40 dBmp, the INMD should report a default code indicating noise level is greater than –40 dBmp.

Specification for connections that contain echo cancellers

- If the echo canceller is working correctly, the INMD should report a default code indicating echo loss/echo path loss/speech echo loss is greater than 45 dB.
- If the echo canceller fails to fully remove the echo, the system should:

  i) measure the echo loss/echo path loss/speech echo loss and delay; or

  ii) indicate that echo is present.

Item i) is the most desirable; however, the returned signal may be sufficiently distorted such that an accurate measurement of echo and delay is no longer possible. In these cases item ii) is acceptable.

Specification for connections that contain DCME

- The INMD should report the measured echo loss/echo path loss/speech echo loss and speech echo path delay. The echo signal may be degraded by clipping or low-bit rate coding and hence any measurement will be less accurate.
- If the DCME is operating in conjunction with an echo canceller, the INMD should report measurements as specified for echo cancellers.
- The measured noise level will be affected by the comfort noise inserted by the DCME. The INMD should report the noise level detected.

8.1.4 Class D – Future networks

Class D devices are for use on circuits that use non-linear processing or vocoders.

This is an area for further study.

8.2 Measurement specification

See Tables 1 and 2.

8.3 Requirements for accuracy

For any single measurement the difference between the reference measurement (used to calibrate the test circuit) and the INMD measurement shall not exceed ± 2. This is defined by equation (8.1).

$$Accuracy = INMD\_measurement - REFERENCE\_measurement \qquad (8.1)$$

where the accuracy is measured in dB for active speech level, noise, echo loss, and speech echo path loss and milliseconds for echo path delay.

TABLE 1/P.561

INMD measurement specification for Class A, B and C for linear circuits

Measurement		Range		Mean accuracy (Note 2)	Resolution (Note 4)
		Lower	Upper		
Active speech level (dBm) (Note 1)		–35 –14.9	–15 0	± 0.3 ± 0.3	0.1 0.1
Speech activity factor (%) (Note 5)		0	100		0.1
Noise level (dBmp)		–70	–40	± 0.3	0.1
Echo Loss (dB) Echo path loss (dB) Speech Echo path loss (dB) (Note 3)	A[a)]	6	25	± 0.3	0.1
	B[a)]	6	35	± 0.3	0.1
	C[a)]	6	45	± 0.3	0.1
Speech Echo path delay (ms) (Note 3)	A[a)]	0	50	± 0.3	0.1
	B[a)]	0	150	± 0.3	0.1
	C[a)]	0	1000	± 0.3	0.1

[a)] Refer to different class of device.

NOTES

1 Active speech levels greater than –15 dBm may cause saturation on digital connections, i.e. A-law or μ-law. The requirement for ± 0.3 accuracy may need revision if the reference measurement is not limited by a digital codec.

For measurements on an analogue connection, the analogue to digital converter of the INMD may cause saturation.

2 For any single measurement, the difference between the reference measurement shall not exceed ± 2 dB (or ± 2 ms).

3 The upper limits of the echo loss and echo path delay measurement ranges for classes A, B and C devices are consistent with each other, for in each case, from the point of view of the near end user, the talker echo loudness rating given by:

TELR = (EPL upper range limit + near side SLR + near side RLR).

TELR = (EPL upper range limit + 10)
corresponds to the smallest acceptable loudness ratings according to Figure 2/G.131 [12] if the one-way transmission delay does not exceed the overall mean one-way transmission time given by:

OMOTT = (EPD upper range limit + near side mean one-way transmission time)
where the near side mean one way propagation time is great enough for the INMDs of the corresponding class to take into account any end user connection to the national network covered by the specification of this class.

4 A 0.1 resolution is required to meet the ±0.3 mean accuracy. For reporting measurements a greater resolution of 1 or 0.5 is acceptable.

5 Accuracy and resolution is for further study.

In addition, for each measurement by an INMD, the average of the accuracy's for each reported measurement over all of the test circuits and speech segments will be calculated:

$$MeanAccuracy = \sum Accuracy[i, j, k] / N \tag{8.2}$$

where i indicates the number of the test circuit, j is the number of the speech sample being tested, k indicates which transmission direction is being evaluated, and N is the total number of measurements reported by the INMD. Test cases where the INMD did not report a measurement for a parameter will be ignored in the calculation of Mean Accuracy. The Mean Accuracy shall not exceed ± 0.3.

TABLE 2/P.561

INMD specification for circuits that contain signal processing equipment

Signal processing equipment	Echo loss and speech echo path delay		
	A	B	C
Echo Canceller	N	Y	Y
DCME	N	N	Y

NOTES

1 N) INMD is not specified to work with this equipment in circuit.

2 Y) INMD is specified to work with this equipment in circuit.

For each test circuit, at least 30% of the measurements shall report a result for each parameter. The average yield over all circuits shall be better than 60%.

8.4 Description of reference circuits

Three reference circuits are defined for testing of the INMD classes. The circuits are designed to simulate a local analogue network connected to a digital transmission network.

8.4.1 Reference circuit 1

The reference circuit in Figure 1 is designed such that INMD with a digital interface is connected at points A and B. The codecs are either μ-law or A-law depending whether the equipment is specified to work on T1 or E1 circuits.

8.4.1.1 Near and far speech

The near and far speech are two-way conversations digitally recorded in a quiet background environment. The method of producing suitable speech material is described in Annex A.

8.4.1.2 Speech band limiting filters (FR1 and FR2)

Filters FR1 and FR2 should be PCM filters [14] to band limit the speech.

8.4.1.3 Attenuators (A1 and A2)

Attenuators A1 and A2 are used to set the reference speech levels.

8.4.1.4 Near and far noise (N1 and N2)

For a digital reference circuit the near and far noise sources are band limited by the codec. N1 and N2 should be uncorrelated.

For an analogue reference circuit, the near and far noise sources are 3 kHz flat band limited random white noise. The band limiting filter is specified in Recommendation G.712 [14].

8.4.1.5 Echo path channel

The echo path channel is simulated in the test circuit by discrete circuit elements – a variable delay unit (D1 or D2); an attenuator (A3 or A4) to simulate the loss; and a filter (FR3 or FR4) to simulate the frequency response. A summing amp is used to add the echo path to the transmit path.

The characteristics of filters FR3 and FR4 for use in the echo path channel are:

1) flat frequency response in the range 300 Hz to 3.4 kHz with ± 3 dB ripple.

2) positive slope of 10 dB per decade with 0 dB at 3.2 kHz with ± 3 dB ripple.

These filters can be implemented using a digital filter or standard test and measurement equipment.

