INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.562

(05/2004)

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

Objective measuring apparatus

Analysis and interpretation of INMD voice-service measurements

ITU-T Recommendation P.562

ITU-T P-SERIES RECOMMENDATIONS

TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

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

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

ITU-T Recommendation P.562

Analysis and interpretation of INMD voice-service measurements

Summary

This Recommendation provides advice on the analysis and interpretation of INMD voice-service measurements. It describes methods to analyse individual measurement parameters over single and multiple calls. The effects of the location of an INMD on measurements are discussed and the use of customer opinion models, and how these can be used with INMD measurements, described. This Recommendation also looks at how INMD measurements can be applied to network planning and maintenance.

Source

ITU-T Recommendation P.562 was approved on 14 May 2004 by ITU-T Study Group 12 (2001-2004) under the ITU-T Recommendation A.8 procedure.

FOREWORD

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

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

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

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

NOTE

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

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

INTELLECTUAL PROPERTY RIGHTS

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

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

CONTENTS

Page

ITU-T Recommendation P.562

Analysis and interpretation of INMD voice-service measurements

1 Scope

This Recommendation provides advice on the analysis and interpretation of voice-service measurements as produced by an in-service non-intrusive measurement device (INMD). It should be used in association with ITU-T Rec. P.561, *In-service, non-intrusive measurement device – Voice service measurements* [1].

INMDs are utilized primarily for the measurement of voice-grade parameters such as speech, noise and echo. INMDs may also be used to measure parameters associated with digital transmission systems, in both circuit switched and packet switched networks, that impact the performance of the voice-grade channels being transported. 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 telecommunication networks to measure the in-service performance parameters of voice grade services, and to locate and analyse network anomalies. For the circuit-switched network (i.e., using INMDs of classes A, B or C [1]), 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. This stands also for the packet-switched networks (i.e., using INMDs of class D [1]), with connection information as well as protocol information.

This Recommendation is divided into the following clauses. Clause 2 gives a list of references to related standards. Clause 3 provides abbreviations and definitions used within this Recommendation. Clause 4 describes how individual INMD measurements should be interpreted and describes limitations of this method. Clause 5 discusses the impact of Class C INMD location within the network on measurements. Clause 6 shows how INMD measurements can be used to predict average customer opinion, and how these customer opinion predictions should be interpreted. Clause 7 looks at how INMD measurements can be applied to network planning through the use of the E-model [2]. Clause 8 shows how INMD measurements can be used for network maintenance. Full details of the recommended model for predicting average customer opinion are given in Annex A and details of how to map INMD measurements into the E-model are given in Annex B.

2 Normative references

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

[1] ITU-T Recommendation P.561 (2002), *In-service, non-intrusive measurement device – Voice service measurements.*

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

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

[4] ITU-T Recommendation G.131 (2003), *Talker echo and its control.*

[5] ITU-T Recommendation G.169 (1999), *Automatic level control devices.*

[6] ITU-T Recommendation P.800 (1996), *Methods for subjective determination of transmission quality*.

[7] ITU-T Recommendation G.108 (1999), *Application of the E-mode: A planning guide.*

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

[9] ITU-T Recommendation G.113 (2001), *Transmission impairments due to speech processing.*

[10] ITU-T Recommendation G.114 (2003), *One-way transmission time*.

[11] ITU-T Recommendation G.120 (1998), *Transmission characteristics of national networks.*

[12] ITU-T Recommendation G.121 (1993), *Loudness ratings (LRs) of national systems.*

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

[14] ITU-T Recommendation P.79 (1999), *Calculation of loudness ratings for telephone sets*.

[15] ITU-T Recommendation P.76 (1988), *Determination of loudness ratings; fundamental principles*.

[16] ITU-T Recommendation G.223 (1988), *Assumptions for the calculation of noise on hypothetical reference circuits for telephony.*

[17] ITU-T Recommendation G.212 (1988), *Hypothetical reference circuits for analogue systems.*

3 Abbreviations and definitions

3.1 Abbreviations

This Recommendation uses the following abbreviations:

ALC Automatic Level Control

CCI Call Clarity Index

CDR Call Data Record

CME Circuit Multiplication Equipment

CMS Circuit Multiplication System

EC Echo Canceller

EL Echo Loss

EPL Echo Path Loss

INMD In-service Non-intrusive Measurement Device

ISC International Switching Centre

LE Local Exchange

MOS Mean Opinion Score

OLR Overall Loudness Rating

PCM Pulse Code Modulation

RLR Receiving Loudness Rating

SEPL Speech Echo Path Loss

SLR Sending Loudness Rating

TELR Talker Echo Loudness Rating

TSG Trunk SubGroup

3.2 Definitions

For definitions not listed here, please refer to ITU-T Recs P.561 [1] and G.100 [3].

3.2.1 customer opinion model: Customer opinion models aim to predict the average subjective opinion of customers using objective measures.

3.2.2 subjective opinion: A subjective opinion is a personal view and varies from person-to-person. A subjective opinion will not necessarily be the same when repeated under the same external conditions. Call quality is an example of a subjective measure.

3.2.3 objective measurement: An objective measurement is made using measurement equipment and is repeatable given the same external conditions. Active speech level is an example of an objective measure.

3.2.4 E1: A G.703 interface operating at 2048 kbit/s capable of carrying 32 channels of 64 kbit/s each.

3.2.5 T1: A G.703 interface operating at 1544 kbit/s capable of carrying 24 channels of 64 kbit/s each.

4 Interpreting INMD measurements

INMDs continuously monitor the network and have the potential to generate vast amounts of data. Interpreting this data is crucial to understanding the performance of the network being monitored. This clause provides guidelines on how individual voice-service measurements should be interpreted and how measurements from many calls can be collated. For each case the applications, benefits and limitations are described.

4.1 Single call measurements

Voice-service measurements made by an INMD are described in ITU-T Rec. P.561 [1], the minimum required measurements being: speech level, noise level, echo path delay and at least one type of echo loss measurement, plus IP delay variation and IP packet loss ratio for Class D INMDs.

Each of these parameters allows some aspect (or aspects) of network performance to be determined or predicted for that particular call. The parameters measured by an INMD characterize the network connection from each talker to the INMD equipment in that direction only. The network connection in the opposite direction, INMD to listener, is not measured. The only exception to this is the measurement of the echo-path which provides some information about the network connection from the INMD to source of reflected echo (usually the 4-wire to 2-wire hybrid) and back to the INMD. This means that the majority of impairments in the receive path, INMD to listener, cannot be detected by an INMD.

Some aspects of network performance that can be derived from single voice-service measurement parameters are shown in Table 1. Any assumptions made in deriving network performance are also listed.

Table 1/P.562 – Aspects of network performance derived from INMD measurement parameters

Parameter	Aspect of network performance	Assumptions made
Active speech level	Network SLR	Speaker's vocal level
Psophometric noise level	Circuit noise level introduced by the network	Room noise level
Echo loss Echo path loss Speech echo path loss	Operation or presence of echo cancellers Hybrid performance	
Echo path delay	Transmission delay of the connection	Local delay
Speech activity factor	Accuracy of other parameters (e.g., 90% activity in both directions could mean noise is being classified as speech) Type of call (e.g., recorded message)	Normal conversational habits
Front end clipping	Performance of voice activity detectors (e.g., in CMEs)	
Saturation clipping	Amplitude clipping and distortion	
Double-talk	Rough indicator of delay of the connection	Normal conversational habits
One-way transmission	Short network outages	Two-way conversation
IP packet loss ratio	IP node congestion	Correct IP network dimensioning
IP delay variation	IP node congestion IP route flapping	

When interpreting INMD voice-service measurements, call results should ideally be viewed as a set. Investigating single parameters in isolation may give rise to misleading conclusions regarding the quality of the connection. In addition to this, data from a single call is prone to variations in customers' voices and customers' equipment which should be taken into account when considering the data. The examples below show how possible measurements can be misinterpreted.

Example 1

Measurements from a call show that the echo path loss is low and the echo-path delay is low. The low echo path loss indicates that there is significant echo present and the call is assumed to be of poor quality.

However, since the echo path delay is also low the user hears the echo only as sidetone, and perceives the call as good quality. This call would be correctly classified when using a customer opinion model (such as described in clause 6 – Using INMD measurements to predict average customer opinion).

For further information on echo and delay, and guidelines on the control of echo see ITU-T Rec. G.131 [4].

Example 2

Measurements from a call show no discernible noise present and average active speech levels. This initially appears to be a good quality call. Inspection of the speech activity factor reveals speech activity of over 90% in both directions.

The cause of this abnormal speech pattern may be due to high levels of noise being interpreted as speech or high levels of echo being interpreted as speech. In fact, this call may be severely degraded and could require investigation.

Example 3

Measurements from a call show no IP packets are apparently lost, whereas the measurements of IP delay variation from the same call show an important amount of jitter. If the jitter buffer of the IP end-point (VoIP gateway or customer equipment) is configured in such a way that it is not able to cope with this level of delay variation, then some packets will be discarded by the buffer, and will be seen as lost from the customer's point of view.

4.2 Multiple call measurements

Collating measurement data helps to reduce variations due to customers' voices and equipment. When collating measurement parameters, consideration should be given to the purpose for which the information is being gathered.

Measurement data to be used for reporting performance statistics should be selected according to a specific grouping. Recommended groupings to be used are listed in Table 2.

Table 2/P.562 – Recommended groupings for INMD parameter collation

Grouping	Description
Physical link	Single access medium; typically E1/T1 or 10/100 Mbit/s Ethernet
Route	Collection of physical links to the same destination
Province	Collection of routes to a single geographical area within a country
Carrier	Collection of routes delivered by a specific company to a country
Country	Collection of routes to a specific country

Once data on a single parameter has been gathered for a particular grouping, it should be processed using the following procedure:

Stage 1 Exclude all invalid measurements. An invalid measurement is defined as either a default code or a value outside the specified range of the measurement device. However, any default codes that can be reliably translated to a valid measurement should be translated and counted as valid. For example, a default code representing 'no echo detected' could be translated to an echo-path loss equal to the maximum limit of the specified range of the measurement device.

Stage 2 A sample size should be used that gives statistically valid results. For more information see 4.3 – Sample size.

Stage 3 Calculate the sample mean and sample standard deviation.

Stage 4 Calculate the percentage of valid measurements that exceed maximum and minimum pre-set threshold values. Table 3 lists recommended threshold values for the most common parameters. However, due to varying network performance goals, different Administrations may wish to set their own thresholds. The threshold values used should always be stated together with the results.

Stage 5 The following data should be reported:

- Grouping used.
- Date and time over which the measurements were collected.
- Threshold values used.
- Sample size.
- Sample mean, median and standard deviation.
- Percentage of samples above the maximum threshold value.
- Percentage of samples below the minimum threshold value.

Table 3/P.562 – Recommended threshold limits for multiple call measurements

Parameter	Recommended threshold values	
	minimum	**maximum**
Active speech level	–35 dBm0	–6 dBm0
Psophometric noise level	None	–50 dBmp
Echo path loss	15 dB (Note 1) or 35 dB (Note 2)	None
Echo path delay (round-trip) (Note 3)	None	40 ms (Note 1) or 800 ms (Note 2)
IP delay variation	0 ms	200 ms
IP packet loss ratio	0%	10%
NOTE 1 – For connections without echo cancellers. NOTE 2 – For connections with echo cancellers. NOTE 3 – This is the sum of the near and far end delay values.		

4.3 Sample size

When considering confidence intervals for a set of measurements, the following factors influence the minimum sample sizes required:

- Standard deviation of the population distribution.
- Significance level required for the confidence interval (e.g., 95%).
- Required accuracy of the confidence interval (e.g., ±1 dB or ±5%).

There are two different types of analysis: measurement averages and threshold percentages, each requiring different minimum sample sizes. Each of these are considered in the following subclauses.

4.3.1 Confidence intervals for measurement averages

Determining a confidence interval for a measurement average involves calculating the mean and standard deviation of a sample of measurements. The minimum number of samples required depends on the above factors. Examples of using these factors to calculate the minimum sample size required are given in Appendix II. Typical minimum sample sizes are also shown in Table 4.

Table 4/P.562 – Example minimum sample sizes for average calculations

Parameter	Standard deviation dB, % or ms	Confidence interval %	Required accuracy (±)	Minimum sample size
Speech level	7	95	1 dB	188
Noise level	6	95	1 dB	138
Echo path loss	8	95	1 dB	246
Echo path delay	4	95	1 ms	61
IP delay variation	TBD	95	1 ms	TBD
IP packet loss ratio	TBD	95	0.1%	TBD

4.3.2 Confidence intervals for thresholding

Thresholding involves calculating the percentage of measurements above or below a threshold for a sample of measurements. A suitable sample size depends on the accuracy required and the confidence factor that the specified accuracy is met. The number of samples can be calculated theoretically and these calculations are detailed in Appendix III. A graph of minimum sample sizes for confidence intervals, α, of 95% and 98% is shown in Figure 1.

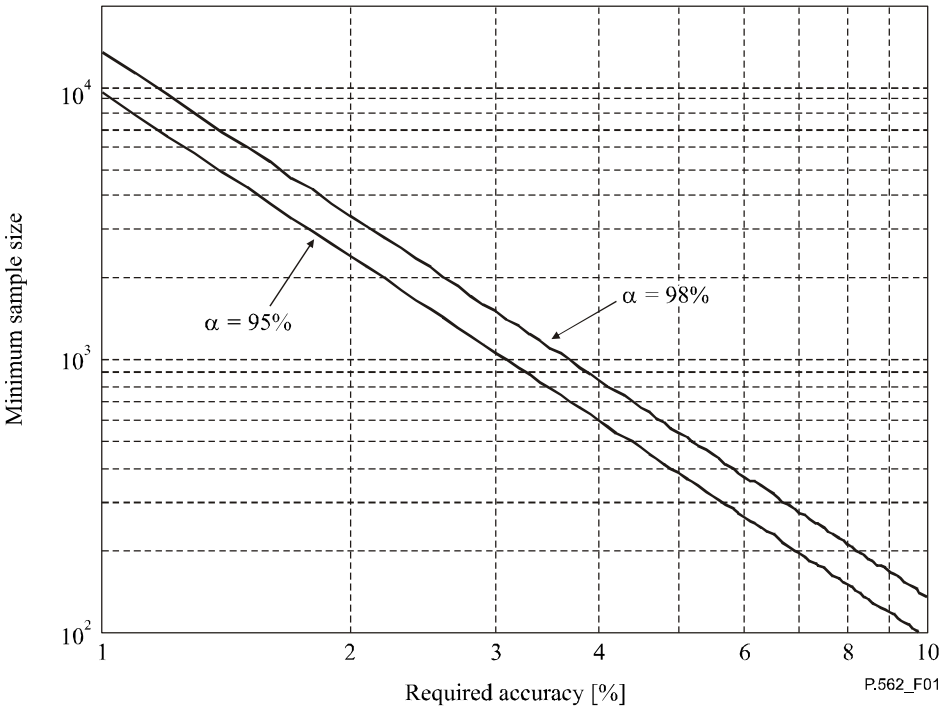

Figure 1/P.562 – Recommended minimum sample sizes for threshold calculations

5 Impact of Class C INMD location in the network

Although as stated in ITU-T Rec. P.561 [1] In-service Non-intrusive Measurement Devices (INMD) of Class C (i.e., for TDM links with echo path delays up to 1000 ms) can be connected at any four-wire DS1 interface on a link they are most commonly installed in international gateways.

In this clause, two different locations are considered for a Class C INMD in an international gateway: at the outgoing side and at the incoming side. For both scenarios the impact of transmission and/or processing devices inside and outside the international gateway on the non-intrusive measurement of parameters is described. Finally, suggested applications for each scenario are given along with the respective properties of interest to a network operator.

5.1 Context

In this discussion, we focus only on situations where the network of an operator, known as the "near end", is connected with the network of another operator, known as the "far end", through an international link. The hypothetical reference connection of such a link can be represented by the elements shown in Figure 2.

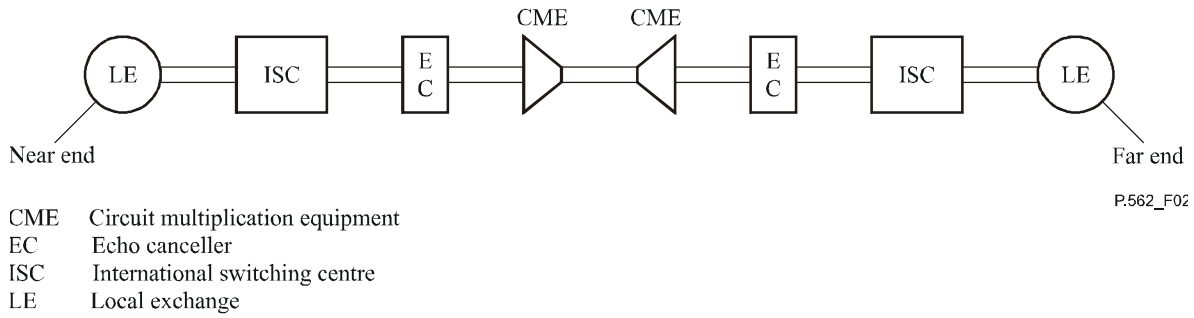

Figure 2/P.562 – Hypothetical reference international connection

In some cases no CMEs or echo cancellers (ECs) are present in the connection, but here we assume that they are present in connections where INMDs are used. The impact of ECs and ALCs in national networks is not considered in the following analysis.

5.2 INMD location at the outgoing side of the international gateway

A common use of INMDs is to connect to an international E1/T1, on the outgoing side of the international switching centre, beyond the echo canceller, as shown in Figure 3.

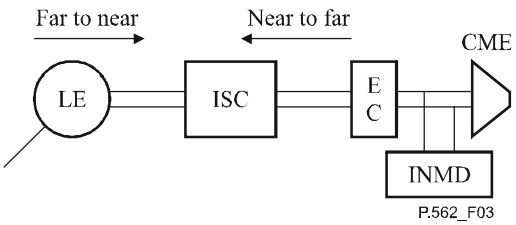

Figure 3/P.562 – Class C INMD implementation at the outgoing side of the international gateway

5.2.1 Advantages

- The measurements are made on both 'near-to-far' and 'far-to-near' directions of the E1/T1 so if a problem is detected, the defective E1/T1 (and the timeslot) is known.
- The measurement of echo is made, taking into account the effects of both the near and far end echo cancellers.

5.2.2 Disadvantages

- The measurements are dedicated to the E1/T1s connected. To monitor quality for all destinations, a high percentage of the E1/T1s must be connected. This requires a large monitoring system which generates huge volumes of data that need to be stored and managed.
- Most systems use signalling information which needs to be decoded. On the international network, several signalling systems still exist (R2, C5, different levels of C7: TUP, TUP+, ISUP) each requiring specific software to be decoded.

5.3 INMD location at the incoming side of the international gateway

An alternative implementation is to connect the INMD to a national E1/T1, on the incoming side of the international switching centre, as shown in Figure 4.

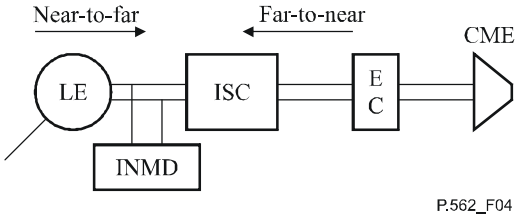

Figure 4/P.562 – Class C INMD implementation at the incoming side of the international gateway

5.3.1 Advantages

- Measurements can be made on calls to many destinations by monitoring a small number of E1/T1s. Choosing a few E1/T1s that are carrying a lot of international traffic, selected to cover a large proportion of the national network, provides the system with a significant volume of calls to many destinations.
- Systems that decode signalling information are less complex since, on most national networks, the number of signalling systems is generally limited to one or two rather than the wider range expected on international systems.

5.3.2 Disadvantages

- If a problem is detected, it is not immediately clear which outgoing E1/T1 is causing the problem. A solution to this problem could be to use Call Data Records (CDR).
- The effects of any near end echo canceller (which affects the person at the far end) are not taken into account. Thus more information regarding near end delay and hybrid performance is gained but no information is available on the performance of near end echo cancellers which can have significant influence on the perceived quality of the call. Measurements of echo and delay are more likely to be made, but do not represent the actual signals reaching the far end listener's ear.

5.4 Comparative impact of both scenarios on the measures recommended in ITU-T Rec. P.561

The implementations described in the previous two clauses can be distinguished not only by their respective advantages and drawbacks, but also by their effect on INMD measured parameters complying to ITU-T Rec. P.561. The effects of equipment, such as echo canceller (EC) [4] and automatic level control (ALC) [5], are shown in Table 5 and described below.

In Table 5, the terms "near to far " and "far to near" are used with the meaning they have in ITU-T Rec. P.561. Thus, the near to far echo path (and its delay and loss) corresponds to the loop followed by an incident speech signal originating from the near end, reflected at the far end and coming back to the near end.

Scenario 1 and 2 represent Class C INMD implementations at the outgoing and incoming sides of the international gateways respectively.

Table 5/P.562 – Comparison of scenarios

	Scenario 1 – Outgoing side		Scenario 2 – Incoming side	
	Near to far	**Far to near**	**Near to far**	**Far to near**
Active speech level Speech activity factor	Measure of levels transmitted by the near end network; includes effects of near end EC and ALC if present.	Measure of levels transmitted by the far end network; does not include effects of near end ALC if present.	Measure of levels from near end access network; does not include effects of near end EC or ALC if present.	Measure of levels transmitted to near end access network; includes effects of near ALC if present.
Noise level (Note 2)	Measure of noise transmitted by near end network; includes noise inserted by EC, ALC and C5 analogue signalling if present.	Measure of noise transmitted by the far end network; does not include effects of near end ALC if present.	Measure of noise from near end access network; does not include effects of near end EC, ALC or C5 analogue signalling if present.	Measure of noise transmitted to near end access network; includes effects of ALC if present.
Echo path delay (Notes 1 and 2)	Measure does not include switching and processing delays in near end gateway.	Measure includes switching and processing delays in near end gateway.	Measure includes switching and processing delays in near end gateway.	Measure does not include switching and processing delays in near end gateway. Echo measured before processed by EC.
Echo loss Echo path loss Speech echo path loss	Measure does not include effects of near end ALC if present. Defective ALC may amplify echo signal.	Measure includes effects of near end EC and ALC if present.	Measure includes effects of near end ALC if present. Defective ALC may amplify echo signal.	Measure does not include effects of near end EC. Echo measured before being processed by EC.
NOTE 1 – For both scenarios, the mean one-way delay value remains the same and is equal to the half of the sum of the two loop delays (if they can both be measured). NOTE 2 – CMEs are part of the international transmission path, and therefore have the same impact on the measurements in both scenarios (comfort noise, additional transmission delay). This is why they are not considered in the table above. The same remark can be made for analogue transmission, which can cause noise and asymmetric loss.				

5.5 Discussion

The difference between the two implementations, presented above, in terms of impact on the measurements results is significant. Depending on a network operator's aim in using an INMD, one of the two solutions will be more suitable. To see this more clearly, consider the following types of international link:

Link 1: with echo cancellers and CME at both sides;

Link 2: with echo cancellers and without CME at both sides;

Link 3: without echo cancellers and without CME at both sides.

For each parameter with each scenario and each kind of link and considering the insertion of ALCs in the international switching centre, it is possible to estimate how the perception of quality is evaluated for the following points of view:

- near end customer;
- far end customer;
- interconnection.

The figures in normal type in Table 6 show how many of the four parameters whose measurement is helped using that scenario. More detailed information for each of the parameters can be found in Appendix I.

Table 6/P.562 – Percentage of measurements which help evaluate the perception of quality

		Scenario 1 – Outgoing				Scenario 2 – Incoming			
		Link 1	**Link 2**	**Link 3**	**TOTAL**	**Link 1**	**Link 2**	**Link 3**	**TOTAL**
With ALC	Near end customer	1	1	2	**33%**	4	4	4	**100%**
	Far end customer	0	1	2	**25%**	0	1	2	**25%**
	Interconnection	3	4	4	**92%**	0	1	2	**25%**
	TOTAL	**33%**	**50%**	**67%**	**50%**	**33%**	**50%**	**67%**	**50%**
Without ALC	Near end customer	4	4	4	**100%**	4	4	4	**100%**
	Far end customer	2	4	4	**83%**	1	1	4	**50%**
	Interconnection	3	4	4	**92%**	1	2	4	**58%**
	TOTAL	**75%**	**100%**	**100%**	**92%**	**50%**	**58%**	**100%**	**69%**

NOTE – Numbers in normal type represent the number of parameters that the scenario helps to measure from a total of four parameters (speech and noise level, echo loss and delay).

Percentages in bold type represent the total percentage of parameters that the scenario helps to measure based on a maximum of four parameters per condition.

If we compare the overall performance of both scenarios, scenarios 1 and 2 are equivalent when ALC is present and scenario 1 provides more useful information than scenario 2 without ALC. Comparing performances for each type of application reveals significant differences between scenarios.

If the aim of transmission quality monitoring is mainly for the supervision of interconnections with other networks, the implementation of INMDs on the outgoing side (scenario 1) is the best solution.

However, the implementation on the incoming side (scenario 2) provides an operator with the most useful information on the voice quality as perceived by near end customers when ALCs are present.

When ALCs are not used, then scenario 2 provides no benefits over scenario 1 in terms of measurements.

The consistent difference between the two implementations is cost. It is cheaper to implement an INMD on the national side (scenario 2) so if two interconnected carriers both implemented systems on the national side and shared the results, this would provide a more accurate indication of the quality perceived by the users at each end of the network. However, locating any faults on the outgoing side would be harder when using in INMD on the national side.

5.6 Conclusion

The analysis on Class C INMD location within an international connection shows that there is no unique use of such a measurement device, and that, according to the ultimate aim of a network operator, each scenario has advantages and disadvantages.

To summarize the advantages of each scenario:

- the implementation at the outgoing side provides the most accurate indication of quality provided by the near end network to the far end network and provides useful information about the operation or presence of echo cancellers;
- the implementation at the incoming side provides the most accurate indication of quality provided to the near end network user in the specific case where ALC is enabled on the near end. This implementation can also be a cheaper solution.

A more complete evaluation of the end-to-end speech transmission quality can be achieved if two interconnected carriers share their measurement results and the knowledge of the characteristics of their respective near end access network.

6 Using INMD measurements to predict average customer opinion

Individual measurement parameters, by themselves, do not provide a complete picture of the connection. A customer opinion model can be used to encapsulate all available information from many measurement parameters into a single figure quality prediction.

The recommended model for predicting average customer opinion from measurements made by an INMD of classes A, B or C is described in this clause. The model, known as the Call Clarity Index (CCI), has been specifically designed for use with non-intrusive measurements and has been shown to be more robust than using planning models for this purpose.

As far as INMDs of Class D are concerned, there is currently no customer opinion model integrating all mandatory measurements of P.561. Nevertheless:

– the model described in this clause is also applicable to class D, provided that the impact of IP impairments is negligible;
– if the impact of IP impairments only (mandatory and optional measurements) is addressed, the parametric model currently studied by ITU-T SG 12, Question 16 group, and soon available as ITU-T draft Recommendation P.VTQ, gives the possibility of predicting customer opinion.

6.1 Using a model to predict customer opinion

Customer opinion models attempt to map objective measures of network performance to subjective opinions. A customer opinion model for INMDs should, therefore, be able to relate the network performance (as represented by the objective measurements such as speech level, echo loss, etc.) to customer perceived performance (represented by an opinion score).

Benefits of using a model to interpret INMD measurements include:

1) The identification of combination effects that are incorrectly classified when using individual measures.
2) Reduction in data volume (a single figure now represents the measured quality compared to many individual measurements).
3) The model encapsulates expert knowledge about the effects of impairments on customer perception.

Benefit 1 is illustrated in Figures 5 and 6 below. Here, as an example, just two parameters are considered, echo loss and echo path delay, whereas, in reality, the problem is multidimensional. Use of independent thresholds for each parameter would identify the measurement combinations marked with crosses in Figure 5 as failures, and the ticks as passes. The true perceptual threshold, however, would look more like the curved line and result in the passes and failures shown in Figure 6. It can be seen that several false positives and false negatives can be avoided by the use of a model, helping to enable more efficient application of investment in network repairs or upgrade.