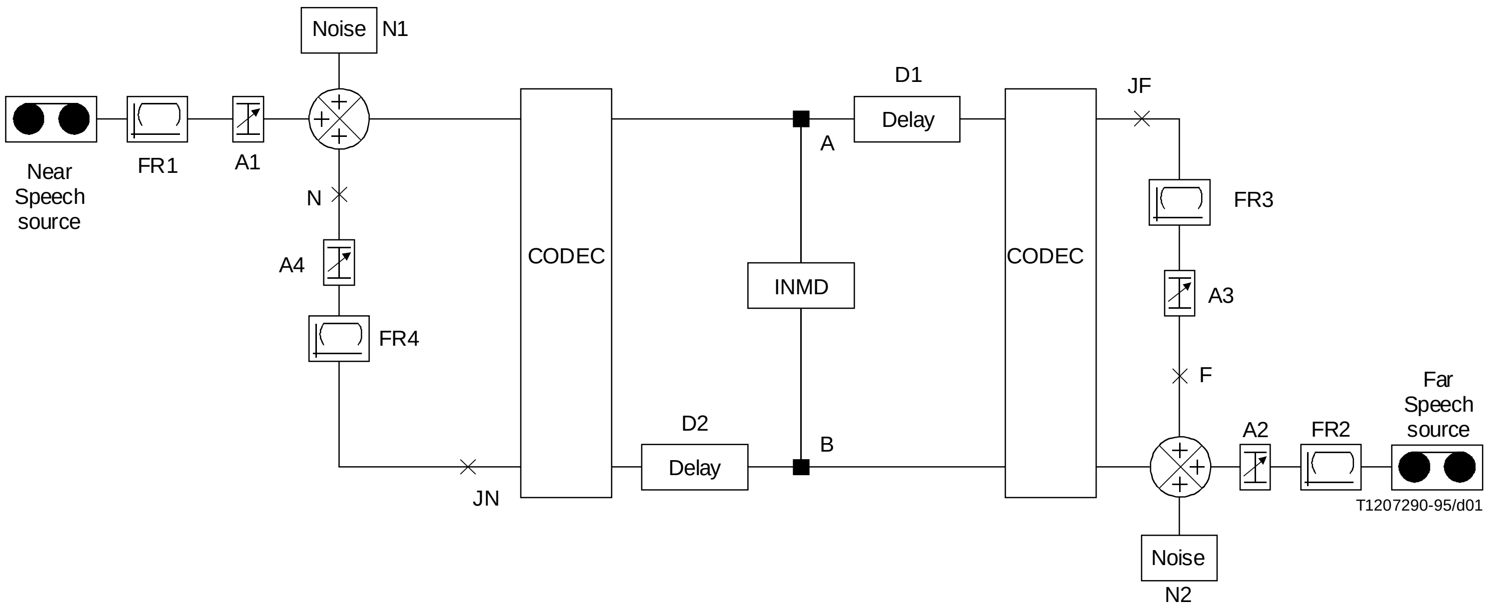

FIGURE 1/P.561

Reference circuit 1 – for linear test conditions

8.4.2 Reference circuit 2

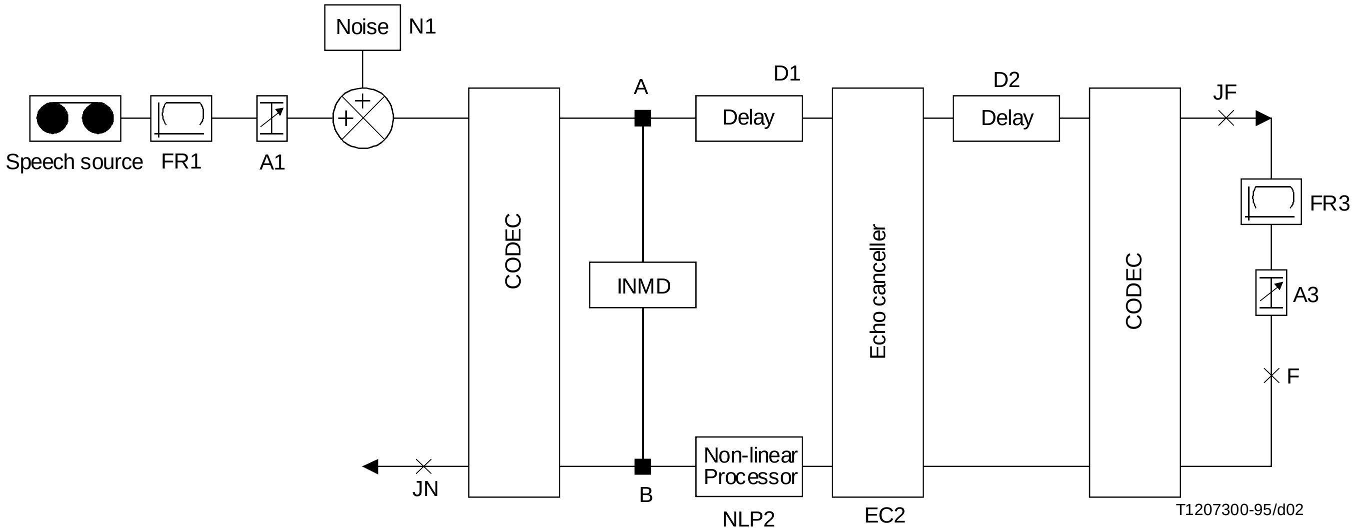

FIGURE 2/P.561

Reference circuit 2 – for test conditions with echo canceller

Figure 2 shows a possible reference circuit for testing Class B and C devices. Items FR1, A1, N1, FR3, A3 are as described for Figure 1. The delay is split into two values D1 and D2. D1 represents a transmission delay, whereas D2 represents the local delay to be cancelled.

The echo canceller is split into two parts EC2 which is the linear part of the canceller and NLP2 which is the non-linear part of the canceller.

8.4.3 Reference circuit 3

A reference circuit for testing Class C devices with DCME is similar to the circuit shown in Figure 2 with the exception that a pair of back-to-back DCME devices are added to the right of points A and B.

8.5 Reference measurements

The resolution for all reference measurements shall be within a tenth of a dB. If an analogue termination is used, then when jumpers JN and/or JF in Figure 1 are disconnected, both sides of the disconnected jumper(s) shall be terminated in 600 ohms. All analogue measurements are represented as bridged measurements, with bridging loss no greater than 0.1 dB.

For each test case the following tolerances are acceptable.

Active Speech Level:	Each reference speech measurement should be within 1 dB of level given in clause 9.
Noise Level:	Each reference measurement should be within 0.5 dB of the level given in clause 9.
Echo Loss:	Each reference measurement should be within 0.5 dB of the level given in clause 9.
Echo path loss:	Each reference measurement should be within 0.5 dB of the level given in clause 9.
Speech echo path loss:	Each reference measurement should be within 2 dB of the level given in clause 9.
Echo Path Delay:	Each reference measurement should be within 0.5 ms of the level given in clause 9.

8.5.1 Analogue versus digital measurements

An INMD connected via a digital point will have its range governed by the codec, unlike analogue measurement equipment. This may result in differences in results for high speech levels (greater than –15 dBm) as the INMD may measure a PCM (A-law or μ-law) saturation clipped signal at the analogue to digital conversion point whereas analogue equipment will measure an unclipped (non-distorted) signal.