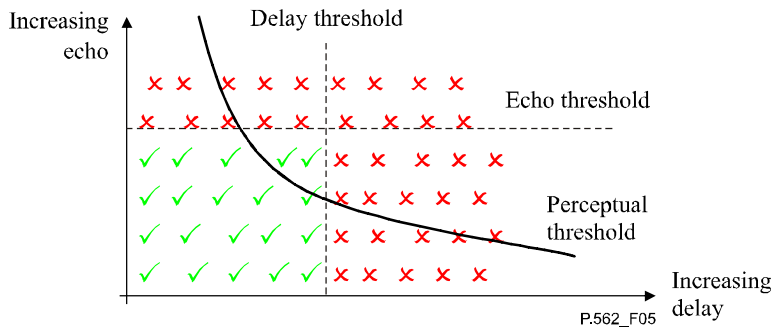

Figure 5/P.562 – Using individual measures for thresholding

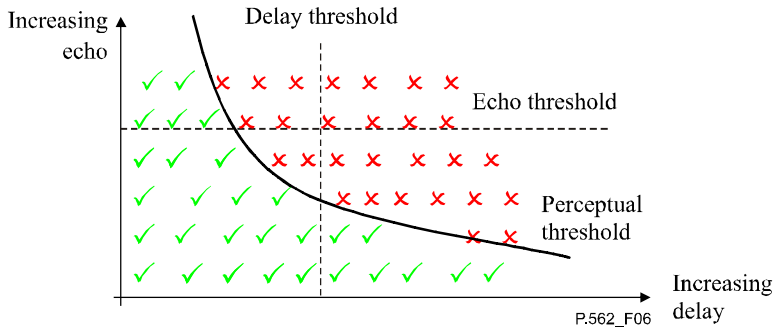

Figure 6/P.562 – Using a model for thresholding

6.2 Assumptions

Figure 7 illustrates the location of the INMD at a non-intrusive four-wire monitoring point. INMDs measure the speech level (SL), noise level (NL), echo loss (EL), and echo path delay (EPD) of both directions of a connection. These parameters can be used to derive the impact of loss, noise, and echo on customer opinion. Because INMDs do not make end-to-end performance measurements, however, it should be noted that using INMD measurements in a customer opinion model requires estimating some parameters which cannot be derived from the INMD's measurements. In particular, referring to Figure 7, the INMD's far SL measurement (SL_f) can be used to derive the combination of the far SLR (SLR_F) and the transmit loss in the far to near direction (T_F) provided that an assumption about the far end vocal level is made. However, the near RLR (RLR_N) and the receive loss in the far to near direction (R_N) cannot be derived from the INMD's measurements because they affect the connection's performance after the point at which the INMD makes measurements. These parameters have to be estimated by the user and typically are selected to represent an average value or distribution of values for the network being evaluated.

The model has been designed to reduce the dependence on assumptions as far as possible. The main way it does this is by allowing direct input of the speech level SL and noise level NL measurements into its core loss/noise model. The primary effect on call clarity of noise masking speech is independent of any assumptions about send vocal level or absolute send loss (although an average frequency shape is assumed) within the model.

Subjective effects where the absolute loudness needs to be known (such as echo loudness) do still require assumptions in order to model them appropriately. These assumptions need to be chosen to represent the typical or average expected conditions on the network.

Measured SL is dependent upon send loss (SLR+T) and send vocal level (VL). The send loss and the send vocal level will have a statistical mean and a distribution about that mean. The measured SL will, therefore, usually differ from the average expected SL. The difference in levels could be caused by one of two things:

a) the speaker talking at a different level than assumed; or

b) the network loss (SLR+T) being a different value than assumed.

The model takes this into account when estimating network loss.

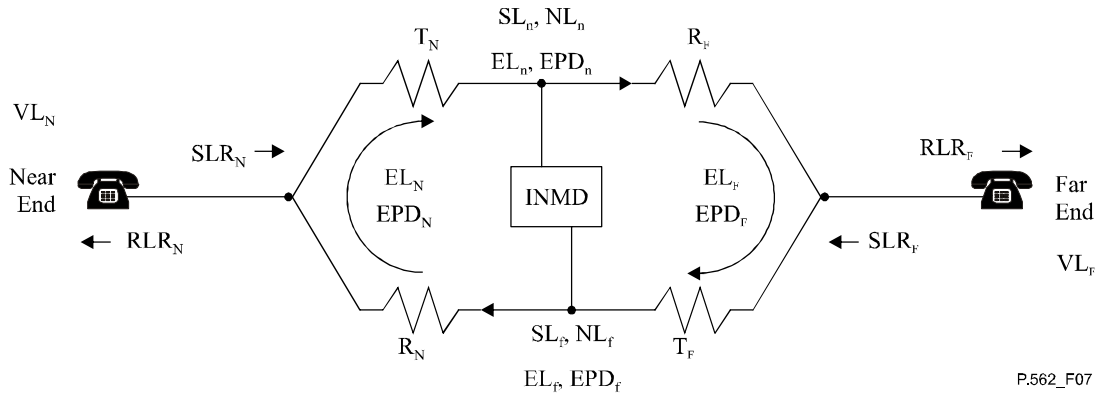

Figure 7/P.562 – Network diagram

The full set of assumptions used by the model are described together with the model itself in Annex A – Call Clarity Index model description.

Some of the assumptions can be network and country specific. Included in this Recommendation is a suggested set of values for the assumptions based on different European countries' data. In the absence of a comprehensive set of data for each country or region, it is recommended that the standard set of assumptions contained within this Recommendation be used. This allows accurate comparison of average predicted opinions between results taken on the same route at different times but care should be taken when comparing different routes. Typically, if a country has different average losses to those in the standard set of assumptions, then its averaged predicted opinion score will be offset compared to another country that has assumptions closer to the standard set. However, the distribution of predicted opinion scores around this average will still reveal highly valuable information.

6.3 Model

The Call Clarity Index (CCI) model predicts the call clarity (also known as speech transmission quality) from INMD measurement parameters on a call-by-call basis. This clause describes the functional blocks, shown in Figure 8, that form the CCI.

The overall operation of the CCI is to use the non-intrusive measurement parameters in conjunction with assumptions about the network and the users at either end to predict the signals arriving at each user's ear. These predicted signals, along with knowledge of the human auditory system, are then transformed into listening and conversational speech quality opinion predictions of the call as perceived by each user.

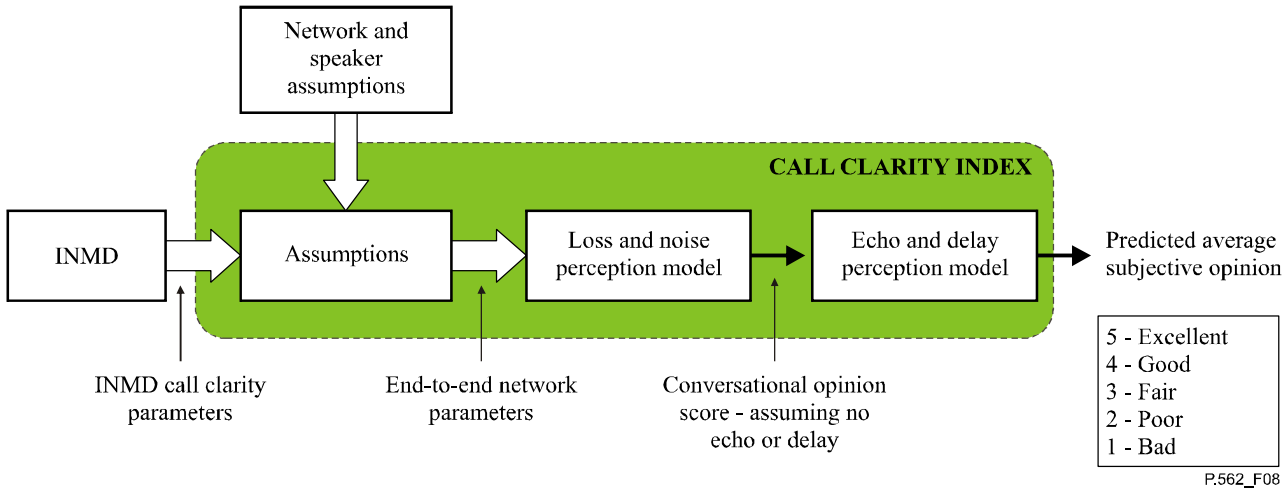

Figure 8/P.562 – The functional blocks that form the CCI

6.3.1 Network and speaker assumptions

To be able to predict the call clarity of a telephone call as perceived by the customers at either end the model requires the following information that is not available from INMD measurements:

1) The overall sensitivity-frequency response characteristic of each transmission path (talker's mouth-to-INMD and INMD-to-listener's ear).

2) The sensitivity-frequency response characteristic of each sidetone path (each talker's mouth to his own ear).

3) The room noise spectra and levels at each end of the connection.

4) The average speech spectrum and threshold of hearing.

6.3.2 Assumptions model

The assumptions model takes the INMD call clarity parameters together with the network and speaker assumptions to produce a complete description of the end-to-end network. From this description, the signal levels at each listener's ear can be calculated. Measuring speech level and noise level in the centre of the network means that only the path from the INMD to the listener's ear needs to be completely assumed. This reduces the amount of uncertainty in the model's predictions.

6.3.3 Loss and noise perception model

The loss and noise perception model accounts for the frequency selectivity of the human ear and noise masking effects on the connection. Sidetone and room noise are also accounted for in the prediction of listening effort and conversational speech quality.

Firstly, the listening opinion index (LOI) is calculated for each listener. This takes into account the effects of loss from speaker to listener and the masking of speech by noise. The listening opinion index is then transformed to a listening effort score and finally to a prediction of conversational speech quality.

6.3.4 Echo and delay perception model

The echo and delay perception model modifies the conversational speech quality prediction to account for delay or echo present on the connection. The effects of echo and delay are taken into account by considering the power of the echo reaching the listener's ear in combination with the delay, sidetone level and overall loss of the connection.

6.3.5 Model output

The outputs from the CCI are predictions of conversational speech quality for an average user at each end of the connection. The predictions are on a continuous scale from one to five based on the ITU-T scale of conversational speech quality in ITU-T Rec. P.800 [6] shown in Table 7.

Table 7/P.562 – Speech quality scale

5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

6.4 Predictions

This clause gives guidance on how to interpret single and multiple CCI predicted quality scores. The CCI produces two scores, one for each end of the connection. These should be combined separately and reported separately.

6.4.1 Single call clarity index values

The output of the CCI represents the predicted subjective opinion of an average customer at either end of the connection. This prediction is based on assumptions made about the network and the users at either end. If the actual network and speaker conditions vary greatly from these assumptions, then the accuracy of the predicted score will be compromised.

A single CCI value can be taken as an indicator of a poor quality call, but should not be used on its own as a measure of network performance.

6.4.2 Multiple call clarity index values

Combining many CCI values provides benefits due to statistical averaging. However, comparisons between different countries should be treated with caution since the relevance of the assumptions made may vary from country to country. Trends over time on a country-by-country basis are a more useful indicator of network performance.

Multiple values should be combined using the methods described below.

6.4.2.1 Averages

A sample size should be used that gives statistically valid results. For more information see 4.3, Sample size. The sample of CCI values from a particular grouping (see Table 2) is taken and the statistical average and standard deviation of the sample calculated. The following data should be reported:

- CCI assumptions used.
- Grouping used.
- Date and time over which the CCI values were collected.
- Sample size.
- Sample mean and standard deviation for each end of the connection.

Taking the average value shows changes in overall trends. The average will not reveal a small number of very poor calls and hence some measure of distribution is useful.

6.4.2.2 Distribution

A sample size should be used that gives statistically valid results. For more information see 4.3, Sample size. The sample of CCI values from a particular grouping (see Table 2) is taken and the percentage of CCI values exceeding pre-set thresholds calculated. Table 8 lists recommended threshold values to be used.

Table 8/P.562 – Recommended CCI threshold values

Threshold name	Value
Upper threshold	3.5
Lower threshold	2.5

The following data should be reported:

- CCI assumptions used.
- Grouping used.
- Date and time over which the CCI values were collected.
- Sample size.
- Threshold values used.
- Percentage of CCI values above and below each threshold for each end of the connection.

Note that temporary irregularities, possibly due to network faults, become less distinct as the period of time over which data is collected increases. This has to be balanced with the need to make sufficient measurements.

Looking at the percentage of calls with CCI values above or below certain thresholds gives more information about the distribution of call quality and can reveal the presence of a small number of poor calls.

6.5 Diagnostics using an opinion model for INMDs of classes A, B or C

As well as indicating that the quality of a connection is poor, it is also desirable to know the reason for this reduced quality. An opinion model, because it combines all information for a call, has the potential to act as a fault diagnosis platform.

Any diagnostic measure should indicate (for each single call) the probability that the poor quality is due to each of the following factors:

- Signal level.
- Total noise level.

  This could, if possible, be further split into:

  – Circuit noise level.

  – Room noise level.
- Echo level.
- Delay.
- Amplitude saturation.
- Temporal clipping.
- Non-linear coding.

The source of the problem should also be indicated as:

- Near end (e.g., national).
- Far end (e.g., international).

Wherever possible, this should be stated as one of the following sub-categories:

- Access network.
- Echo canceller.
- Automatic level control.
- CME.
- Noise cancellation.
- Other.

The probability should be stated from zero to one. Zero indicating that there is no possibility that this is the cause of poor quality, and one indicating that this factor is definitely the cause for poor quality. A value of –1 should indicate that there is no method of knowing whether or not this factor is the cause of poor quality

7 Using INMD measurements for network planning

Outlined in this clause is a method for using INMD measurements in the E-model [2], which is the ITU-T recommended model for network planning. Annex B provides equations for mapping INMD parameters to some of the parameters used by the E-model. In particular, the INMD's active speech level, noise level, IP delay variation, IP packet loss ratio (together with some other easy to find IP information), echo loss, and echo path delay measurements are mapped to the SLR, noise, end-to-end delay, low bit-rate coding, TELR, and echo path delay parameters used in the E-model. To get ratings for the performance of the end-to-end loss, noise and echo performance of connections though, the RLR and receive loss of connections must be estimated from network averages because they are not included in the INMD's measurements.

Analyses of the ratings produced using this mapping have shown that they do accurately measure the performance of connections. When ratings derived from multiple INMD measurements from a network are averaged, they can be used to accurately evaluate the performance of the network. This will provide network planners with a useful tool for determining how changes to the loss, noise, echo loss, or delay in their network will affect performance.

ITU-T Recs G.108 [7] and G.109 [8] provide guidance for using subjective ratings from the E-model to do network planning. The guidance provided by G.108 and G.109 could be used to assess the acceptability of the performance of a network or route or to plan changes to a network.

Because the E-model can also additively include the impact of other impairments not measured by the INMD on performance, the mapping provided in Annex B can be used to assess how adding new technologies to a network will affect performance. This technique can be used to determine how adding echo control devices, low bit rate codecs, circuit multiplication systems (CMS), or other technologies to connections will change performance.

8 Using INMD measurements to maintain networks

This clause provides two subclauses (8.1 and 8.2) that give techniques for determining when networks require maintenance. The first subclause (8.1) discusses how the guidance of some of the Recommendations in the G series can be used to provide objectives for assessing network performance and determining when maintenance is required. The second subclause (8.2) provides techniques for using INMD measurements to set thresholds that can be used to direct maintenance activities.

8.1 Objectives provided by Recommendations in the G series

A number of Recommendations in the G series provide objectives for assessing network performance. In particular, ITU-T Recs G.113, G.114, G.120, G.121, G.122 and G.131 ([9] to [13] and [4]) include objectives that can be used with measurements made by INMDs. As was mentioned in the previous clause, ITU-T Recs G.108 [7] and G.109 [8] also provide objectives for using subjective ratings from the E-model to assess performance.

This clause discusses how those objectives can be used with INMD measurements to determine whether performance is acceptable. Although the objectives given in the Recommendations discussed in this clause provide general guidance for performance planning and assessment, network providers may want to determine their own objectives for acceptable performance levels based on experience.

8.1.1 Objectives for speech level measurements

ITU-T Rec. G.121 provides objectives for the SLRs of national networks. It indicates that the long term objective for the SLR is 7 to 9 dB, but that some networks that cannot achieve this objective can use a short term objective of 7 to 15 dB. ITU-T Rec. G.121 also provides objectives for the minimum (+2 dB[1]) and maximum (17 dB) SLRs for national networks.

The speech levels measured by INMDs can be converted to SLRs using the mappings provided in Annex B. Averaging a number of SLRs determined in this manner using the techniques described in the previous clause will provide an accurate estimate of the SLR of a particular route or in a particular network. This estimate can be compared to the objectives provided in G.121 to determine if the SLR is acceptable.

8.1.2 Objectives for noise measurements

ITU-T Rec. G.120 provides objectives for circuit noise in national networks and on international circuits respectively. This Recommendation provides objectives based on the type of circuits in a connection (analog or digital) and the length of the circuits. INMD noise measurements over a route or network could be averaged and compared to these objectives to determine if noise performance was acceptable.

The objectives given in ITU-T Rec. G.120 are primarily driven by analog circuits which have worse noise performance than digital circuits. Idle channel noise in a connection with all digital links should be quite low. In terms of customer perception, a noise level of –62 dBm0p is considered audible but not objectionable. On a fully digital network or route, initially setting a threshold at –62 dBm0p would assure acceptable performance. This threshold could be adjusted based on user experience with INMD noise measurements on the network or route.

However, it should be noted that the measurement of idle channel noise by an INMD is complicated for two reasons. INMDs actually measure a combination of idle channel noise, any ambient noise that may be transmitted from the distant end, and CME noise fill if compression devices are present.

Ambient noise may or may not get transmitted through the network depending on its loudness. If it is very low, it could be filtered by mechanisms in the network such as the speech detector in a CME. However, loud ambient noises would be transmitted. One method to estimate the impact of ambient noise would be to analyse measurements on a per circuit basis and examine the high end of the noise distribution. If the noise measurements on a circuit are generally well behaved, but there are a few outliers, these outliers could have come from high ambient noise.

In addition, CME often introduce comfort noise if speech interpolation is active. In this case, the INMD is often measuring the power of the CME's noise fill and not the actual idle channel noise.

---

[1] This objective is still under further study.

The above points make setting thresholds for noise complicated. The threshold must be set so that it is above the noise fill level. However, it must also not be set too high so that genuinely noisy circuits escape detection. These points indicate that users will need to base thresholds for determining if noise levels are acceptable on their experience with INMD noise measurements on the route or network being analysed.

8.1.3 Objectives for echo power measurements

ITU-T Rec. G.122 provides objectives for echo loss (EL) in national networks. Currently no objectives are provided for EPL or SEPL measurements. However, the objectives for EL provided in G.122 could be used for initial guidance with these measurements and users could then develop objectives based on their experience with EPL and SEPL measurements. Because the EL and EPL are not affected by variations in speech level, SLR, and background noise, individual measurements can be used to identify performance problems. SEPL measurements, however, are affected by variables outside the control of network providers and must be averaged, like the SL and noise measurements, to limit the impact of these variables.

For circuits that do not include echo control, ITU-T Rec. G.122 provides objectives which vary with the number of 4-wire analog or mixed analog-digital circuits in the connection being measured. In many modern networks, there are no circuits of this type. In this case, ITU-T Rec. G.122 indicates that the mean EL measured should be no less than 15 dB and the standard deviation of the EL measurements must not exceed 3 dB. Users of INMDs may also want to set a threshold for a minimum acceptable EL on any measurement over a route. This threshold could initially be based on the expected distribution of ELs, the loss engineered into the connection being tested, and the double talk sensitivity margin of ECs which may be used on the connection.

For connections with long round trip delays (exceeding 100 ms), different EL objectives will be required. ITU-T Rec. G.131 indicates acceptable talker echo loudness ratings (TELRs) as a function of transmission time. TELR can be calculated as the sum of SLR, RLR and EL, so if nominal values of SLR and RLR are assumed as 7 dB and 3 dB respectively, then EL can be calculated as TELR – 10 dB. For connections with 100 to 200 ms of round trip propagation delay, the minimum acceptable TELR is in the range 40 to 47 dB giving a minimum EL in the range 30 to 37 dB (assuming SLR and RLR values of 7 and 3 dB respectively). Such connections should have active echo control. It may be beneficial to set a uniform threshold independent of the echo path delay for such connections. Since G.131 also indicates that connections with round trip delays exceeding 50 ms should have active echo control, a reasonable initial objective would be to have ELs exceeding 35 dB on any connection with a round trip EPD exceeding 50 ms.

8.1.4 Objectives for echo path delay

ITU-T Rec. G.114 provides objectives for absolute delay and ITU-T Rec. G.131 provides guidelines for the EPD levels at which active echo control is needed. ITU-T Rec. G.114 indicates that round trip delays under 300 ms are acceptable for most applications, delays of between 300 and 800 ms are acceptable, but impact performance, and delays exceeding 800 ms are unacceptable. Users could develop objectives for round trip delay based on these objectives and their knowledge of the facilities used in the connection being tested.

ITU-T Rec. G.131 states that connections with round trip delays of 50 ms or less may not require active echo control. However, it may be beneficial to consider implementing echo control when the two-way delay exceeds 40 ms. The rationale for this more conservative approach is that customers appear to begin noticing and complain about echo at 40 ms.[2]

---

[2] *Assessment of customer perception of echo from INMD measurements*, France Telecom, QSDG 14/97-21.

8.2 Additional thresholding techniques used to direct maintenance

In addition to using thresholds to evaluate the performance of individual measurements or the average of a number of measurements, it is also useful to develop thresholds based on the distribution of a group of measurements. By focusing on the worst measurements made over a route or network, sources of performance problems can be identified. A model for directing maintenance actions based on this concept is the pre-baseline corrective action model.

The pre-baseline corrective action model can be used before the baseline performance of a route or network is known. The model ranks trunk subgroups (TSG) by the percentage of measurements made on the TSG which exceed a threshold. An example is to rate TSGs on a particular route high, medium or low. Where, in a case where EL was being evaluated on a long connection that included active echo control, high would indicate that 90% of the calls over the route monitored by INMDs had a measurable echo. Medium would indicate between 30% and 89% had measurable echoes and low would indicate less than 30% of the calls had measurable echoes.

When a suitable number of calls on the route had been monitored, the TSGs ranked in the high category would be referred for further analysis and possible maintenance. If very few or no TSGs were in the high category, the threshold for being classified in the high category could be reduced to allow a reasonable number of TSGs to be identified for further analysis. Over time, the threshold could potentially be reduced as problems are fixed, until a point is reached where performance is at a level where maintenance is not required.

Similar thresholding techniques could be used for SL, noise, and EPD measurements.

A second model for directing maintenance activities is the baseline exceptions model. As an INMD program matures, a substantial database gets created and the baseline performance of routes/networks becomes known. An exceptions model can now be utilized to drive maintenance actions. This can be done in two ways: exceptions on similar routes or networks and historical exceptions on a route or network.

Exceptions on similar routes or networks addresses the question: is the route that is currently being monitored significantly different than other similar routes? Hypothesis tests, such as those described in Appendix II, could be used to compare two routes or networks, or users could set a threshold representing a significant difference based on experience. One possible initial threshold that could be used is differences of larger than 1 standard deviation. Routes that were significantly worse would be referred for further analysis and possible maintenance.

Historical exceptions identify cases when a route or network has a significant degradation over time. The same process of hypothesis testing or threshold setting described above could be used to determine when a significant degradation has occurred and a route or network should be referred for further analysis and possible maintenance.

Annex A

Call Clarity Index model description

A.1 Introduction

The Call Clarity Index is described in terms of its equations and assumptions. The equations and assumptions relate to different ends of the network and different directions of transmission and so it is necessary to define naming conventions for each quantity. Each end-dependent, or direction-dependent, quantity is defined for each end, or direction, by appending a superscript A or B denoting either end A or B. Directional quantities are described by the end they originate from. For example the OLR from end A to end B is termed OLR^A and the INMD measured speech level originating from end A is termed SL^A. This A/B notation can be converted to Near/Far end notation by replacing A with N (Near End) and B with F (Far End). Figure A.1 shows how the major quantities used are named. These quantities are explained throughout this annex.

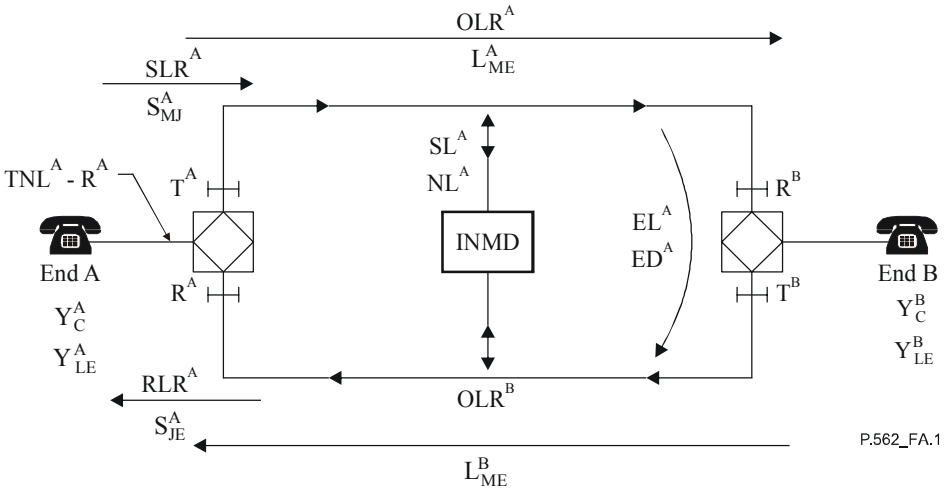

Figure A.1/P.562 – Naming conventions for directional quantities used in the CCI

The inputs to the model from the INMD are as follows:

Name	Description	Units
SL^A	INMD measured active speech level in direction A → B	dBm0
SL^B	INMD measured active speech level in direction B → A	dBm0
NL^A	INMD measured psophometric noise level in direction A → B	dBm0p
NL^B	INMD measured psophometric noise level in direction B → A	dBm0p
EL^A	INMD measured echo path loss for echo path A → B → A	dB
EL^B	INMD measured echo path loss for echo path B → A → B	dB
ED^A	INMD measured echo path delay for echo path A → B → A	ms
ED^B	INMD measured echo path delay for echo path B → A → B	ms

The outputs from the model are as follows:

Name	Description	Units
Y_C^A	Conversational speech quality as perceived from end A	1 to 5 scale [Bad, Poor, Fair, Good, Excellent]
Y_C^B	Conversational speech quality as perceived from end B	

Much of the model uses frequency-dependent data. This data is described in terms of the ISO frequency bands, shown in Table A.1. For narrow-band calculations, bands 4 to 17 are used while, for wideband calculations, all the bands (1 to 20) are used. Most of the calculations in the model are narrow-band, but some, STMR for example, need to be wideband.

Table A.1/P.562 – ISO defined frequency bands

Band	Centre frequency (Hz)	Band	Centre frequency (Hz)
1	100	11	1000
2	125	12	1250
3	160	13	1600
4	200	14	2000
5	250	15	2500
6	315	16	3150
7	400	17	4000
8	500	18	5000
9	630	19	6300
10	800	20	8000

All the equations shown are defined to predict the customer's opinion at end B of the connection. To change the orientation to predict the customer's opinion at end A, all references to end A should be swapped to end B and vice versa.

The assumptions made by the CCI are shown here in the form of a reference connection, which consists of the following components (Table A.2):

Table A.2/P.562 – Reference connection components assumption for CCI

Description	Quantity
Average speaker's speech spectrum density	β'_S
Average network SLR	SLR_{REF}
Average network mouth-to-junction sensi tivity for 0 dB SLR	$S_{MJ\ REF}$
Average network RLR	RLR_{REF}
Average network junction-to-ear sensitivity for 0 dB RLR	$S_{JE\ REF}$
Average listener's threshold of audibility	β_0
Average STMR	STMR
Average sidetone mouth-to-ear loss (frequency dependent)	L_{MEST}
Average handset microphone excess sensitivity to room noise	Δ_{SM}
Average room noise spectrum density and level	RNS, RN
Average circuit noise spectrum density	CNS

A.2 Model equations

The output from the model is a prediction of customer opinion, Y_c, for each end of the connection. The equations in this annex show how to calculate the predicted customer opinion for end B, Y_C^B.

$$Y_C^B = 1 + (E^B . Y_{Cpre\text{-}echo}^B) \tag{A-1}$$

where:

E^B is an 'echo & delay' multiplier, between zero and one, to modify the pre-echo opinion score to take account of echo and delay impairments

$Y_{Cpre\text{-}echo}^B$ is the calculated pre-echo opinion score, on a zero-to-four scale, which takes into account effects of noise and loss

The addition of one converts Y_C^B to a one-to-five scale. All intermediate opinion score values are based on a zero-to-four scale for ease of calculation.