Therefore if the INMD interface is at a digital point, analogue reference measurements should reflect the impact of codecs. If a fully digital approach is used in testing the conversion to A-law or μ-law will take into account the clipping phenomena. Digital characterization of the resulting signals will be the reference.

8.5.2 Reference Speech Level (RSL)

The reference speech level measurements shall be made with both noise sources and echo paths (JN and JF) disconnected. The measurement algorithms shall be those specified by Recommendation P.56 [3] method B. It is recommended that the software implementation of Recommendation P.56 in Recommendation G.191 [18] be used as a reference.

8.5.3 Reference Noise Level (RNL)

The reference noise level measurements are made with speech sources disconnected and echo paths (JN and JF) connected. The measurement algorithms shall be those specified by Recommendation O.41 [15].

8.5.4 Reference Speech Activity Factor (RSAF)

The reference speech level measurements shall be made with both noise sources and echo paths (JN and JF) disconnected. A suitable reference for speech activity factor is for further study. However, a software implementation of Recommendation P.56 exists in Recommendation G.191 and can be used as guidance for the INMD measurements.

8.5.5 Reference Echo Loss (REL)

The REL is measured using a weighted noise signal. The frequency weighting is specified in Recommendation G.122 [4]. Referring to Figure 1, the Near-to-Far REL is measured by transmitting weighted noise from noise generator N1 at a level equal to the speech level that will be used on the test circuit. The speech sources and noise generator N2 are either disconnected from the circuit and replaced with 600 ohm terminations or set to transmit no signal. The Far-to-Near echo path is disconnected at JN and terminated in 600 ohms. The REL is determined by making unweighted noise measurements using digital measurement equipment at points A and B. Subtracting the measurement at point B from that at point A gives the REL. The noise measurements can also be made using analogue measurement equipment at points JF and JN.

The Far-to-Near REL is measured by reversing this arrangement.

8.5.6 Reference Echo Path Loss (REPL)

The REPL is measured using an unweighted noise signal. Referring to Figure 1, the Near-to-Far REPL is measured by transmitting noise from noise generator N1 at a level equal to the speech level that will be used on the test circuit. The speech sources and noise generator N2 are either disconnected from the circuit and replaced with 600 ohm terminations or set to transmit no signal. The Far-to-Near echo path is disconnected at JN and terminated in 600 ohms. The REPL is determined by making unweighted noise measurements using digital measurement equipment at points A and B. Subtracting the measurement at point B from that at point A gives the REPL. The noise measurements can also be made using analogue measurement equipment at points JF and JN. The Far-to-Near REPL is measured by reversing this arrangement.

8.5.7 Reference Speech Echo Path Loss (RSEPL)

The RSEPL is measured for each speech segment used in the testing. When the Near-to-Far RSEPL is measured, the Far speech source and both noise generators shown in Figure 1 are either disconnected from the circuit and replaced with 600 ohm terminations or set to transmit no signal. The Far-to-Near echo path is also terminated in 600 ohms at JN and both delay generators should be set to provide no delay. Each of the Near speech segments is then played, while speech level measurements are made using an algorithm meeting the requirements of Recommendation P.56 [3] Method B at points JF and JN or at points A and B, if a digital implementation of the speech level algorithm is available. The RSEPL is determined by subtracting the speech level measurement made at either point JN or B from that made at point JF or A. In some of the test conditions, it may be necessary to amplify the echo signal because of the limitations of the speech level measurement algorithm. The amplifier used should provide linear gain between 0 and 4000 Hz.

The Far-to-Near RSEPL is measured by reversing this arrangement.

8.5.8 Reference Speech Echo Path Delay (RSEPD)

Delay measurement methods are currently under study within various standards bodies. In the interim, *de facto* delay calibration methods can be used.

9 Test conditions for Class A, B and C devices

Test conditions for Class A, B and C devices are discussed here. Tests for Class D devices are for further study.

9.1 Test design

Test should be designed to:

1) test all parameters at their range limits;
2) test the worst case combination of parameters not to be less than a speech-to-noise ratio of 20 dB or a minimum echo-to-noise ratio of 10 dB including quantization noise; and
3) minimize the number of tests to accomplish 1) and 2).

9.2 Multiple measurements

The INMD shall be tested using speech material produced as specified in Annex A.

Each test condition shall be repeated at least three times for each conversation (in order to have sufficient total number of measurements for the accuracy requirements, see 8.3). This also performs the function of checking the reliability of the test circuits and the algorithms.

9.3 Circuit conditions for Class A device

The following circuit conditions should be met for a Class A device. The test circuits should be symmetric (i.e. the same conditions for near and far ends) for conditions in Table 3 and asymmetric for conditions in Table 4. Conditions described in Tables 3 and 4 should be used in conjunction with reference circuit 1.

TABLE 3/P.561

Symmetric test circuit conditions for Class A device

Circuit	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	SNR(p)	ENR(p)	Speech echo path delay (ms)	Speech echo path frequency response (Note 1)
1	–32	–52	10	20	10	6	Flat
2	–35	–55	6	20	14	50	Flat
3	–10 (Note 2)	–45	25	35	10	12	Flat
4	–27	–50	13	23	10	44	Slope
5	–35	–70	16	35	19	19	Slope
6	–5 (Note 2)	–40	6	35	29	38	Flat
7	–15	–50	25	35	10	25	Slope
8	–0 (Note 2)	–50	20	50	30	31	Slope

NOTES

1 Speech echo path frequency responses are defined in 8.4.1.5.

2 Due to peak clipping the speech level will differ from Method B of Recommendation P.56 [3] which assumes that the ratio of peak power to mean power in speech is 18 dB, this is not the case for these conditions.

9.4 Circuit conditions for Class B device

Echo cancellers will “leak” echo if the echo path is non-linear or the near delay is too great. In addition if the non-linear processor of an echo canceller is switched off, the canceller will still “leak” some low level residual echo. If the modelling part of the canceller is switched off this will leave a non-linear echo path with the NLP switched on.

The required specification for a Class B INMD is given in Tables 1 and 2.

TABLE 4/P.561

Asymmetric test circuit conditions for Class A device

Circuit	Near end					Far end				
	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	Speech echo path delay (ms)	Speech echo path frequency response	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	Speech echo path delay (ms)	Speech echo path frequency response
1	–15	–55	15	30	Flat	–25	–60	20	50	Flat
NOTE – Near end speech level, noise level, loss of the echo path and speech echo path delay are adjusted by A1, N1, A3 and D1 respectively (see Figure 1). Far end speech level, noise level, loss of the echo path and speech echo path delay are adjusted by A2, N2, A4 and D2 respectively (see Figure 1).										

9.4.1 Description of test elements

A Class B device should meet all the requirements of a Class A device. Hence the same reference measurements are used in conjunction with the ranges specified in Table 3. Additional tests for the Class B ranges are specified.

Further additional tests are specified for circuits that contain echo cancellers. These tests are for characterization purposes only.

9.4.2 Circuit conditions

The following circuit conditions, Tables 5 and 6, are applicable to Figure 1. The test circuits should be symmetric, i.e. the same conditions for near and far ends for conditions in Table 5 and asymmetric for conditions in Table 6.

TABLE 5/P.561

Symmetric test circuit conditions for Class B device

Circuit	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	SNR(p)	ENR(p)	Speech echo path delay (ms)	Speech echo path frequency response (Note 1)
1	–32	–52	10	20	10	150	Flat
2	–35	–55	6	20	14	20	Flat
3	–10 (Note 2)	–60	35	50	15	38	Flat
4	–20	–50	15	30	15	130	Slope
5	–35	–70	20	35	15	56	Slope
6	–5 (Note 2)	–40	6	35	29	112	Flat
7	–15	–55	30	40	10	75	Slope
8	–0 (Note 2)	–50	25	50	25	94	Slope

NOTES

1 Speech echo path frequency responses are defined in 8.4.1.5.

2 Due to peak clipping the speech level will differ from Method B of Recommendation P.56 [3] which assumes that the ratio of peak power to mean power in speech is 18 dB, this is not the case for these conditions.

TABLE 6/P.561

Asymmetric test circuit conditions for Class B device

Circuit	Near end					Far end				
	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	Speech echo path delay (ms)	Speech echo path frequency response	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	Speech echo path delay (ms)	Speech echo path frequency response
1	–15	–70	30	30	Flat	–25	–60	25	150	Flat
NOTE – Near end speech level, noise level, loss of the echo path and speech echo path delay are adjusted by A1, N1, A3 and D1 respectively (see Figure 1). Far end speech level, noise level, loss of the echo path and speech echo path delay are adjusted by A2, N2, A4 and D2 respectively (see Figure 1).										

9.4.3 Circuit conditions with echo cancellers

No pass/fail criteria are specified for the following tests. Their purpose is to characterize the INMDs performance under real network conditions.

Echo cancellers are voice operated devices placed in a 4-wire portion of a circuit. They enhance the echo loss of the circuit by subtracting an estimation of the echo from the reflected signal (Recommendation G.165 [16]). Echo canceller performance will degrade depending on network conditions, the aim of these tests is to create a circuit where the echo canceller will not fully reduce the echo and may add some additional distortion.

Echo cancellers are constructed from two parts, a cancellation process and a non-linear processor. The cancellation process is not perfect and some residual echo level will be “leaked”. A non-linear processor is used to suppress the “leaked” signal by attenuating all signal detected below a defined suppression threshold level.

Three tests are defined here:

i) echo canceller in circuit;

ii) echo canceller without non-linear processor;

iii) echo canceller without cancellation process.

9.4.3.1 Characterization test i)

If an echo canceller is placed in the circuit within its operating range no echo path should be present in the circuit. The aim of this test is to ensure that the INMD does not give false echo detection. To ensure that no double-talk, etc., occurs in the test it is only necessary to carry out this test with a single speech source.

Using the conditions in Table 7, the INMD should report no echo path delay or echo signal loss. The speech echo path delay is equal to the bulk delay + the local echo path delay. The echo canceller should be placed in the local echo path.

The test circuit is given in Figure 2.

9.4.3.2 Characterization test ii)

Test ii) is to simulate a faulty echo canceller with the non-linear processor disabled. After the cancellation some residual echo will be present in the circuit. The amount of residual echo present in the circuit will differ for different echo cancellers. Hence before this test is carried out the echo cancellers performance needs to be characterized. If the residual echo loss is greater than the range of the INMD then a linear amplifier should be used on the output of the canceller to increase the echo.

The tests in test i) should be repeated using the conditions in Table 7.

TABLE 7/P.561

Test circuit conditions for Class B device with echo canceller

Circuit	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	Bulk Delay (ms)	Local speech echo path delay (ms)	Speech echo path frequency response (Note 1)
1	–28	–52	14	59	16	Flat
2	–35	–55	6	7	8	Flat
3	–10 (Note 2)	–60	35	122	28	Slope
4	–18	–50	22	12	25	Slope
5	–21	–70	30	92	20	Flat

NOTES

1 Speech echo path frequency responses are defined in 8.4.1.5.

2 Due to peak clipping the speech level will differ from Method B of Recommendation P.56 [3] which assumes that the ratio of peak power to mean power in speech is 18 dB, this is not the case for these conditions.

9.4.3.3 Characterization test iii)

Test iii) is to simulate a faulty echo canceller with no canceller. With just the non-linear processor present in the circuit the reflected speech (echo) signal will become non-continuous, i.e. non-linear. The suppression algorithm for the non-linear device will be different for different devices – hence values need to be chosen for echo loss that will insure that some echo will be leaked. Suggested values for reflected speech level are given in Table 8.

TABLE 8/P.561

Reflected speech levels for use with non-linear processors

Reflected speech level
Threshold level +10 dB
Threshold level +5 dB
Threshold level
Threshold level –5 dB
Threshold level –10 dB
NOTE – The suppression threshold level will depend on the echo canceller that is being tested.

In general speech has peaks approximately +18 dB above the average speech level; hence even for an r.m.s. level 10 dB below the threshold some of the signal will not be suppressed. The threshold level will be dependent on the echo canceller used.

9.5 Circuit conditions for Class C device

Class C INMDs are specified for use on longer delay routes. It is expected that such routes will include signal processing devices such as echo cancellers and DCMEs. These devices may under some circumstances add additional degradation to the circuit performance.

For the effect of echo cancellers, see 9.4.

DCMEs have two main effects on the echo path, low-bit rate coding will degrade the reflected speech signal; and a voice activity switch will replace silence periods with comfort noise. These silence periods are likely to include the echo signal.

It is not feasible to specify exhaustive tests for such devices.

The required specification for a Class C INMD is given in Tables 1 and 2. Tables 11 and 12 give the additional requirements for echo and delay measurements with circuits that have signal processing devices.

9.5.1 Description of test elements

A Class C device should meet all the requirements of a Class A and B device. Hence the same reference measurements are used in conjunction with the ranges specified in Tables 3 and 5. Additional tests for the Class C ranges are specified.

Further additional tests are specified for circuits that contain DCME. These tests are for characterization purposes only.

9.5.2 Circuit conditions

The following circuit conditions, Tables 9 and 10, are applicable to Figure 1. The test circuits should be symmetric, i.e. the same conditions for near and far ends for conditions in Table 9 and asymmetric for conditions in Table 10.

TABLE 9/P.561

Symmetric test circuit conditions for Class C device

Circuit	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	SNR(p)	ENR(p)	Speech echo path delay (ms)	Speech echo path frequency response (Note 1)
1	–32	–52	10	20	10	125	Flat
2	–35	–55	6	20	14	1000	Slope
3	–5 (Note 2)	–65	45	60	15	50	Flat
4	–27	–50	13	23	10	875	Slope
5	–35	–70	20	35	15	10	Slope
6	–10 (Note 2)	–45	6	35	29	750	Flat
7	–15	–60	35	45	10	300	Slope
8	–0 (Note 2)	–55	40	55	15	800	Flat
9	–15	–40	15	25	10	800	Flat

NOTES

1 Speech echo path frequency responses are defined in 8.4.1.5.

2 Due to peak clipping the speech level will differ from Method B of Recommendation P.56 [3] which assumes that the ratio of peak power to mean power in speech is 18 dB, this is not the case for these conditions.