A.2.1 Echo and delay multiplier, E

The echo and delay perception model modifies the opinion score output from the loss and noise perception model to account for the effects of echo and delay on the quality of the connection.

The effects of echo and delay on perceived opinion are incorporated through the use of a multiplier value from zero-to-one. This multiplier is based on predictions of opinion scores for standard connections with variable delay. The multiplier value, shown in Equation A-2, is the ratio of the predicted score under the specified echo and delay conditions (y_{MOPT}) to the predicted score at zero delay (y_0).

$$E^B = \frac{Y_{MOPT}^B}{Y_0^B} \tag{A-2}$$

$$MOPT = \frac{ED^A + ED^B + LD^A + LD^B}{2} \tag{A-3}$$

where:

LD^A, LD^B are the assumed local delay values at each end of the connection (in milliseconds)

$MOPT$ is the mean one-way propagation time (in milliseconds), and is independent of end

ED^A, ED^B are the INMD echo-path delay inputs to the model (ms)

The opinion score predictions, y_x, use the empirical formulae described in Equation A-4. Firstly y_0 is calculated by setting $x = 0$, then y_{MOPT} is calculated, setting $x = MOPT$.

$$
\begin{aligned}
ln\left(\frac{y_x^B}{4-y_x^B}\right) = a & \\
& + b \times OLR_{EST}^B \\
& + c \times STMR^B \\
& + d \times OLR_{EST}^B \times STMR^B \\
& + e \times \left(10^{\frac{-OELR^B}{10}} + 10^{\frac{-TCL-OLR_{EST}^A-OLR_{EST}^B}{10}}\right)^i \times (1+x)^j \\
& + f \times \left(10^{\frac{-STMR^B}{10}}\right)^i \times (1+x)^j \\
& + g \times \left(10^{\frac{-OELR^B}{10}} + 10^{\frac{-TCL-OLR_{EST}^A-OLR_{EST}^B}{10}}\right)^i \times (1+x)^{2j}
\end{aligned}
\tag{A-4}
$$

where:

OLR_{EST}^A, OLR_{EST}^B are the estimated overall loudness ratings of the connection, in each direction of transmission

$STMR^B$ is the sidetone masking rating at end B

$OELR$ is the overall echo loudness rating (of the talker)

TCL is the terminal coupling loss. The recommended value is given in Table A.3.

x is the mean one-way propagation time.

The coefficient values (a … g, i, j and TCL) are given in Table A.3.

Table A.3/P.562 – Echo and delay formula coefficient values

Coefficient values
a = 2.033147
b = -0.098411
c = -0.025504
d = 0.002805
e = 12.038429
f = 0.938353
g = -12.970093
i = 0.195657
j = 0.078805
TCL = 37.0

A.2.1.1 Overall loudness rating, OLR_{REF} and OLR_{EST}

Two versions of OLR are used in the CCI. The first is the OLR of the reference connection (OLR_{REF}), and the second is an estimated value (OLR_{EST}) based on the received speech level at the INMD. The basic OLR equations are based on those in ITU-T Rec. P.79 [14].

$$OLR_{REF}^{A} = SLR_{REF}^{A} + RLR_{REF}^{B} - \frac{10}{m}\sum_{i=4}^{17} 10^{\frac{m}{10}\left(S_{MJi}^{A} + S_{Jei}^{B} - L_{Ei}^{B} - W_{oi}\right)} \quad \text{(A-5)}$$

$$OLR_{EST}^{A} = OLR_{REF}^{A} + SF^{A}.\left(SL_{REF}^{A} - SL^{A}\right) \quad \text{(A-6)}$$

where:

m is a constant and is equal to 0.175

SLR_{REF}, RLR_{REF} are the send and receive loudness ratings for the reference connection

S_{MJ} is the sending sensitivity – mouth-to-junction (for 0 dB SLR)

S_{Je} is the receiving sensitivity – junction-to-ear (for 0 dB RLR)

L_E is the earcap leakage correction factor

W_o is the OLR weighting.

SL_{REF}^{A} is the expected speech level at the INMD calculated from the reference connection and assumptions and is shown in Equation A-7

SF^{A} is a scaling factor between zero and one. This is included to apportion any difference between the reference speech level and the INMD measured speech level. The deviation of the INMD measured speech level from the reference speech level is due to two factors: the speaker's vocal level and the sending loss of the network. If the speaker's vocal level was constant then all variations in INMD measured speech level would be due to the sending loss of the network. In this case, SF would be set to equal one, to reflect this fact and correct the estimated OLR. In reality, SF depends on the distribution of network losses together with the distribution of vocal levels.

A.2.1.2 Reference speech level, SL_{REF}

SL_{REF} is the expected speech level at the INMD, calculated for the reference connection.

$$SL_{REF}^{A} = 10\log_{10}\sum_{i=4}^{17} 10^{0.1\left(\beta'^{A}_{S_i} + S_{MJ_i}^{A} - SLR_{REF}^{A} + 10\log_{10}[\Delta_i]\right)} \quad \text{(A-7)}$$

where:

β'_S is the level of spectrum density of speech emitted by the talker

S_{MJ} is the mouth-to-junction sensitivity of the reference condition at 0 dB SLR

SLR_{REF} is the SLR of the reference connection

Δ is the width of the frequency band

A.2.1.3 Sidetone masking rating, STMR

The STMR is the sidetone masking rating for the reference connection

$$STMR^B = -\frac{10}{m}\log_{10}\sum_{i=1}^{20} 10^{\frac{m}{10}\left(-L^B_{MeSTi}-L^B_{Ei}-W_{MLi}\right)} \qquad \text{(A-8)}$$

where:

m is a constant and is equal to 0.225

L_{MeST} is the mouth-to-ear loss via the sidetone path

L_E is the earcap leakage correction factor

W_{ML} is the STMR weighting for an unsealed ear

A.2.1.4 Overall echo loudness rating (OELR)

$$OELR^B = SLR^B_{EXP} + EL^B + RLR^B_{REF} \qquad \text{(A-9)}$$

where:

SLR_{EXP} is a predicted SLR value based on the received speech level at the INMD

EL is the INMD measured echo-path loss

RLR_{REF} is the RLR of the reference connection

A.2.1.5 Predicted sending loudness rating, SLR_{EXP}

$$SLR^B_{EXP} = SLR^B_{REF} + SF^B.\left(SL^B_{REF} - SL^B\right) \qquad \text{(A-10)}$$

where:

SLR_{REF} is the SLR of the reference connection

SF^B is a scaling factor between zero and one. Described after Equation A-6

A.2.2 Pre-echo opinion score, $Y_{Cpre\text{-}echo}$

The pre-echo conversational opinion score prediction is derived from an intermediate opinion score, Y_{Cint} as shown in Equation A-11.

$$\ln\left(\frac{Y^B_{Cpre\text{-}echo}}{4-Y^B_{Cpre\text{-}echo}}\right) = 0.8541.\ln\left(\frac{Y^B_{C\text{int}}}{4-Y^B_{C\text{int}}}\right) - 0.2727 \qquad \text{(A-11)}$$

A.2.2.1 Intermediate opinion score, Y_{Cint}

Determination of Y_{Cint}, the intermediate opinion score.

$$\ln\left(\frac{Y^B_{C\text{int}}}{4-Y^B_{C\text{int}}}\right) = 0.7\left[\ln\left(\frac{Y^B_{LE}}{4-Y^B_{LE}}\right) + 0.5 - \frac{K\times\left(13-STMR^B\right)}{20}\left(\frac{4-Y^B_{LE}}{Y^B_{LE}}\right)^2\right] \qquad \text{(A-12)}$$

where:

STMR is calculated in Equation A-8.

K is equal to one if STMR < 13 dB and equal to zero otherwise

Y_{LE} is the listening effort opinion score

A.2.2.2 Listening effort opinion score, Y_{LE}

Listening effort is calculated using the following formula.

$$\ln\left(\frac{Y_{LE}^{B}}{4-Y_{LE}^{B}}\right)=1.465\left[\ln\left(\frac{LOI^{B}}{LOI_{LIM}-LOI^{B}}\right)-0.75\right] \quad \text{(A-13)}$$

where:

LOI is the listening opinion index

LOI_{LIM} is usually set to 0.885, to account for the 'enhancement effect'. This is the effect whereby subjects tend to allocate the maximum available score to a condition which is less than ideal but which is equal to the best commonly experienced. Thus the maximum Y_{LE} value is attributed to a connection for which $LOI = LOI_{LIM}$.

A.2.2.3 Listening Opinion Index, LOI

The listening opinion index (LOI) is calculated using Equations A-14 to A-28.

$$LOI^{B} = A^{B}D^{B}\sum_{i=4}^{17} B_i^{'}\Delta_i P\left(Z_i^{B}\right) \quad \text{(A-14)}$$

where:

A is the A-Factor used to compensate for deviation from optimum listening level

D is the D-Factor used to compensate for received circuit noise

$B_i^{'}$ is the LOI frequency weighting for the ith frequency band

Δ_i is the width of the ith frequency band

$P(\,)$ is a growth function

Z_i is the effective sensation level for the ith band

The listening opinion index is calculated narrow-band over bands 4 to 17.

A.2.2.4 A-Factor

The A-Factor is a multiplier depending on the received speech level, with the value 1 for a small range of levels around the optimum but decreasing rapidly outside this range. The optimum received speech level corresponds to an optimum OLR of 8 dB.

$$A^{B} = 1.00125 - 0.0005556\left\{\Delta RSL^{B} - 1.50\right\}^{2} \quad \text{when} \quad \Delta RSL < 0 \quad \text{(A-15)}$$

$$A^{B} = 1.01005 - 0.0002571\left\{\Delta RSL^{B} - 1.56\right\}^{2} \quad \text{when} \quad \Delta RSL \geq 0 \quad \text{(A-16)}$$

where:

$$\Delta RSL^{B} = SL_{REF}^{A} - SL^{A} + OLR_{REF}^{A} - 8 \quad \text{(A-17)}$$

and:

SL^A_{REF} is the expected speech level at the INMD calculated from the reference connection and assumptions and is shown in Equation A-7

SL is the INMD measured speech level

OLR_{REF} is the OLR of the reference connection

The following limits should also be applied to the A-Factor:

- if A > 1 then A = 1.
- if A < 0.001 then A = 0.001.

The value '8' in Equation A-17 represents the currently accepted optimum value of OLR. The OLR value in question is the one for the path through which the subject is listening (not speaking).

A.2.2.5 D-Factor

D is a multiplier depending on the received total noise level with a value decreasing slowly from 1 at negligible noise levels towards 0 at very high noise levels. The total noise level is dependent on the noise levels measured in both directions of transmission by the INMD.

$$D^B = \frac{1}{1+\exp\left(0.1125\cdot(TNL^B - RLR^B_{P.XXE}) + 2.25\right)} \tag{A-18}$$

$$TNL^B = 10\log_{10}\left(10^{\left(N^A - R^B\right)/10}\right) + R^B \tag{A-19}$$

where:

RLR_{CCI} is the receive loudness rating according to ITU-T Rec. P.79, calculated with weighting factors shown in Table A.4

R, T are the send and receive exchange pad loss settings in dB

TNL is the total noise level reaching the listener's ear referred back to the INMD (or 0 dBr point). This assumes that the noise reaching the listener's ear is the product of noise emanating from the local networks at each end and so is a function of both INMD noise level measurements (in both directions).

A.2.2.6 RLR_{CCI}

RLR calculated with slightly different weighting values to ITU-T Rec. P.79.

$$RLR^B_{P.CCI} = \frac{-10}{m}\log_{10}\sum_{i=4}^{17} 10^{\frac{m}{10}\left(S^B_{Jei} - L^B_{Ei} - W_{rp\cdot xxei}\right)} \tag{A-20}$$

where:

m is a constant and is equal to 0.225

S_{Je} is the junction-to-ear sensitivity

L_E is the earcap leakage correction factor

W_r are weighing values given in Table A.4

A.2.2.7 Growth function

P( ) is a growth function which derives LOI (a quantity related to listening effort) from the sensation level Z.

$$P(Z) = 10^{\frac{Z+3.8}{10}} \quad \text{when} \quad Z < -11 \tag{A-21}$$

$$P(Z) = 1 - 10^{\frac{-0.3(Z+14)}{10}} \quad \text{when} \quad Z \geq -11 \tag{A-22}$$

A.2.2.8 Effective sensation level

The effective sensation level, Z, is a frequency-dependent quantity and is calculated for each frequency band. The formulae given to calculate Z are applied to each frequency band.