9.5.3 Circuit conditions with echo cancellers

The three characterization tests specified in 9.4 should be repeated using the conditions in Tables 11 and 12.

TABLE 10/P.561

Asymmetric test circuit conditions for Class C device

Circuit	Near end					Far end				
	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	Speech echo path delay (ms)	Speech echo path frequency response	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	Speech echo path delay (ms)	Speech echo path frequency response
1	–10	–70	40	30	Flat	–20	–65	25	1000	Flat

NOTE – Near end speech level, noise level, loss of the echo path and speech echo path delay are adjusted by A1, N1, A3 and D1 respectively (see Figure 1). Far end speech level, noise level, loss of the echo path and speech echo path delay are adjusted by A2, N2, A4 and D2 respectively (see Figure 1).

TABLE 11/P.561

Test circuit conditions for Class C device with echo canceller

Circuit	Speech level (dBm)	Noise level (dBmp)	Loss of the echo path (dB)	Bulk Delay (ms)	Local speech echo path delay (ms)	Speech echo path frequency response (Note 1)
1	–28	–52	14	734	16	Slope
2	–35	–55	6	92	8	Flat
3	–10 (Note 2)	–70	45	972	28	Flat
4	–18	–50	22	225	25	Slope
5	–21	–70	35	480	20	Flat

NOTES

1 Speech echo path frequency responses are defined in 8.4.1.5.

2 Due to peak clipping the speech level will differ from Method B of Recommendation P.56 [3] which assumes that the ratio of peak power to mean power in speech is 18 dB, this is not the case for these conditions.

TABLE 12/P.561

Reflected speech levels for use with non linear processors

Reflected speech level
Threshold level +10 dB
Threshold level +5 dB
Threshold level
Threshold level –5 dB
Threshold level –10 dB

NOTE – The suppression threshold level will depend on the echo canceller that is being tested.

9.5.4 Circuit conditions with DCME

No pass/fail criteria are specified for the following tests. Their purpose is to characterize the INMDs performance under real network conditions.

DCMEs are used to multiplex voice circuits together, by reducing the bit rate of the speech and not transmitting silence periods only their length. This will have two effects on INMDs:

i) the echo path is degraded by low bit rates;

ii) any unwanted signals are suppressed and comfort noise injected.

For testing the DCME can be broken into two separate devices; a low bit rate codec and a voice switch. The voice switch is similar to the effect of a non-linear processor.

Different levels of traffic loading are applied to the transmit sides of the DCME devices to generate various DCME-related impairments such as front-end clipping and freeze-out. Traffic loading can consist of controlled mixtures of multiple voiceband data channels and speech channels with controlled speech activity factor.

Three DCME loading levels are suggested:

a) Light traffic (4.0 bits/sample, no clipping).

b) Peak-hour traffic (3.7 bits/sample, mild clipping).

c) Heavy traffic (3.0 bits/sample, severe clipping).

9.6 Class D – Future networks

Class D devices are for use on circuits that use non-linear processing or vocoders.

The tests for this class are for further study.

Annex A

Speech material

(This annex forms an integral part of this Recommendation)

This annex gives the details for producing suitable speech material intended for use with the reference circuits described in this Recommendation. It uses, as much as possible, the principles described in Recommendation P.80 [19]. Other useful information can be found in 2.5 of the *Handbook on Telephonometry* [20]. It is desirable that standardized speech material be made available and this is for further study.

A.1 Parameters

The following parameters shall be included to test the INMD as described in 9.2.

A.1.1 Subjects

The following combinations are recommended:

- Male → Male.
- Male → Female.
- Female → Female.

A.1.2 Conversations

A minimum of five different conversations are required (incorporating the combinations of subjects in A.1.1).

A.1.3 Language

The minimum set of languages to be used is:

- English.
- Optionally any other language.

It is suggested that the requirement may not guarantee that INMDs will meet the accuracy requirement given in 8.3 for every language. INMD users should take care to assure that languages appropriate to their networks are tested.

A.1.4 Duration of conversations

Duration of conversation shall be 3 minutes ± 30 seconds.

A.1.5 Activity factor of conversations

Activity factor of conversations in both directions:

- Greater than 25%.

A.2 Telephone connection

The sending and receiving systems shall conform to the Modified Intermediate Reference System (IRS) as specified in [21].

NOTE – This can also be achieved by using wideband recordings and shaping the tape output with the appropriate filter to produce the required frequency characteristic.

The total connection is formed by two Modified Intermediate Reference Systems or equivalent connected to each other with an appropriate PCM filter [14] inserted at the junction.

The following loudness ratings from Recommendation P.310 [22] are recommended:

- SLR, 8 dB;
- RLR, 2 dB;
- OLR, 10 dB;
- STMR in the range 10 to 15 dB;
- LSTR >15 dB.

A.3 Source recordings

A.3.1 Recording environment

The subjects should be seated in separate quiet rooms with volume between 30 and 120 m^3 and a reverberation time less than 500 ms (preferably in the range 200 – 300 ms). The room noise level must be below 30 dBA with no dominant peaks in the spectrum.

A.3.2 Recording system

The recording system shall be of high (studio) quality and can be any of the following:

- A conventional two-track recorder with IEC equalization. High grade tape (low-print through, low noise) should be used at all times.
- A two-channel digital audio processor with a high quality Video Cassette Recorder (VCR) or Digital Audio Tape (DAT) machine.
- A computer-controlled digital storage system.

The third system is preferred.

A.3.3 Recording procedure

The following recording scheme is recommended.

The conversation is recorded from the sending output of the Modified IRS (see A.2), with the handset held in the normal manner.

Two separate channels are used to record the conversation from both ends simultaneously.

Care is taken during the recording that the active speech level is between 20 to 30 dB below the peak overload point of the recording system to reduce the possibility of overload; the recommended level is 26 dB.

A.3.4 Speech level

The active speech level, as defined in Recommendation P.56 [3], shall be measured at the end of the recording.

A.3.5 Subjects

As a minimum there shall be one pair of subjects per language (see A.1.2). The subjects should have no speech deficiencies such as stutter. They should be natives of the language they are speaking.

A.3.6 Conversation task

The conversation task should be meaningful and designed to meet the needs of obtaining the required activity factor (see A.1.5).

A.3.7 Calibration signal

After recording a calibration tone of 20 seconds duration is inserted at the beginning at a r.m.s. level that is a known relationship to the mean active level of the conversation. This tone is normally at 1000 Hz, but can be of any frequency. This process is best done at a re-recording stage.