The effective sensation level Z, is the difference in dB between the speech and total noise spectra reaching the listener's ear.

$$Z_i^B = Z_{ROi}^B - M_i^B - \left(L_{MEi}^A - SLR_{REF}^A - RLR_{REF}^B\right) + \left(SL^A - SL_{EXP}^A\right) \tag{A-23}$$

where:

Z_{RO} is the sensation level (in dB above threshold of hearing) for 0 dB mouth-to-ear loss at a given frequency in the absence of noise

L_{ME} is the mouth-to-ear loss of the connection for 0 dB OLR

M is a threshold shift to account for the fact that in the presence of noise, the loudness of the sound heard is altered

$$Z_{ROi}^B = \beta'^A_{Si} - \left(\beta_{0i}^B - 10\log_{10}[\Delta_i]\right) \tag{A-24}$$

where for a given frequency:

β'_S is the level of the spectrum density of the speech emitted by the talker in dB Pa^2/Hz. Note that the values currently used are based on the arithmetic mean of the male and female spectra

β_0 is the pure-tone threshold of audibility of the standard listener in dB Pa^2

Δ is the width of the frequency band centred at a given frequency in Hz

$$L_{MEi}^A = -\left(S_{MJi}^A + S_{Jei}^B\right) + L_{Ei}^B \tag{A-25}$$

The threshold Shift M, at any given frequency is given by:

$$M_i^B = 10\log_{10}\left[10^{\frac{\beta_{CNi}^B}{10}} + 10^{\frac{\beta_{RNi}^B}{10}} + 10^{\frac{\beta_{0i}^B - 10\log_{10}[\Delta_i]}{10}}\right] - \beta_{0i}^B - 10\log_{10}[\Delta_i] \tag{A-26}$$

where:

β_{CN} is the spectrum density of circuit noise from all sources combined (including room noise from the far end) at the ear reference point in dB Pa^2/Hz

β_{RN} is the spectrum density of room noise at the ear reference point after transmission through the sidetone path, combined with that arriving by leakage past the earcap of the earphone, in dB Pa^2/Hz

The spectrum density of circuit noise at the ear reference point is given by adding the level of circuit noise, ICN, to the noise spectrum, ICNV0 (which is measured with psophometric weighting), taking into account the loss from junction to the ear. This is shown in Equation A-27.

$$\beta_{CN}^B = CNS_i^A + TNL^B + S_{Jei}^B - L_{Ei}^B - RLR_{REF}^B \tag{A-27}$$

where:

CNS is the spectrum density of circuit noise at a level of 0 dBm0p

TNL is the total level of circuit noise referred to the INMD as given in Equation A-19

S_{Je} is the sensitivity from junction to ear

L_E is the ear leakage correction factor

RLR_{REF} is the RLR of the reference connection

The spectrum density of room noise at e.r.p. is due to two components: the spectrum of noise reaching the ear through the sidetone path of the telephone terminal; and the spectrum of noise reaching the ear through the air gap between the telephone ear piece and the ear.

$$\beta'^{B}_{RNi} = 10\log_{10}\left(10^{\frac{\left(RNS_i^B + RN^B - L^B_{RNEi}\right)}{10}} + 10^{\frac{RNS_i^B + RN^B + \Delta^B_{SMi} - L^B_{MeSTi} - L^B_{Ei}}{10}}\right) \qquad \text{(A-28)}$$

where:

L_{MeST} is the loss from mouth reference point to ear reference point for transmission via the sidetone path

Δ_{SM} is the excess microphone sensitivity to room noise

L_{RNE} is the air-to-air transmission loss of earcap leakage path for room noise to e.r.p.

RN is the room noise level in dBA

RNS_i is the spectrum density of Hoth room noise at 0 dBA

L_E is the artificial to real ear correction factor

A.3 Model output

The output from the call clarity index model is a prediction of the conversational mean opinion score, Y_c, for each end of the connection. This score gives a measure of the speech quality of the connection for a conversational task.

The prediction of Y_c is given as a floating point number from 1 to 5 based on the five-point speech quality scale in ITU-T Rec. P.800 [6] and shown below:

5 Excellent

4 Good

3 Fair

2 Poor

1 Bad

A.4 Weighting values

A.4.1 Loudness ratings weightings

See Table A.4.

Table A.4/P.562 – Loudness ratings weighting values

Band	Frequency Hz	W_s	W_r	W_o	W_{ML}	W_{rcci}
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
1	100				94.0	
2	125				91.0	
3	160				90.1	
4	200	76.9	85.0	65.8	86.4	90.8
5	250	62.6	74.7	60.4	81.9	79.0
6	315	62.0	79.0	68.2	78.5	71.2
7	400	44.7	63.7	55.3	78.2	64.3
8	500	53.1	73.5	66.6	72.8	58.0
9	630	48.5	69.1	63.0	67.6	56.9
10	800	47.6	68.0	63.1	58.4	56.1
11	1000	50.1	68.7	65.0	49.7	57.6
12	1250	59.1	75.1	72.8	48.0	57.2
13	1600	56.7	70.4	69.8	48.7	56.2
14	2000	72.2	81.4	81.7	50.7	58.0
15	2500	72.6	76.5	78.3	49.8	56.8
16	3150	89.2	93.3	95.1	48.4	58.3
17	4000	117.0	113.8	76.6	49.2	103.5
18	5000				47.7	
19	6300				48.0	
20	8000				50.7	

A.4.2 LOI weightings

See Table A.5.

Table A.5/P.562 – LOI weighting values

Band	Frequency Hz	Δ_i Hz	B'_I	$B'_i \Delta_I$ (3) × (4)
(1)	*(2)*	*(3)*	*(4)*	*(5)*
1	100			
2	125			
3	160			
4	200	46.0	3.694E-04	0.016994
5	250	58.0	4.097E-04	0.023763
6	315	73.0	4.872E-04	0.035565
7	400	92.0	5.139E-04	0.047282
8	500	115.0	5.084E-04	0.058466
9	630	146.0	4.999E-04	0.072979
10	800	183.0	4.491E-04	0.082186
11	1000	229.0	3.894E-04	0.089166
12	1250	290.0	3.459E-04	0.100300
13	1600	370.0	2.728E-04	0.100925
14	2000	460.0	1.809E-04	0.083234
15	2500	580.0	1.128E-04	0.065449
16	3150	730.0	6.725E-05	0.049091
17	4000	920.0	3.989E-05	0.036703
18	5000			
19	6300			
20	8000			

A.5 Assumptions/Data files

This clause contains the assumption data values used by the CCI.

NOTE – Where values describe/refer to the network or equipment (such as telephones) they are based on data obtained from different European countries.

A.5.1 Speaker and listener data values

See Table A.6.

Table A.6/P.562 – Frequency dependent speaker and listener assumption data values

Band	Frequency Hz	β'_S dB Pa^2/Hz	β_0 dB Pa^2/Hz $\times 10^{-10}$
(1)	*(2)*	*(3)*	*(4)*
1	100	–38.0	223.8000
2	125	–32.9	109.6000
3	160	–31.5	39.8100
4	200	–29.7	12.59
5	250	–30.0	7.079
6	315	–33.3	3.631
7	400	–34.2	1.995
8	500	–34.6	1.259
9	630	–37.7	0.9333
10	800	–41.6	0.6310
11	1000	–46.4	0.5012
12	1250	–48.8	0.5623
13	1600	–50.0	0.6310
14	2000	–52.6	0.5012
15	2500	–55.2	0.2818
16	3150	–59.3	0.1660
17	4000	–63.0	0.1995
18	5000	–66.2	0.2239
19	6300	–67.9	0.3090
20	8000	–68.5	0.5012

A.5.2 Telephone data values

See Table A.7.

Table A.7/P.562 – Frequency dependent telephone assumption data values

Band	Frequency Hz	S_{MJ} dB V/Pa	S_{Je} dB Pa/V	Δ_{SM} dB	L_E dB	L_{MeST} dB	L_{RNE} dB Pa^2/Hz
(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
1	100				20.0	47.76	
2	125				16.5	35.46	
3	160				12.5	46.66	
4	200	–12.55	0.99	–6.6	8.4	0.06	3.6
5	250	–10.22	5.72	–6.8	4.9	–1.33	4.9
6	315	–8.48	9.42	–6.9	1.0	–0.18	6.6
7	400	–7.10	11.33	–6.6	–0.7	0.49	8.7
8	500	–6.03	11.49	–4.8	–2.2	2.27	10.6
9	630	–5.04	11.84	–6.6	–2.6	2.68	13.1
10	800	–4.49	11.99	–8.4	–3.2	2.33	16.8
11	1000	–4.22	12.22	–8.9	–2.3	0.79	20.2
12	1250	–3.57	12.04	–11.0	–1.2	1.16	23.1
13	1600	–3.16	11.19	–13.3	–0.1	7.27	24.4
14	2000	–3.16	9.63	–14.7	3.6	10.70	23.3
15	2500	–2.41	10.08	–10.8	7.4	4.90	20.6
16	3150	–5.45	11.35	–12.9	6.7	4.39	18.8
17	4000	–38.69	–28.39	–12.1	8.8	24.82	18.4
18	5000				10.0	66.26	
19	6300				12.5	60.96	
20	8000				15	73.26	

S_{MJ} and S_{je} defined for 0 dB SLR and RLR.

A.5.3 Noise data values

See Table A.8.

Table A.8/P.562 – Assumed frequency spectra for room and circuit noise

Band	Frequency Hz	RNS dB Pa2/Hz	CNS dB V/√Hz
(1)	*(2)*	*(3)*	*(4)*
1	100		
2	125		
3	160		
4	200	–116.4	–42.3
5	250	–118.0	–39.7
6	315	–119.6	–36.9
7	400	–121.3	–35.7
8	500	–122.9	–35.1
9	630	–124.5	–34.8
10	800	–126.2	–34.6
11	1000	–127.8	–34.6
12	1250	–129.4	–34.7
13	1600	–131.1	–34.7
14	2000	–132.7	–34.6
15	2500	–134.4	–34.1
16	3150	–136.2	–37.5
17	4000	–138.6	–85.0
18	5000		
19	6300		
20	8000		

For room noise level of 0 dBA and circuit noise level of 0 dBm0p.

Table A.9/P.562

Quantity	Value	Unit
RN	40	dBA
LC	1.0	ms
SF	0.5	–
SLR_{REF}	8	dB
RLR_{REF}	3	dB
T	2.5	dB
R	7.5	dB

A.6 Miscellaneous equations

In various places in the CCI the following functions are used to map points between a boundary-limited and a boundary-unlimited domain. The boundary limits used within the model are zero and four. The equations to transform from a boundary limited domain to a boundary unlimited domain, and back are shown below.

Boundary-limited to boundary-unlimited transformation:

$$y = \ln\left(\frac{x}{4-x}\right) \tag{A-29}$$

Where x is a boundary limited value and y is boundary unlimited.

The inverse of this function is a boundary-unlimited to boundary-limited transformation:

$$x = \frac{4}{1+\exp(-y)} \tag{A-30}$$

Where x is a boundary limited value and y is boundary unlimited.

Annex B

Mapping INMD measurements to the E-model

B.1 Algorithms relating INMD measurements to E-model parameters

INMDs measure the speech level (SL), noise level (NL), echo loss (EL), and echo path delay (EPD) of both directions of a connection. INMDs of class D can also measure the IP packet loss ratio (IPLR) and delay variation (IPDV) in both directions. These parameters can be used to derive the impact of loss, noise, delay, packet loss, and echo on customer opinion. Because INMDs do not make end-to-end performance measurements, however, it should be noted that using INMD measurements in a customer opinion model like the E-model, requires estimating some parameters which cannot be derived from the INMD's measurements. In particular, referring to Figure 7, the INMD's far SL measurement (SL_f) can be used to derive the combination of the far SLR (SLR_F) and the transmit loss in the far-to-near direction (T_F). However, the near RLR (RLR_N) and the receive loss in the far-to-near direction (R_N) cannot be derived from the INMD's measurements because they affect the connection's performance after the point at which the INMD makes measurements. These parameters have to be estimated by the user and typically are selected to represent an average value or distribution of values for the network being evaluated.

The following three subclauses provide algorithms mapping the INMD's measurements to the loss, noise, delay, packet loss, and echo parameters used by the E-model [2]. Clause B.1.6 provides a full set of equations for using the E-model to evaluate loss, noise, delay, packet loss, and echo performance with INMD measurements.

B.1.1 Algorithms relating INMD measurements to the E-model's loss parameters

The E-model uses the Overall Loudness Rating (OLR) to evaluate the acoustic to acoustic loss of a connection. The OLR is the sum of the SLR and RLR of the loops and telephones in a connection, plus any additional losses in the network. In most descriptions of the E-model, any transmit or receive losses (T_N, T_F, R_N, R_F) are included in either the SLR or the RLR of the connection.

An algorithm was developed relating the INMD's SL measurement to TOLR. Using the relationship: SLR = TOLR + 56, a new algorithm relating the INMD's SL measurement to SLR is created:

$$SRL = -18.6 - 0.962 \times SL \qquad \text{(B-1)}$$

In terms of the connection given in Figure 7, the SLR includes both the SLR of the loop and telephone (SLR_F) and the transmit loss in the 4-wire portion of the network (T_F). In most networks, this transmit loss is set to 0 which was the case during the testing which produced the algorithm in Equation B-1.

The algorithm given in Equation (B-1) was developed using a least squares regression analysis on the SL and SLR (converted from TOLR) measurements. The accuracy of this regression was very good. The square of the multiple correlation coefficient (R^2) for the regression was 0.876. R^2 is a measure of the goodness of fit of the regression. An R^2 of 1.0 would indicate the regression explained the variation between the SL and the SLR perfectly. The regression had errors ranging between –3.4 and +5.5 dB. However, an analysis of the accuracy of the regression indicated that about 50% of the SLRs predicted from the INMD's speech levels were within 1 dB of the actual SLRs and more than 75% were within 2 dB of the actual SLRs.

Because the average SL and average SLR simulated in the test differ from those of most actual networks, Equation B-1 must be modified for use with INMD measurements made on actual networks. In the laboratory test, the average SL simulated was –33.8 dBm and the average SLR was 13.7 dB. Administrations using the algorithm in Equation B-1 to derive customer opinion ratings

from INMD measurements should modify the equation by adding a constant value, C, which is given by:

$$C = 0.962 \times (SL_{AVG} + 33.8) + (SLR_{AVG} - 13.7) \quad \text{(B-2)}$$

Where, SL_{AVG} and SLR_{AVG} are the average SL and SLR in the network being evaluated. Note that SLR_{AVG} must include the network loss T_{AVG}.

The other two parameters included in the OLR, the receive loss and the RLR, cannot be estimated from the INMD's measurements. They both impact the speech level after it has been measured by the INMD. Based on this, the following equation can be used to estimate the OLR of a connection from an INMD's speech level measurement:

$$OLR = -18.6 + C - 0.962 \times SL + RLR_{AVG} + R_{AVG} \quad \text{(B-3)}$$

where RLR_{AVG} and R_{AVG} are the average RLR and receive loss for the network being evaluated.