This tone can then be used later to adjust the mean input levels (see 9.3 for example).

Appendix I

Review of measurement techniques

(This appendix does not form an integral part of this Recommendation)

I.1 Speech echo path delay

I.1.1 Correlation analysis

This method is based upon an analysis of the incident speech in the transmit direction and its correlated reflected speech in the receive direction.

Cross correlation can be thought of as describing the dependence of one waveform on another. Equation (I.1) calculates the cross correlation, r:

$$r = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i (X_i - \bar{X})^2} \sqrt{\sum_i (Y_i - \bar{Y})^2}} \tag{I.1}$$

where X and Y are two time series of data and $\bar{X}$ and $\bar{Y}$ are the mean values.

The correlation coefficient r takes a value between –1 and +1; a value of +1 signifying complete positive correlation which occurs when the two series are identical.

For a time τ there exists a cross correlation function $R_{xy}(\tau)$ which, in its discrete form, is given by the equation:

$$R_{xy}(\tau) = \frac{1}{N - \tau} \sum_{i=1}^{N-\tau} x_i y_{i+\tau} \quad \text{(I.2)}$$

where τ = 0, 1, 2, ..., m; m is the number of lags required and where x_i and y_i are the deviations from the mean values of the two time series given by:

$$x_i = X_i - \bar{X}$$
$$y_i = Y_i - \bar{Y}$$

To calculate the delay between the incident speech signal and its echo, Equation (I.2) is applied and the value of τ that gives the maximum R_{xy} is then the delay.

I.1.2 Adaptive filter analysis

The impulse response is a mathematical representation of the transfer function of one half of the connection. It comprises a set of coefficients whose shape characterizes the connection and whose actual values give an indication of the transmission performance. For example, the echo path loss in a circuit can be determined by a comparison of the values of the coefficients for the circuit under test with those for a circuit with a known echo loss.

The impulse response for an unknown circuit can be calculated from an adaptive filter. An adaptive filter is a device that is self-designing in the sense that it automatically adjusts its coefficient $h(k)$ based upon an estimated statistical characteristics of the input signal. This technique is used in echo cancellers, hence it is worth considering the canceller operation as an example.

Referring to Figure I.1:

$$y(n) = \sum_{k=0}^{\infty} h(k)\,u(n - k) + v(n) \quad \text{(I.3)}$$

where $u(n)$, $u(n–1)$, ..., are speech samples from speaker A, $v(n)$ is the speech signal from speaker B plus any additive noise at time n, and $\{h(k)\}$ is the impulse response of the echo path. The adaptive filter in the echo canceller makes an estimate $\{h'(k)\}$ of the impulse response of the echo path and then estimates the echo as the convolution sum:

$$y'(n) = \sum_{k=0}^{M} h'(k)\,u(n - k) \quad \text{(I.4)}$$

which can be realized as a FIR filter with coefficients $h'(0)$, $h'(1)$, ..., $h'(\mathrm{M})$. The error signal $e(n)$ is formed by subtracting the estimate $y'(n)$ from the return signal $y(n)$, as shown by:

$$e(n) = y(n) - y'(n) \quad \text{(I.5)}$$

The error signal is then used to adaptively control the canceller coefficients $h'(0)$, $h'(1)$, ..., $h'(\mathrm{M})$, so that after a small number of iterations the echo is minimized. When the filter is optimized, the filter holds coefficients which accurately model the transmission path.

Consider the diagrammatic representation of the impulse response shown in Figure I.2 below:

It is possible to determine the delay from this response by analysing the coefficients from $n = 0$ and looking for the peak amplitude coefficient. The value of n to which this coefficient corresponds is then the delay in samples from which the delay in seconds can be calculated.

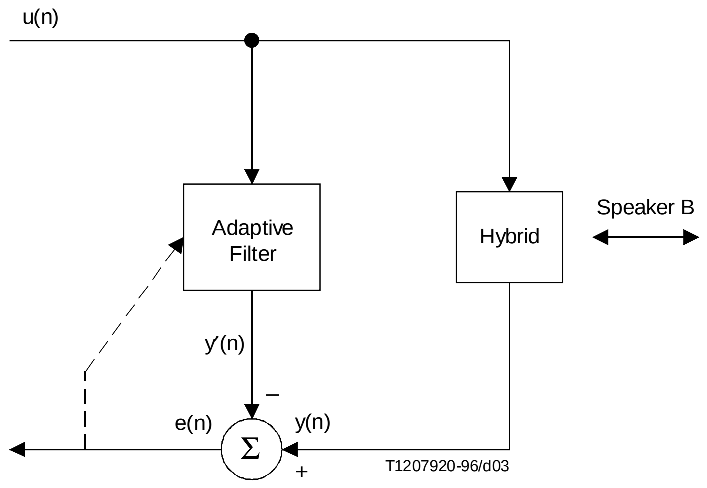

FIGURE I.1/P.561

Signal definitions

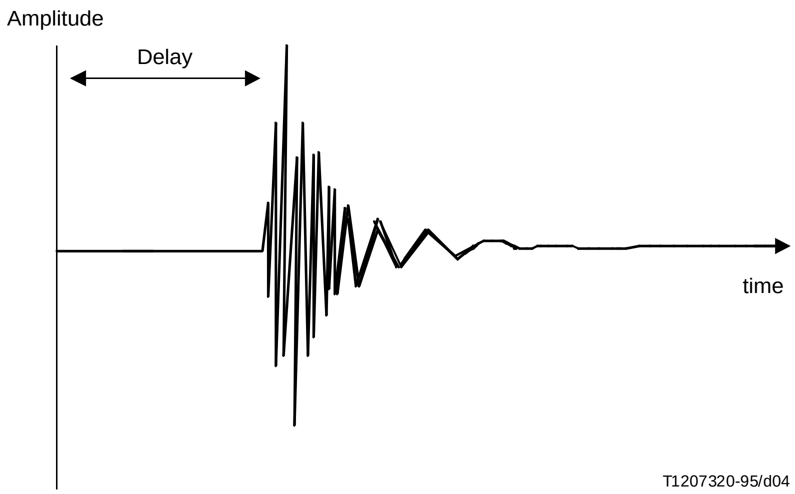

FIGURE I.2/P.561

Impulse response

I.2 Verification of delay measurements

The two methods described above are both search techniques for finding the delay. Therefore a means of determining whether the INMD has found the correct delay needs to be specified. If the delay is determined incorrectly then the echo loss measurement will also be incorrect.

For correlation analysis the correlation coefficient can be used to determine if the correct delay is measured.

For the impulse response monitoring the convergence of the adaptive filter can be used to determine if the correct delay is measured.

I.3 Failure of delay measurement

Should the INMD fail to correctly determine the echo path delay due to failure of the adaptive filter to converge or the magnitude of the correlation coefficient to be low this failure shall be reported.

Such a failure could be caused by excessive intra-system crosstalk; presence of echo cancellers; time varying or non-linear circuits; noise or delays beyond the range of the INMD.