B.1.2 Algorithms relating INMD measurements to the E-model's noise parameters

As Figure 7 shows, the INMD measured the near and far noise levels in the 4-wire portion of the test circuit. Because the test simulated no losses in the 4-wire portion of the connection, the noise measured at the 2-wire side of the 2-wire to 4-wire interface should be equal to the power sum of the near and far INMD noise measurements. This provides the following equation:

$$N = 10 \times \log_{10}\left(10^{\frac{NL_n}{10}} + 10^{\frac{NL_f}{10}}\right) \quad \text{(B-4)}$$

where, N is the Psophometric weighted circuit noise measured at the network side of the near loop and NL_n and NL_f are the near-to-far and far-to-near INMD noise level measurements respectively. An analysis of the error made by Equation A-4 showed that more than 90% of the predicted noise values were within 1 dB of the actual noise. In addition, the mean of the predicted noise values was within 0.2 dB of the mean of the actual measured noise values.

The noise values given by Equation B-4 are the electrical noise measured at the network side of the near loop. These values are converted to acoustic noise values within the E-model. It should be noted that in actual networks, if there is any transmit or receive loss in the 4-wire portion of the connection being evaluated, the INMD noise measurements must be corrected as shown in the equation below:

$$N = 10 \times \log_{10}\left(10^{\frac{NL_n + T_{AVG}}{10}} + 10^{\frac{NL_f - R_{AVG}}{10}}\right) \quad \text{(B-5)}$$

Where, T_{AVG} and R_{AVG} are the average transmit and receive losses for the network being evaluated which cannot be derived from the INMD's measurements. It should be noted that Equation B-5 does not include the impact of noise inserted at the four-wire near end (such as quantization noise, etc.). This noise cannot be measured by the INMD.

B.1.3 Algorithms relating INMD measurements to the E-model's absolute delay parameters

Because the echo path delay parameter used by the E-model (we will call it loop delay, or LD) is equal to the time it takes a transmitted signal to return to the caller, the sum of the INMD's two EPD measurements provides an excellent estimate of LD. It is only missing the amount of time it takes the signal to travel across the near end evaluator's loop twice. This provides the following algorithm for this parameter:

$$LD = EPD_n + EPD \quad \text{(B-6)}$$

where:

EPD_n is the INMD's EPD measurement for echoes reflected from the hybrid at the near end of the connection

EPD_f is the INMD's EPD measurement for echoes reflected from the far end of the connection. No accuracy analysis was performed on the predicted EPD values because the reference measurements and the INMD measurements were the same.

The absolute one-way delay is the sum of the following factors:

– One-way network delay: in circuit switched scenarios, it is equal to LD/2. We will assume that it is also the case for packet switched scenarios (i.e., packets follow the same path in both transmission directions), which cannot be verified with an INMD.

– Processing delay in terminals: this is restricted to the impact of codec and jitter buffer (i.e., in the case of class D INMDs only).

  - The delay attributable to codec-related processing in IP-based systems is:

    (N + 1) × frame size + look-ahead, where N is the number of frames in each packet and can be easily derived from the analysis of IP protocols. The frame size and the length of the look ahead are specific to each type of codec (which can be easily derived from the analysis of IP protocols) and can be found in Table I.4/G.114 [10].

  - The delay attributable to the de-jitter buffer (BD) will depend on the configuration of the buffer (i.e. the maximum audio delay variation that the buffer can support before discarding packets). If the de-jitter buffer is static, BD will be equal to this maximum delay, and this information can be easily derived from the analysis of IP protocols. It is more difficult in the case of a dynamic buffer, where the delay introduced by the buffer will also depend on the measurement of IP delay variation (IPDV) itself, but this cannot be measured with an INMD.

This provides the following algorithm for the absolute one-way delay parameter used by the E-model (OWD):

$$OWD = \frac{LD}{2} + (N+1)\,framesize + lookahead + BD \qquad \text{(B-7)}$$

B.1.4 Algorithms relating INMD measurements to the E-model's packet loss parameters

The overall packet loss is taken into account in the E-model by combining the raw measurement of IP packet loss ratio (IPLR) with the impact of the de-jitter-buffer in case of late packets discarded when their relative delay (reflected by the measurement of IPDV) exceeds the capacity of the buffer (BD, see B.1.3). This provides the following formula:

$$Ppl = IPLR + 100 \cdot \left(1 - \frac{BD}{IPDV}\right) \qquad \text{(B-8)}$$

B.1.5 Algorithms relating INMD measurements to the E-model's echo parameters

The E-model uses the TELR as its parameter for the acoustic to acoustic echo loss. For the caller at the near end of the connection in Figure 7, the TELR experienced by the caller at the near end is equal to the sum of SLR_N, EL_F, and RLR_N. Where the intrusive echo loss measurement, EL_F, is made at the two- to four-wire interface and includes all of the losses in the four-wire portion of the connection (T_N, R_F, T_F, and R_N). As Figure 7 shows, the INMD's EL_f measurement only includes the sum of the transhybrid loss of the far two- to four-wire hybrid and R_F and T_F. The INMD's SL_n measurement can be used to derive the sum of SLR_N and T_N using Equation B-1 modified by Equation B-2. Because the INMD's measurements cannot be used to derive R_N or RLR_N, they

should be set equal to averages for the network being evaluated. Equation B-9 provides an algorithm for determining the TELR:

$$TELR = -18.6 + C - 0.962 \times SL_n + EL_f + R_{AVG} + RLR_{AVG} \quad \text{(B-9)}$$

Because the reference and INMD measurements for the EL were the same in the laboratory test, no accuracy analysis was performed on the algorithm given in Equation B-9.

B.1.6 Equations for the E-model using INMD measurements

The E-model generates a rating R for a connection which can be translated into a customer opinion rating. The E-model for loss, noise, and talker echo uses five terms to generate R:

$$R = R_o - I_{OLR} - I_{DD} - I_{e\text{-}eff} - I_{DTE} \quad \text{(B-10)}$$

where:

R_O is the signal-to-noise ratio at a 0 dB reference point. In the equations provided here, the 0 dB reference point is at the 2-wire input to the telephone receiving system at the near end of the connection

I_{OLR} is the impairment term for the overall loudness rating

I_{DD} is the impairment term for the absolute one-way delay

$I_{e\text{-}eff}$ is the impairment term for the low bit-rate coding under random packet loss conditions

I_{DTE} is the impairment term for the delayed talker echo

Using Equation B-1 modified by Equation B-2 and Equation B-5, R_O can be derived from the INMD's SL and NL measurements:

$$R_O = 15 - 1.5 \times \left(-18.6 + C - 0.962 \times SL_f + R_{AVG} + N_O\right) \quad \text{(B-11)}$$

where:

C is given by Equation B-2

R_{AVG} is the average receive loss in the network being evaluated and is included because the 0 dBr point used here is at the 2-wire input to the telephone receiving system

N_O is the overall noise which is given by:

$$N_O = 10 \times \log_{10}\left(10^{\frac{N}{10}} + 10^{\frac{N_F + RLR_{AVG}}{10}}\right) \quad \text{(B-12)}$$

where:

N is given by Equation B-5 and includes T_{AVG} and R_{AVG} which are the average transmit and receive losses in the network being evaluated

N_F is a noise floor which represents the impact of the room noise at the listening party's location. This noise floor is referred to an RLR of 0 dB which is corrected for the actual average RLR of the network being evaluated in Equation B-12. N_F is usually assumed to be equal to –64 dBmp.

The impairment for the Overall Loudness Rating, I_{OLR}, is:

$$I_{OLR} = 20 \times \left[\left(1 + \left(\frac{X}{8} \right)^8 \right)^{\frac{1}{8}} - \frac{X}{8} \right] \tag{B-13}$$

where the variable X is given by:

$$X = OLR + 0.2 \times (64 + N_t) \tag{B-14}$$

which using Equation B-3 can be rewritten as:

$$X = -18.6 + C - 0.962 \times SL_F + RLR_{AVG} + R_{AVG} + 0.2 \times (64 + N_t) \tag{B-15}$$

where N_t is given by:

$$N_t = N_O - RLR_{AVG} \tag{B-16}$$

The impairment for the Absolute Delayed, I_{DD}, is:

For OWD < 100 ms:

$$I_{DD} = 0$$

For OWD > 100 ms:

$$I_{DD} = 25 \left\{ \left(1 + X^6 \right)^{\frac{1}{6}} - 3 \left(1 + \left[\frac{X}{3} \right]^6 \right)^{\frac{1}{6}} + 2 \right\} \tag{B-17}$$

where:

$$X = \frac{\log\left(\frac{OWD}{100} \right)}{\log 2} \tag{B-18}$$

where:

OWD is given in Equation B-7.

The impairment for the low bit-rate coding under random packet loss conditions, $I_{e\text{-}eff}$, is:

$$I_{e\text{-}eff} = I_e + (95 - I_e) \cdot \frac{Ppl}{Ppl + Bpl} \tag{B-19}$$

where:

P_{pl} is given in Equation B-8

I_e is the codec specific value for the equipment impairment factor at zero packet-loss

B_{pl} is the Packet-loss Robustness Factor of the codec

I_e and *Bpl* values are both listed in Appendix I/G.113 [9] for several codecs.

The impairment for the Delayed Talker Echo, I_{DTE}, is:

$$I_{DTE} = \left\{ \left(\frac{R_{OE} - R_E}{2} \right) + \left[\sqrt{\frac{(R_{OE} - R_E)^2}{4} + 100} - 1 \right] \right\} \times \left(1 - e^{-\frac{LD}{2}} \right) \tag{B-20}$$

where:

$$R_{OE} = -1.5 \times N_t \tag{B-21}$$

$$R_E = 80 + 2.5 \times (TERV - 14) \tag{B-22}$$

where:

$$TERV = TELR - 40 \times \log_{10}\left(\frac{1 + \frac{LD}{20}}{1 + \frac{LD}{300}}\right) + 6 \times e^{-0.3 \times \left(\frac{LD}{2}\right)^2} \tag{B-23}$$

where:

TELR is given in Equation B-9

LD is given in Equation B-6

Appendix I

Details on the comparison of class C INMD location within the network

Tables I.1 and I.2 show if the measurements made by an INMD of class C help to evaluate the perceived quality for each measurement parameter. These are shown for each link, and in relation to the near end customer, the far end customer and the interconnection. If a measurement provides useful information as to the perceived quality, a '1' is used and, where possible network equipment effects could not be measured, a '0' is used. Scenario 1 refers to an INMD located on the outgoing side of the international gateway and scenario 2 refers to the incoming side. The link descriptions are:

- Link 1 – with echo cancellers and CME at both sides;
- Link 2 – with echo cancellers and without CME at both sides;
- Link 3 – without echo cancellers and without CME at both sides.

Table I.1/P.562 – Comparison of INMD location scenarios for the case when ALC is enabled

With ALC			Near end customer	Far end customer	Inter-connection	Total
Active speech level	Sc. 1	Link 1	0	0	1	1
		Link 2	0	0	1	1
		Link 3	0	0	1	1
	Sc. 2	Link 1	1	0	0	1
		Link 2	1	0	0	1
		Link 3	1	0	0	1
Noise level	Sc. 1	Link 1	0	0	1	1
		Link 2	0	0	1	1
		Link 3	1	1	1	3
	Sc. 2	Link 1	1	0	0	1
		Link 2	1	0	0	1
		Link 3	1	1	1	3
Echo delay	Sc. 1	Link 1	1	0	0	1
		Link 2	1	1	1	3
		Link 3	1	1	1	3
	Sc. 2	Link 1	1	0	0	1
		Link 2	1	1	1	3
		Link 3	1	1	1	3
Echo loss	Sc. 1	Link 1	0	0	1	1
		Link 2	0	0	1	1
		Link 3	0	0	1	1
	Sc. 2	Link 1	1	0	0	1
		Link 2	1	0	0	1
		Link 3	1	0	0	1
TOTAL	Sc. 1	Link 1	1	0	3	4
		Link 2	1	1	4	6
		Link 3	2	2	4	8
		Total	**4**	**3**	**11**	**18**
	Sc. 2	Link 1	4	0	0	4
		Link 2	4	1	1	6
		Link 3	4	2	2	8
		Total	**12**	**3**	**3**	**18**

Table I.2/P.562 – Comparison of INMD location scenarios for the case when ALC is disabled

Without ALC			Near end customer	Far end customer	Inter-connection	Total
Active speech level	Sc. 1	Link 1	1	1	1	3
		Link 2	1	1	1	3
		Link 3	1	1	1	3
	Sc. 2	Link 1	1	1	1	3
		Link 2	1	1	1	3
		Link 3	1	1	1	3
Noise level	Sc. 1	Link 1	1	0	1	2
		Link 2	1	1	1	3
		Link 3	1	1	1	3
	Sc. 2	Link 1	1	0	0	1
		Link 2	1	0	0	1
		Link 3	1	1	1	3
Echo delay	Sc. 1	Link 1	1	0	0	1
		Link 2	1	1	1	3
		Link 3	1	1	1	3
	Sc. 2	Link 1	1	0	0	1
		Link 2	1	0	1	2
		Link 3	1	1	1	3
Echo loss	Sc. 1	Link 1	1	1	1	3
		Link 2	1	1	1	3
		Link 3	1	1	1	3
	Sc. 2	Link 1	1	0	0	1
		Link 2	1	0	0	1
		Link 3	1	1	1	3
TOTAL	Sc. 1	Link 1	4	2	3	9
		Link 2	4	4	4	12
		Link 3	4	4	4	12
		Total	**12**	**10**	**11**	**33**
	Sc. 2	Link 1	4	1	1	6
		Link 2	4	1	2	7
		Link 3	4	4	4	12
		Total	**12**	**6**	**7**	**25**

Appendix II

Statistical techniques for use on multiple INMD measurements

The following three statistical techniques can be used for evaluating multiple INMD measurements.

II.1 Confidence intervals

If it is assumed that a set of INMD measurements is from a normal distribution, then the following provides an α percent confidence interval for the mean of the measurements:

$$P\left[\overline{X} - \frac{\sigma}{\sqrt{n}}(Z_\alpha) < \mu < \overline{X} + \frac{\sigma}{\sqrt{n}}(Z_\alpha)\right] = \alpha$$

where:

$\overline{X}$ is the mean of the measurements

σ is the standard deviation of the measurements

n is number of measurements

Z_α is the normal deviate for α

μ is the true mean of the distribution

If the estimate for the standard deviation of the distribution is known, then the above equation can be used to determine the sample size needed to generate a confidence interval of a particular size:

$$n = \frac{\sigma^2 Z_\alpha^2}{M^2}$$

where M is the size of the confidence interval desired.