The INMD should be able to distinguish between conditions that may contain an echo path and those which do not. For example:

1) if the failure is caused by excessive intra-system crosstalk; time varying or non-linear circuits; noise or delays beyond the range of the INMD, the circuit may still contain an echo signal and the measurement should be reported as such.

2) if the failure is due to high echo loss beyond the range of the INMD, i.e. due to echo cancellers, the circuit is not in error and the measurement should be reported as such.

Appendix II

Degradation of telephone transmission quality due to non-stationary noise

(This appendix does not form an integral part of this Recommendation)

II.1 Introduction

The "stationary noise" power of idle channels is measured by means of psophometers, devices suited to the purpose; on the other end it is well-known that there is sometimes a small but still important percentage of non-stationary noise present in real channels, that influences the quality and needs a more detailed measurement.

Additionally, since mixed analogue/digital PSTN is foreseen to interwork with ISDNs, actual digital values of stationary and non-stationary noise levels and their statistical distributions in modern telephone networks should be investigated, in order to set suitable limits to public and private operators to safeguard the overall quality of a connection.

Therefore, considering that a wide range of new systems that utilize digital processes and radio transmission are being considered for introduction into telephone networks (examples include PABXs, multimedia and advanced telephone terminals for personal communications, Digital and Packetized Circuit Multiplication Equipments, Mobile Radio and Satellite systems, etc.), and that non-stationary noise (e.g. isolated spikes, transient error bursts, clipping or mutilation effects, signalling crosstalk, etc.) due to both existing systems and new processes, including digital switching, might be present in addition to well-known kinds of noise (quantization, idle, etc.), the ITU-T Study Group 12 has promoted an investigation in the field.

In 1993, a Questionnaire was forwarded by ITU-T to all Administrations, requesting information on noise measurements and methods in use until June 1993.

The questions were:

A What kind of non-stationary noise can be identified in telephone channels?

B What impulsive noise measuring apparatus is possibly in use for routine measurements?

C What kind of measurement techniques are adopted to detect and assess non-stationary noise?

D Are there any level and frequency distributions or other statistics on non-stationary noise in your possession that can be provided to ITU-T for further study?

E What noise level limits are allowed in your national network for existing technologies and what limits are foreseen for possible new services?

From the collection of responses received on the subject, the following types of noise attracted the interest of the operators and could be considered for future INMDs. See Table II.1.

TABLE II.1/P.561

Types of noise of interest for operators
(from the ITU-T SG 12 Questionnaire 1993)

Clicking sound
Crackling sound
Crosstalk
Crowded stadium
Dial pulse
Echo
Fixed telephone network only
“Hash” noise
Hissing
Howl and tone
Humming
Impulsive noise
Induction by humidity
Lightnings
Multi-frequency tones
Noise at frequencies 470 Hz, 1300-1500 Hz and 40 kHz
Noise during handsfree operation
Noise from switching action of analogue switch
Overhead power lines noise
Popping
Rattle
RADAR signals
Radio broadcasted signals
Radio Overhearing
Rotary dialler
Rotating motors
Sharp clicks produced by magnetic induction
Static noise
Thermal and semiconductor noise
Thunderstorms
Tones of exchanges
Tracking problems
Transients
TV cables

II.2 Classification of non-stationary noise

For the time being no automatic methods exist to estimate the effects of all kinds of non-stationary noise on the telephone quality, as perceived from the customer.

INMD could contribute to the achievement of this objective: to the purpose, the development of an algorithm for the classification of the telephone circuit noise would require that essential statistical parameters of the noise signal are measured during an actual conversation.

As regards the classification of non-stationary noise, a major role seems to be played by the so called “impulsive” noise.

Indeed, for this kind of noise it is rather difficult to determine a set of parameters modelling the non-stationary noise and well cut to perform an automatic measure. Computers can make very short term analyses, which is not the case for analogue instruments; therefore, computers are well-suited for detecting impulsive noise.

II.2.1 The noise signal database and its analysis

II.2.1.1 Structures and characteristics

The development of a classifier requires a phase for the study of the available signals, a phase for working out the algorithm and a phase for the performance evaluation of the algorithm itself. Likewise, the used database can be split into three subsets:

- Signals for the study of the problem.
- Signals for the algorithmic development.
- Signals for the performance evaluation.

Each of these subsets has different features depending on its use. The first two of them include analogue signals, while the third is usually made up of PCM signals.

II.2.1.2 Signals for the study of the problem

The files used for the study of the problem include noise segments of different length, taken either from the Public Switched Telephone Network and/or from other networks, e.g. the mobile radio network, and recognized to belong to classes of noise, labelled with easy comprehension names, e.g.:

- Rotary dialling pulses.
- Spikes.
- Isolated impulses.
- Bursts.
- Mixed (impulses+bursts).
- etc.

II.2.1.3 Signals for the algorithmic development of the classifier

Signals from different connections can be collected, the noise samples being successively segmented, measured (e.g. by using a digital psophometer to express every measure in dBmp), and processed through the algorithms under study to show their capability to identify the proper class of noise.

A mere statistical analysis of the mean, mean power and standard deviation on the noise files may not lead to significant results. Then the autocorrelation and covariance functions can be used; if a periodicity is noticed in every file of the database, it will reveal the presence of a predominant frequency (this is the case, for example, of intermittent dialling pulses). LPC analysis and use of templates may be of help.

Modern speech coding or expert "pattern matching" techniques can also be employed to the purpose of developing the classifier.

As the final aim is to predict the impact on the perceived quality of the different classes of non-stationary noise, the algorithm used for the noise classification can be used in combination with objective methods and opinion models, so that each file of detected noise can be classified, measured and considered for its importance as concerns the quality degradation it causes during a normal conversation.

It should be noticed that the great variety of possible kinds of noise seems to conduct towards rather complex (and complicated) algorithms for such classification.

II.2.1.4 Signals for the performance evaluation of the classifier

Signals for the performance evaluation can be either acquired directly from actual conversations, respecting privacy, or built up on purpose in a laboratory. The noisy signals are segmented and classified.

Figure II.1 shows a possible example of non-intrusive instrument allocation to make the acquisition of the PCM stream. A noise discriminator will detect the segments of noise signal to be classified.

Then a verification phase is needed, using the available material, e.g. by calculating, for each class of non-stationary noise, the percentage of correct identification of the proper class, by checking the accuracy of each measure of noise level (or, better, of noise power), etc.