So, as an example, if it is known that the speech level measurements on a network are typically normally distributed with a standard deviation of 5 and a 95% confidence interval of ±1 dB is desired then:

$$n = \frac{5^2 \times 1.96^2}{1^2} \geq 96 \text{ measurements}$$

II.2 Hypothesis test for the mean of a set of measurements versus a fixed value

A second useful statistical technique that can be used is to develop a hypothesis test determining whether the mean of a set of measurements is greater than a fixed value. This fixed value could be a threshold used to determine when maintenance activities are required. This hypothesis is represented as:

$$H_0 : \mu \leq \mu_0 \qquad H_1 : \mu > \mu_0$$

where:

H_0 is the null hypothesis

H_1 is the alternate hypothesis

μ is the true mean

μ_0 is the threshold

If it can be assumed that the measurements are taken from a normal distribution, then the following statistic, T, can be used to determine if the null hypothesis can be rejected:

$$T = \frac{\left(\overline{X} - \mu_0\right)}{\left(\sigma / \sqrt{n}\right)}$$

where:

$\overline{X}$ is the mean of measurements

σ is the standard deviation of the measurements

n is the number of measurements

The statistic T has a Student's t distribution and can be compared to values from that distribution to determine if the null hypothesis can be rejected. In particular, if T is larger than $t_\alpha(n-1)$, then the null hypothesis can be rejected and it can be concluded that the mean of the measurements is larger than the threshold at a level of significance. The value $t_\alpha(n-1)$ is the αth deviate from the t distribution with n – 1 degrees of freedom.

This equation can also be used to determine the sample size required to assure that if a mean is a particular amount larger than a threshold then the true mean of the measurements is larger than the threshold at a given significance level. As an example, if the standard deviation of typical speech levels is 5, then from the T distribution we note that:

$$t_{.95}(60) = 1.67 \qquad t_{.95}(120) = 1.66$$

Using this and the above equation we find if we want a 1 dB difference to indicate the mean is greater than the threshold at a 95% significance level that:

$$T = \frac{\left(\overline{X} - \mu_0\right)}{\sigma / \sqrt{n}} = \frac{1}{5 / \sqrt{n}} \text{ must be greater than } 1.67$$

which implies:

$n \geq 70$ measurements for a 1 dB difference to imply the threshold is truly exceeded at a 95% confidence level.

II.3 Hypothesis test for the means of two sets of measurements

A third statistical technique useful with multiple INMD measurements is to use hypothesis tests to compare the means of two sets of measurements made over different facilities or different routes. In this case the hypothesis is:

$$H_0 : \mu_1 \leq \mu_2 \qquad H_1 : \mu_1 > \mu 2$$

If both sets of measurements are from normal distributions and the standard deviations of the two sets of measurements are equal, then the statistic used to determine if the null hypothesis can be rejected is:

$$T = \frac{\left(\overline{X_1} - \overline{X_2}\right)\sqrt{n_1 n_2 / (n_1 + n_2)}}{\sqrt{\left[(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2\right] / (n_1 + n_2 - 2)}}$$

where:

n_1 and n_2 are the sample sizes of the two sets of measurements

$\overline{X}_1$ and $\overline{X}_2$ are the means of the two sets of measurements

σ_1 and σ_2 are the standard deviations of the two sets of measurements

If T is greater than $t_\alpha(n_1 + n_2 - 2)$, then the null hypothesis can be rejected at a αth significance level and it can be concluded the mean of the first set of measurements is greater than the mean of the second set of measurements.

If we assume that the number of measurements made in the two sets and the standard deviations of the 2 sets are equal, then we can reduce the statistic T to:

$$T = \frac{\left(\overline{X}_1 - \overline{X}_2\right)\sqrt{n/2}}{\sigma}$$

If we again assume that the standard deviation of both sets of measurements is 5 and we use the 95th percentile values of the t distribution taken from the previous clause. We can then determine the number of measurements required to reject the null hypothesis for a 1 dB difference in the means at a 95% confidence level:

$$T = \frac{(1)\sqrt{n/2}}{5} \quad \text{must be greater than } 1.66 = t_{.95}(120)$$

which implies:

$n \geq 2\ (1.66)^2\ 5^2 \geq 138$ measurements for each set of data.

Appendix III

Statistical techniques for use on INMD threshold calculations

III.1 Introduction

Thresholding applied to INMD data involves calculating the percentage of measurements above or below a threshold for a sample of measurements. This appendix provides information on the theoretical approach used to determine suitable sample sizes.

III.2 Theoretical approach

With the aim of developing a statistical approach to the matter, a random process X can be considered associated to the measurement of one of the parameters amongst ASL, PNL, ED, EPL. With good approximation, well confirmed in practice, the ith realization of X has a *Normal* distribution:

$$x_i \sim \mathrm{N}\left(\mu, \sigma^2\right) \qquad \text{(III-1)}$$

where μ and σ are the mean and the standard deviation respectively.

If the mean and the standard deviation are known, then the percentage of measurements above a fixed threshold x_{th}, as depicted in Figure III.1, can be easily calculated using the following formula:

$$p_e = \Pr\{x_i \geq x_{th}\} = 1 - \Phi\left(\frac{x_{th} - \mu}{\sigma}\right) \qquad \text{(III-2)}$$

where:

$$\Phi(z) = \Pr(Z \le z), \quad Z \sim \mathrm{N}(0,1) \tag{III-3}$$

It should be noted that the same holds true, *mutatis mutandis*, for the percentage of measurements below a fixed threshold, due to the symmetry of the *Normal* distribution.

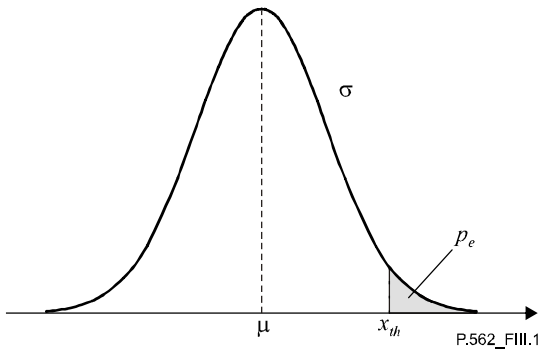

Figure III.1/P.562 – Normal distribution of the measurements

In practical applications, however, the mean μ is not *a priori* known while for the standard deviation σ typical values are available. This implies that the percentage p_e is in its turn not known, thus requiring its estimate through the measurements.

Considering a sample of measurements, whose size n is related to the desired accuracy and is to be determined, a rather simple estimate of p_e, say $\hat{p}_e$, can be calculated as the ratio of the number of measurements that exceed the fixed threshold x_{th} to the sample size n.

Given the required accuracy and the associated confidence interval for $\hat{p}_e$, the sample size n to be used can be determined in the way hereafter presented.

It is emphasized that the following approach is applicable in general, without necessarily making the assumption that the measurements can be approximated by a *Normal* distribution. Equation III-2 can, therefore, be generalized as follows, as concerns the percentage of measurements above a fixed threshold x_{th}:

$$p_e = \Pr\{x_i \ge x_{th}\} \tag{III-4}$$

If x_i is the ith realization of the measurement process, and y_i is obtained using Equation III-5,

$$y_i = \begin{cases} 1 & if\ x_i \ge x_{th} \\ 0 & if\ x_i < x_{th} \end{cases} \tag{III-5}$$

then it follows that y_i is distributed as a *Bernoulli* random variable with unknown success probability equal to p_e:

$$y_i \sim \mathrm{B}(1, p_e) \tag{III-6}$$

Applying Equation III-5 to a sample of n measurements, $x_1, x_2, \ldots, x_n$, the correspondent random variables $y_1, y_2, \ldots, y_n$ can be obtained, and the best maximum likelihood unbiased estimator for p_e turns out to be:

$$\hat{p}_e = \frac{\sum_{i=1}^{n} y_1}{n} \tag{III-7}$$

where the numerator corresponds to the number of measurements exceeding the fixed threshold,

$$\hat{y} = \sum_{i=1}^{n} y_i \tag{III-8}$$

If the measurements $x_1, x_2, \ldots, x_n$ can be considered as being independent and identically distributed (i.i.d), then $\hat{y}$ can be described by a *Binomial* random variable having order equal to n and unknown success probability equal to p_e:

$$\hat{y} \sim \mathrm{B}(n, p_e) \Rightarrow \Pr\{\hat{y} = t\} = \binom{n}{t} \cdot p_e^t \cdot (1 - p_e)^{n-t} \qquad \text{with } t = 0,1,\ldots,n \tag{III-9}$$

If an estimate of p_e with a 95% confidence interval of $\pm\Delta_p$ is desired, then the minimum sample size n_{min} must satisfy the following equation:

$$\Pr\{\hat{p}_e - \Delta_p \le p_e \le \hat{p}_e + \Delta_p\} = 0.95 \tag{III-10}$$

that, utilizing Equation III-9 and exploiting some basic statistical properties provides the results reported in Figure III.2.

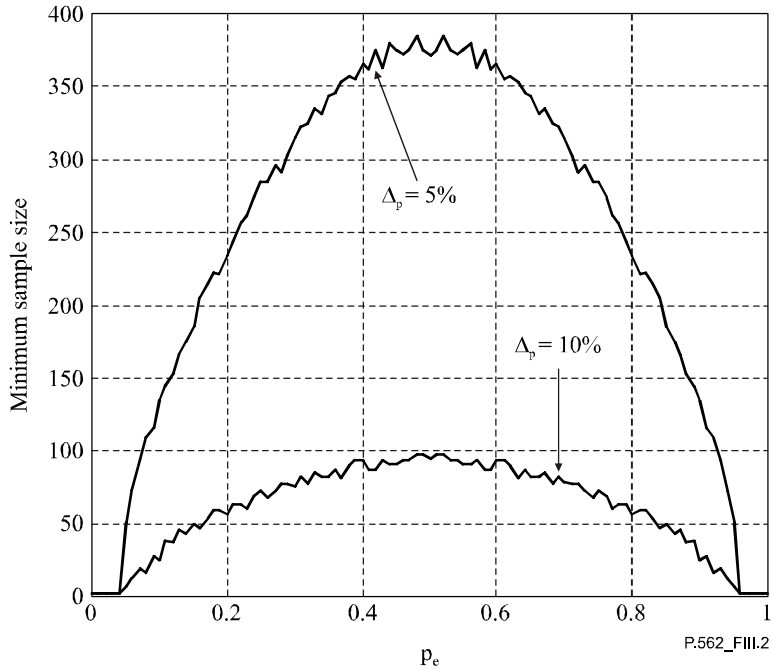

Figure III.2/P.562 – Minimum sample size for a 95% confidence interval of $\pm\Delta_p$

III.3 Approximation

The theoretical approach presented in Equation III-2 can be simplified introducing some approximations.

It is known from the literature, that a *Binomial* random variable can be approximated by a *Normal* distribution when the order is sufficiently high and the success probability is not too close to 0 or 1.

Based on this consideration, and under the same assumptions that have led to Equation III-9, we obtain:

$$\hat{y} \approx \mathrm{N}(np_e, np_e(1 - p_e)) \tag{III-11}$$

and, as a consequence:

$$\hat{p}_e = \frac{\hat{y}}{n} \approx \mathrm{N}\left(p_e, p_e\left(1-p_e\right)/n\right) \tag{III-12}$$

The minimum sample size n_{min} needed to generate an α percent confidence interval of $\pm\Delta_p$ can now be determined using the following formula:

$$\frac{\Delta_p}{\sqrt{\frac{p_e\left(1-p_e\right)}{n_{\min}}}} = Z_\alpha \tag{III-13}$$

where Z_α is the normal deviate for α, which leads to:

$$n_{\min} = p_e\left(1-p_e\right)\cdot\frac{Z_\alpha^2}{\Delta_p^2} \tag{III-14}$$

As an example, Figure III.3 shows the result of the approximation for α = 95% and Δ_p = 5%.

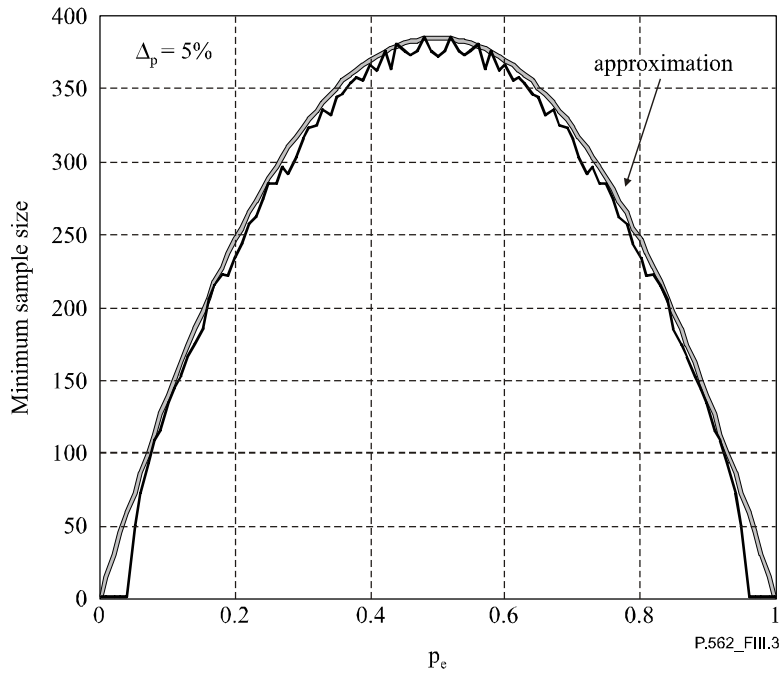

Figure III.3/P.562 – Minimum sample size for a 95% confidence interval of ±5%: approximate result

Since, in practice, no *a priori* knowledge is available as concerns the percentage of measurements of a parameter which are above (or below) a fixed threshold, the worst-case associated with p_e = 1/2 (referring to Equation III-14) should be considered for the determination of the suitable sample size.

The recommended formula is, therefore, the following:

$$\textit{Minimum Sample Size} = \frac{Z_\alpha^2}{4\Delta_p^2} \tag{III-15}$$

where Δ_p is the required accuracy and Z_α is the normal deviate for α.

SERIES OF ITU-T RECOMMENDATIONS

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

Printed in Switzerland
Geneva, 2005