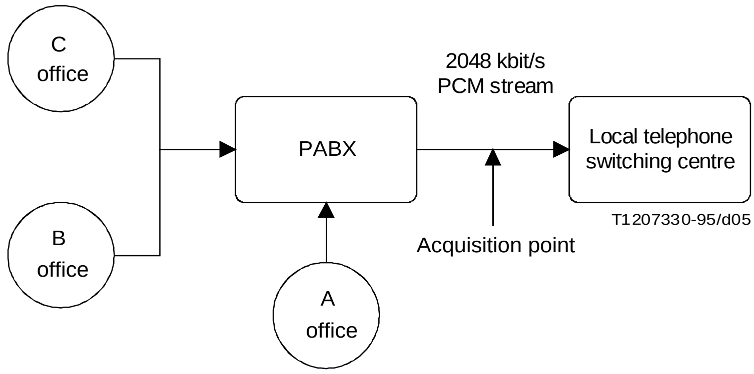

FIGURE II.1/P.561

Example of allocation of INMD for noise classification

Then, if objective measures and/or opinion models are utilized to predict the effect of the classified noises on the perceived quality, subjective tests including those classes of non-stationary noise can be carried out to validate the prediction estimates given from the chosen models.

II.3 Noise classification and measurement algorithm (example)

Figure II.2 gives an example (block diagram) for the classification/measuring of a noise file considering a finite number of classes.

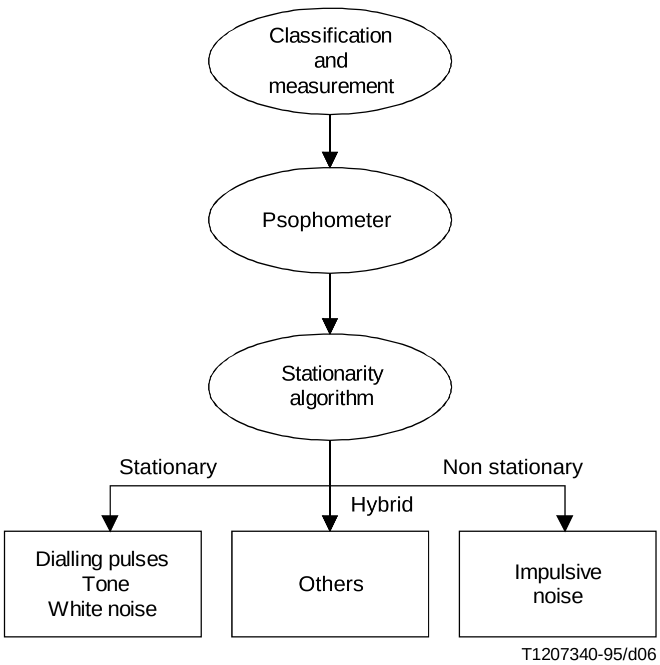

FIGURE II.2/P.561

II.3.1 Main features

First of all, since the noise is psophometrically weighted by the ear, the INMD should do the same. Then the filtered signal will be analysed to estimate its stationarity. As a result it can be classified as:

- Stationary.
- Hybrid.
- Non-stationary.

Depending on this first classification, the following different kinds of noise are distinguished in the above example:

- White (stationary).
- Tone (stationary).
- Decade dialling pulses (stationary).
- Impulsive (non-stationary).
- Others (non-stationary/hybrid).

II.3.2 Stationarity algorithm

Stationarity of noise could be determined running an analysis on the time and frequency domains.

II.3.2.1 Stationarity in time

Stationarity in time of a noise segment can be determined by comparing each short-term average power with the long-term one. Waveforms can always be considered stationary at a very short term. If a segment is stationary but too short it can be classified in a separate class.

Such information can be useful to determine the actual nature of the noise in that frame.

II.3.2.2 Stationarity in frequency

A frequency analysis, determining the LPC coefficients at a short and a long term, can be used, and the stationarity of an active noise segment could be detected by using suitable criteria exploiting, for example, the concept of cepstral distance between each short-term LPC vector and the long-term one.

II.3.3 Stationary noise

White noise and tones can be easily identified because of their characteristics in time and frequency, while the determination of a decadic dialling noise needs more complex analyses.

II.3.4 Non-stationary noise

The recognition of non-stationary noise is of some importance in the development of this activity. Indeed computers may allow very short-term analyses, not possible with the analogue instruments presently available.

II.3.4.1 Impulsive noise

Impulsive noise may be recognized by means of an analysis on the level and duration of each sample, identified as belonging to a burst or a spike, the difference between the two kinds of impulsive noise being determined by their length.

II.3.4.2 Others

If the segment is hybrid or non-stationary (but not recognized as impulsive noise), it may be classified as «others».

II.4 Psophometric measure

Analogue psophometers compute a mean power over intervals of 128 to 200 ms.

The analysis duration is too long for the impulsive noise. The structure of an analogue psophometer can be reproduced by means of a band-pass filter whose specifications follow those described in Recommendation O.41 [15].

II.5 Performance evaluation (example)

The classifier, trained on a set of signals contained in a proper database, could produce results as given in the following examples. See Figures II.3 and II.4.

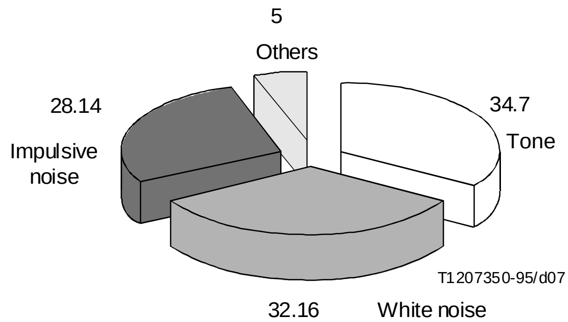

FIGURE II.3/P.561

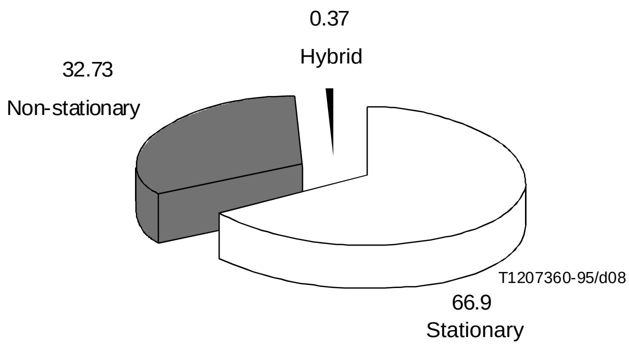

FIGURE II.4/P.561

II.5.1 Analysis of the results

A thorough investigation on the nature and magnitude of the signals could allow some relevant considerations.

For example, if the noise files show higher levels than expected, that could explain why the noise segments of decade dialling signals might be confused with bursts in the classifier, or why portions of the files containing white and impulsive noise could be incorrectly classified.

On the whole the classification should be correct. It can be added that a correct classification depends on the "quality" of the noise. If a segment contains "mixed" noise, the algorithm could detect either one correct kind of noise or classify depending on the frequencies of the noises and on their levels.

II.6 Conclusion

An algorithm for noise classification is useful.

A program can classify only actual circuit noise. In the future the classifier should be supported from a “system” able to recognize echo/circuit noise/environmental noise. In particular, recognizing the environmental noise is really difficult, and at present no algorithm has been found capable to determine if a noise signal is generated by environmental or circuit noise.

The thresholds suggested (provisionally) to classify “active” noise could be –67 dBmp and, for impulsive noise, –38 dBmp, respectively.